OUTLINE

OF

THE OPERATIONAL NUMERICAL WEATHER PREDICTION

AT

THE JAPAN METEOROLOGICAL AGENCY

January 2024

JAPAN METEOROLOGICAL AGENCY

Contents

Preface

The Japan Meteorological Agency (JMA) began numerical weather prediction (NWP) in June 1959 after years of extensive research by Japan's Numerical Weather Prediction Group, following on from similar pioneering efforts by the Swedish Meteorological and Hydrological Institute (SMHI) in September 1954 and the US Weather Bureau in May 1955. Subsequent NWP development has advanced well, with JMA and other operational NWP centers benefitting from enhanced comprehension of meteorological phenomena, improved modeling, increased computing capacity, efficient telecommunications, and better observing systems (especially in the fields of meteorology and satellite-based earth observation). NWP leads itself to a wide range of applications within JMA's scope of operation.

The report prior to this one served as an appendix to the WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research, work covering the period from 2002 to 2022, and was published in 2023 as independent content with the cessation of WMO Technical Progress Report provision. The current report covers the main content of the WMO Technical Progress Report, and is up to date as of March 2023 (or more recent where noted).

The first chapter summarizes the configurations and specifications of the current JMA computer system, followed a description of the related operational suite and job management system.

The second chapter describes Global Analysis, Meso-scale Analysis and Local Analysis as major data assimilation systems for atmospheric fields incorporating observation data, along with data assimilation systems for snow depth and soil moisture, the JMA Climate Data Assimilation System and other related considerations.

The third chapter describes NWP models used for very short-range prediction of meso-scale disturbances and short-/medium-range prediction of synoptic-scale disturbances. The regional deterministic and ensemble prediction system is used especially for heavy rainfall associated with mesoscale convective system, which causes extreme disaster conditions in Japan. The global model is used in ensemble prediction systems for typhoon forecasts and weekly/monthly predictions, and the coupled ocean-atmosphere model is employed for seasonal and El Niño forecasts. Atmospheric transport models are applied to prediction for transport of trace elements such as radioactive materials, Kosa (Aeolian Dust), ozone and volcanic ash for the output of environmental information.

The fourth chapter outlines various applications for NWP products such as weather charts, gridded data, very-short-range precipitation forecasting, half-hourly analysis of wind and temperature, guidance for short-range forecasting, and data relating to aviation services, ensemble prediction and atmospheric angular momentum functions.

The final chapter discusses modeling for ocean waves, storm surges and oil spill prediction along with systems for sea surface temperature analysis and ocean data assimilation.

JMA remains committed to its efforts for enhanced NWP accuracy. Against this background, the Numerical Prediction Development Center was established in 2020 in Tsukuba (the location of the Meteorological Research Institute) for integrated promotion. JMA's NWP Strategic Plan Toward 2030 (https://www.jma.go.jp/jma/en/Publications/JMA_NWP_Strategic_Plan_Toward_2030.pdf), also promotes NWP technology within the context of a recent intensification of natural disasters and societal requirements, defining the direction of NWP development at JMA over the coming years. Updates are provided at https://www.jma.go.jp/jma/en/Activities/nwp.html.

SATO Yoshiaki

Director
Numerical Prediction Division
Japan Meteorological Agency

Chapter 1

Computer System

1.1 Introduction

The Japan Meteorological Agency (JMA) installed its first-generation computer (IBM 704) to run an operational numerical weather prediction model in March 1959. Since then, the computer system at JMA has been repeatedly upgraded, and the current system (Cray XC50) was completed in June 2018 as the tenth-generation computer. Figure 1.1.1 shows the history of computers at JMA, their peak performance, and a change in peak performance calculated using Moore's law[1] from the first computer (IBM 704). The peak performance of the second (HITAC 5020), the third (HITAC 8800), and the eighth (HITACHI SR11000) computers at the beginning of their implementation was almost the same as that projected using Moore's law, while it was lower during the 1980s, 1990s, and the first half of the 2000s. Recent growth is faster and the peak performance of the current computer is higher than the projection.

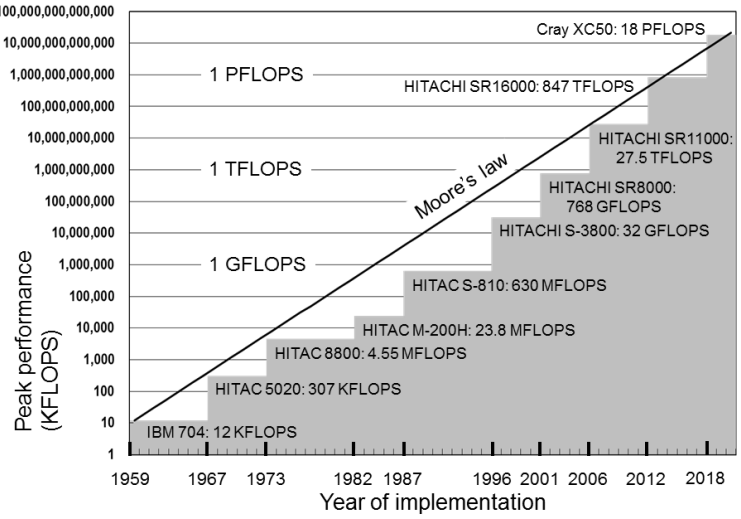


Figure 1.1.1: History of computers used at JMA and their peak performance. The line "Moore's law" represents the projection of peak performance using Moore's law from the first computer (IBM 704).

[1]The term "Moore's law" has many formulations. Here we refer to exponential growth of peak performance which doubles every 18 months.

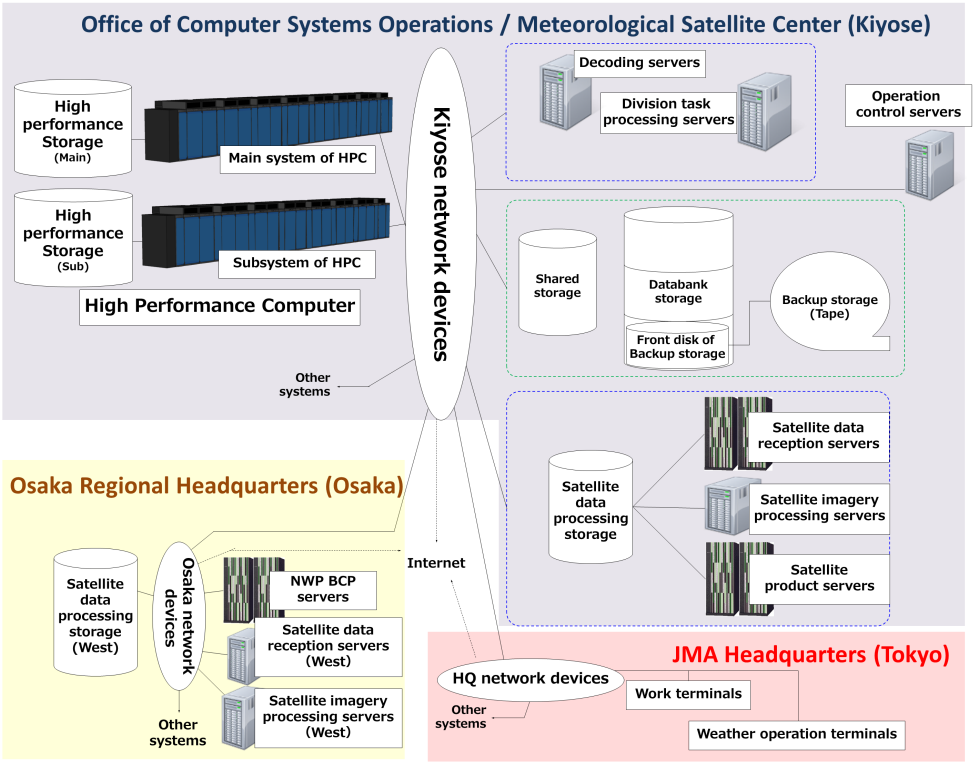


Figure 1.2.1: Schematic illustration of computer system

In this chapter, Section 1.2 briefly describes the configurations and specifications of the current computer system at JMA. Section 1.3 outlines the operational suite and the operational job management system on the current computer system.

1.2 System Configurations and Specifications

1.2.1 Overview

Figure 1.2.1 illustrates major components of the computer system at JMA including Cray XC50 high performance computers, server computers, storages, terminals, and networks. The system has been in operation since 5 June 2018. Most of the computing facilities are located at the Office of Computer Systems Operations and the Meteorological Satellite Center in Kiyose 24 km west of JMA's central-Tokyo HQ, and some servers are located at the Osaka Regional Headquarters for business continuity planning (BCP). A wide area network (WAN) links the Kiyose, HQ and Osaka sites. The specifications of the high-performance computers and server computers are summarized in Table 1.2.1, Table 1.2.2 and Table 1.2.3, respectively.

1.2.2 High Performance Computer

Two independent systems called a main system and a subsystem of a Cray XC50 high performance computer with the same specifications are installed at the Kiyose site. The main system usually runs operational numerical weather prediction jobs, while the subsystem usually runs development jobs. However, in case the main system is under maintenance or out of order, the subsystem runs operational jobs to make the system stable for operational use.

Table 1.2.1: Specifications of high performance computers

Computer	Cray XC50
Number of systems	2
Computational nodes	
Processor, clock frequency	Intel Xeon Platinum 8160, 2.1 GHz
Cores per processor	24
Cores per logical node	48
Logical nodes per system	2741(ESM), 75(MAMU), 8(Tier2), 16(spare)
Peak performance per logical node	3.2256 TFLOPS
Peak performance per system	9,083 TFLOPS
Memory per logical node	96 GiB
Memory per system	264 TiB
Operating system	Cray Linux Environment 6.0(ESM), SUSE 12.2(MAMU,Tier2)
I/O nodes	
Processor, clock frequency	Intel Xeon E5-2699v4, 2.2 GHz
Cores per processor	22
Cores per logical node	22
Logical nodes per system	2(SDB), 4(PBS-MOM), 2(boot), 2(router), 7(network), 15(LNET), 6(data sync), 2(login gateway)
Memory per logical node	128 GiB
Operating system	SUSE 12.2
Login servers	Dell PowerEdge Server
Processor, clock frequency	Intel Xeon Gold 6148, 2.4 GHz
Cores per processor	20
Cores per logical node	40
Number of servers per system	4
Memory per logical node	768 GiB
Operating system	SUSE 12.2

The Cray XC50 consists of computational nodes, I/O nodes, and login servers.

Each of its computational nodes has two sockets for Intel Xeon Platinum 8160 processors with a clock frequency of 2.1 GHz. One socket of the Xeon processor houses a multi-core chip with 24 separate cores, making $2 \times 24 = 48$ cores in each logical node. The theoretical performance per logical node is 3.2256 TFLOPS, and the total memory capacity is 96 GiB per logical node[2]. The computational nodes are ESM[3](2741), MAMU[4](75), Tier2[5](8), and spare(16) types. The theoretical performance per system is 9,083 TFLOPS for only ESM and MAMU nodes. Each ESM node runs the CLE(Cray Linux Environment) 6.0 operating system, and each MAMU and Tier2 node runs SUSE Linux Enterprise Server 12.2 independently. The inter-node communication rate between each node and the hub processor is 14 GB/s for one-way communication.

The I/O nodes consist of an Intel Xeon E5-2699v4 (2.2 GHz) processor. The types and numbers of these nodes are SDB[6](2), PBS-MOM[7](4), boot[8](2), router[9](2), network[10](7), LNET[11](15), data sync[12](6), and login gateway[13](2).

The login system involves four Dell PowerEdge servers with two Intel Xeon Gold 6148 (2.4 GHz) processors. The operating system for both I/O nodes and login servers is SUSE Linux Enterprise Server 12.2.

[2]The International Electrotechnical Commission approved names and symbols for the power of $2^{10} = 1,024$ instead of 1,000 for prefixes of units. Symbols such as GiB or TiB refer to the former. In contrast, symbols such as GB or TB mean the latter.

[3]Extreme Scalability Mode nodes. Used for high performance Massively Parallel Processing(MPP) runs.

[4]Multiple Applications Multiple User nodes. Used for smaller applications.

[5]Distribution of computational environments to ESM and MAMU.

[6]Service DataBase node with PBS installation.

[7]PBS Mom daemon applied for ESM node.

[8]Used for boot step.

[9]Used for connection with surveillance network.

[10]Used for connection with storage network and servers.

[11]Lustre NETwork node. Used for connection with Lustre high performance storage.

[12]Used for connection with main system and subsystem.

[13]Used for connection with login servers.

Table 1.2.2: Specifications of server computers at Kiyose

	Satellite data reception servers	Satellite imagery processing servers	Satellite product servers
Computer	HPE ProLiant DL360 Gen9	HPE ProLiant DL580 Gen9	HPE ProLiant DL380 Gen9
Processor, clock frequency	Intel Xeon E5-2620v3, 2.4 GHz	Intel Xeon E7-8880v3, 2.3 GHz	Intel Xeon E5-2670v3, 2.3 GHz
Cores per processor	6	18	12
Cores per server	12	72	24
Number of servers	5	8	10
Memory per server	64 GiB	256 GiB	192 GiB
Operating system	RHEL 7.3	RHEL 7.3	RHEL 7.3

	Operation control servers	Division task processing servers	Decoding servers
Computer	HITACHI HA8000 RS210AN1	HPE ProLiant DL580 Gen9	HPE ProLiant DL580 Gen9
Processor, clock frequency	Intel Xeon E5-2640v3, 2.6 GHz	Intel Xeon E7-8880v3, 2.3 GHz	Intel Xeon E7-8860v3, 2.2 GHz
Cores per processor	8	18	16
Cores per server	16	72	64
Number of servers	8	12	2
Memory per server	32 GiB	128 GiB	256 GiB
Operating system	RHEL 7.3	RHEL 7.3	RHEL 7.3

The main system and subsystem have high-performance storage configured with a Lustre file system, and have capacities of 1.6x3 PB each. Every time an operational job running on the main system is completed, its output files are copied to the high-performance storage on the subsystem to ensure that the subsystem is ready to run with further operational jobs if operation is switched to it.

1.2.3 Server and Terminal Computers at Kiyose

A number of server computers are used for various tasks, such as processing and decoding of observational data, weather chart analysis and operational suite management.

The satellite data reception servers, satellite imagery processing servers and satellite product servers are used for automatic processing of various kinds of satellite observation data. The five satellite data reception server are HPE ProLiant DL360 Gen9 units with two Intel Xeon E5-2620v3 (2.4 GHz) processors. The eight satellite imagery processing servers are HPE ProLiant DL580 Gen9 units with four Intel Xeon E7-8880v3 (2.3 GHz) processors. The ten satellite product servers are HPE ProLiant DL380 Gen9 units with two Intel Xeon E5-2670v3 (2.3 GHz) processors.

The eight operation control servers used for control of operational suite job groups are HITACHI HA8000 units with two Intel Xeon E5-2640v3 (2.6 GHz) processors.

The division task processing servers are used for weather chart analysis and small operational jobs that are transaction-intensive rather than compute-intensive. The 12 servers of this type are HPE ProLiant DL580 Gen9 units with four Intel Xeon E7-8880v3 (2.3 GHz) processors.

The two servers used for decoding observational data jobs are HPE ProLiant DL580 Gen9 units with four Intel Xeon E7-8860v3 (2.2 GHz) processors.

Other server computers are also used to that manage the operational suite for numerical weather prediction, satellite data processing and other jobs. Server and terminal computers are additionally used to monitor and manage the computer system.

1.2.4 Mass Storage System

Shared, data bank and backup storage systems are used to share data between high-performance computers and server computers.

The shared and databank storage systems are used for jobs running on high-performance computers or server computers. Configuration involves an IBM Spectrum Scale (ISS) file system with RAID 6 magnetic

disks[14] Shared storage comprises three units with a total capacity of 6PB, and databank storage comprises three units with a total capacity of 25PB, one of which is used as a front disk for backup storage.

The backup storage system is used for long-term archiving. It automatically makes backup copies from the front disk of the data bank storage system, and consists of a tape library and four management servers. Its total capacity is about 80 PB.[15]

1.2.5 Networks

The Kiyose network connects the high-performance computers, server computers and other network/server elements in the computer system described above.

The storage network connects the high performance computers, server computers, shared storage system, databank storage system, and backup storage system.

Users at HQ remotely log in to computers at the Kiyose site through a WAN consisting of three independent links with transfer speeds of 100 Mbps, 100 Mbps and 1 Gbps (best effort), respectively. The two 100 Mbps links are used for operational jobs, while the 1 Gbps link is used for development jobs. All network equipment is redundantly configured to prevent single failures from causing catastrophic interruption.

The Osaka site is also connected to the Kiyose site through a WAN with two 100 Mbps links.

1.2.6 Server and Terminal Computers at Osaka

Equipment is located in Osaka for NWP BCP operations and redundancy processing of satellite data. There are two HPC ProLiant DL360 Gen9 servers with two Intel Xeon E5-2680v3 (2.5 GHz) processors, which are used for NWP BCP operations.[16] The satellite data reception servers(West) and satellite imagery processing servers(West) are used for processing of satellite observations data in Osaka. The two satellite data reception servers(West) are HPE ProLiant DL360 Gen9 units with two Intel Xeon E5-2620v3 (2.4 GHz) processors. The four satellite imagery processing servers(West) are HPE ProLiant DL360 Gen9 units with two Intel Xeon E5-2698v3 (2.3 GHz) processors.

Table 1.2.3: Specifications of Osaka server computers

	NWP BCP servers	Satellite data reception severs(West)	Satellite imagery processing servers(West)
Computer	HPE ProLiant DL360 Gen9	HPE ProLiant DL360 Gen9	HPE ProLiant DL360 Gen9
Processor, clock frequency	Intel Xeon E5-2680v3, 2.5 GHz	Intel Xeon E5-2620v3, 2.4 GHz	Intel Xeon E5-2698v3, 2.3 GHz
Cores per processor	12	6	16
Cores per server	24	12	32
Number of servers	2	2	4
Memory per server	256 GiB	64 GiB	128 GiB
Operating system	RHEL 7.3	RHEL 7.3	RHEL 7.3

1.3 Operational Aspects

1.3.1 Operational Suite

The JMA operational suite described in later chapters consists of about 80 job groups, including global analysis and global forecasting, with a total of around 20,600 jobs per day. All jobs are submitted from the Routine Operation and Scheduling Environment (ROSE). There are approximately 4,000 and 17,500 constant and variable datasets, respectively.

[14] RAID stands for redundant array of independent disks or redundant array of inexpensive disks. In particular, RAID 6 utilizes block-level striping with double distributed parity and provides fault tolerance for two drive failures.

[15] The total capacity depends on the volume of the tape cartridge. A capacity of 80 PB is estimated with a 10-TB tape cartridge.

[16] Current NWP BCP operations involve online acquisition of gridded data from overseas NWP center sources and processing to create JMA's product format.

1.3.2 ROSE:Job Scheduler

ROSE is a job flow control computer program that automatically controls execution of all operational jobs. Following on from the start of its development in 2008, it has been used by JMA since 2009 in numerical prediction model development environments and other areas. Based on the current system, the program was adopted for operational NWP control.

ROSE is installed on operational control servers, and is used to control all operational jobs based on submission to PBS[17].

1.3.3 RENS:Operational Job Management System

There are complicated dependencies between jobs in a job group and between input and output datasets. To manage a vast number of operational jobs and datasets systematically and assure that jobs run correctly without human error, JMA developed the comprehensive RENS[18] resource using database management systems (DBMSs).

All job information, input/output datasets, and executables are registered in RENS. Dependencies between these elements can be checked using utility programs.

RENS is comprised of four file types, two DBMSs, and several utility programs to allow registration of information, checking of consistency and other tasks as detailed below.

- Files

  Registration form: Information about job groups, jobs, datasets, executables, and so on. A registration form is submitted when jobs are added or deleted, datasets or executables are updated, or the configurations of job groups or jobs are modified.

  Job definition file: Information about a job group and jobs within the job group such as the job group name, the job name, the schedule (time to run), the order of job groups and jobs (preceding job groups and jobs), and computational resources required (the PBS job class, the number of nodes, the computational time).

  Job control language: Information about executables such as a shell script, a ruby script, an awk script and a load module, and input and output datasets used in each job. A job control language file is converted into a shell script using a utility program to be submitted to PBS.

  Program build file-format: Information about source files, object modules, libraries, options for compilation, and so on. A program build file-format is converted into a makefile using a utility program to compile load modules.

- DBMSs

  DBMS for registration: Information from the above four files is registered using utility programs.

  DBMS for job management: Information from the DMBS for registration is stored and this information is used by job schedulers.

When a job control language file is converted into a shell script, the following procedures are made:

- Existence test: A shell script tests the existence of all non-optional input datasets at the beginning in order to avoid wasting time if the preceding job failed.

- Quasi-atomic output: Every step of a job calling an executable creates output files with temporary names at first and renames them to final names when the step successfully terminates.

The development of the RENS was started in 2004 on the seventh computer system and installed in the operational system in 2006 when the eighth computer system was implemented. The number of man-made errors after the inclusion of this management system was drastically reduced to about one sixth of that before the adoption.

[17] Portable Batch System (the computer program used to perform job scheduling)

[18] RENS : Routine Environment for Numerical weather prediction System.

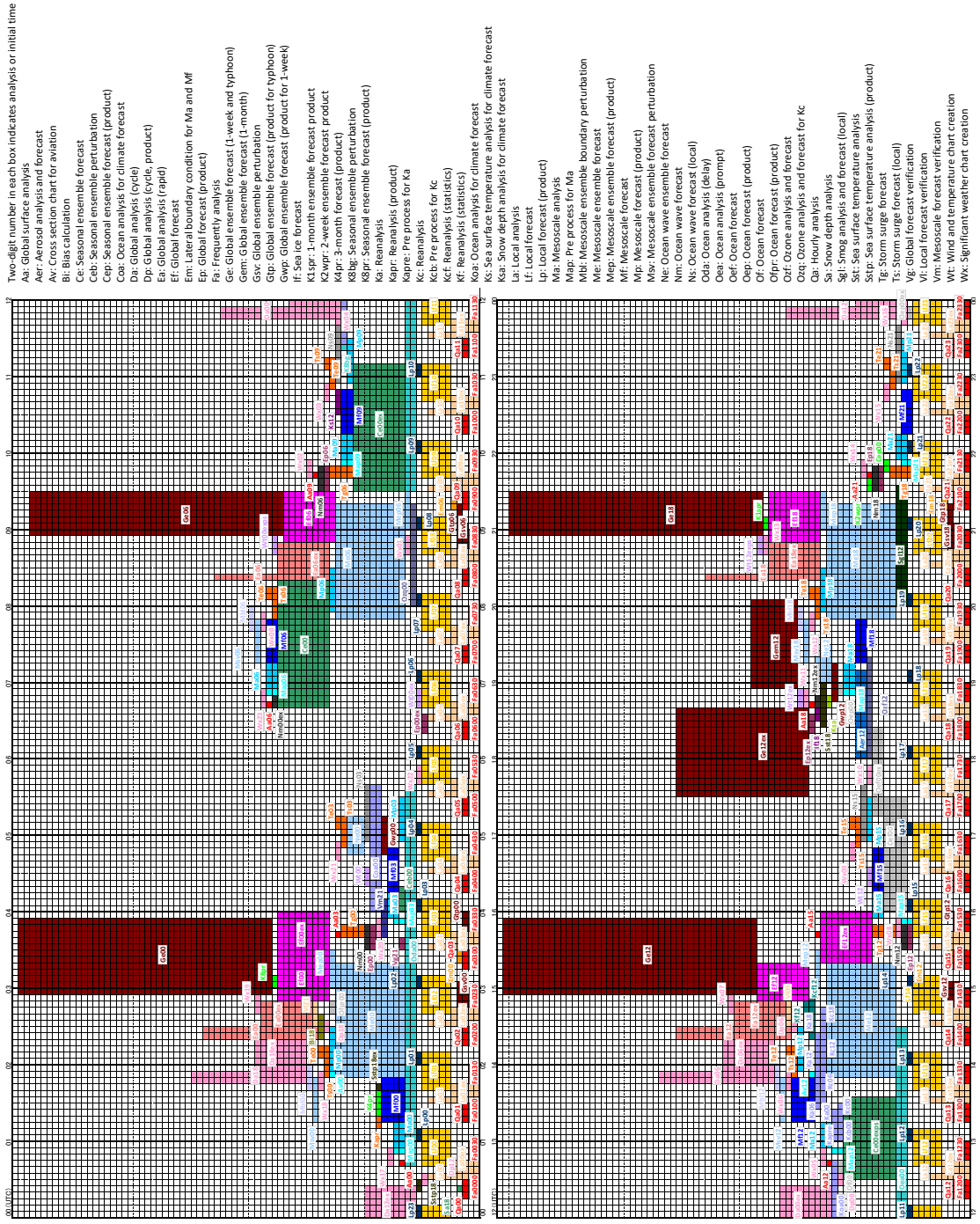


Figure 1.3.1: Daily schedule of the operational suite running on the main system of the high-performance computer as of March 2023. The height and width of each box indicate the approximate number of nodes and the time range, respectively.

Chapter 2

Data Assimilation Systems

2.1 Summary

Three kinds of major data assimilation systems for the analysis for atmospheric fields are operated at JMA: Global Analysis (GA), Meso-scale Analysis (MA) and Local Analysis (LA). Specifications of the JMA data assimilation systems are summarized in Table 2.1.1, Table 2.1.2 and Table 2.1.3. All the analyses are performed by using the procedures shown in Figure 2.1.1.

The following is a brief description of the major components of the analysis systems.

1. Observational data are received from the WMO Information System (WIS) including Global Telecommunication System (GTS), Internet and dedicated networks. The data are decoded according to their code forms. If typhoons exist in the western North Pacific, typhoon bogus profiles are created.

2. Various pre-analysis procedures, such as quality control, data selection and bias correction, are applied to the decoded observational data. In the pre-analysis process, first guess fields retrieved from forecast models are used as a reference of the present atmospheric conditions.

3. The hybrid four-dimensional variational method using the Local Ensemble Transform Kalman Filter is adopted in Global Analysis, while the four-dimensional variational method is adopted in Meso-scale Analysis. To reduce the computational cost, the analysis increment is calculated using a coarser-resolution inner model in each method. The resolution of these analysis type is the same as that of the corresponding forecast models.

4. Local Analysis involves a three-hour cycle based on the hybrid three-dimensional variational method. Its resolution is coarser than that of the corresponding forecast model.

The atmospheric fields analyzed from the data assimilation systems are used as initial conditions of forecast models. First guess fields and boundary conditions of data assimilation systems are provided from forecast models as shown in Figure 2.1.2.

Sea surface temperature fields (see Section 5.2), snow depth fields (see Section 2.8) and soil moisture fields (see Section 2.9) are also analyzed every day.

The Japanese Reanalysis for Three Quarters of a Century project covers the period from September 1947 to the present day on a near-real-time basis (see Section 2.11).

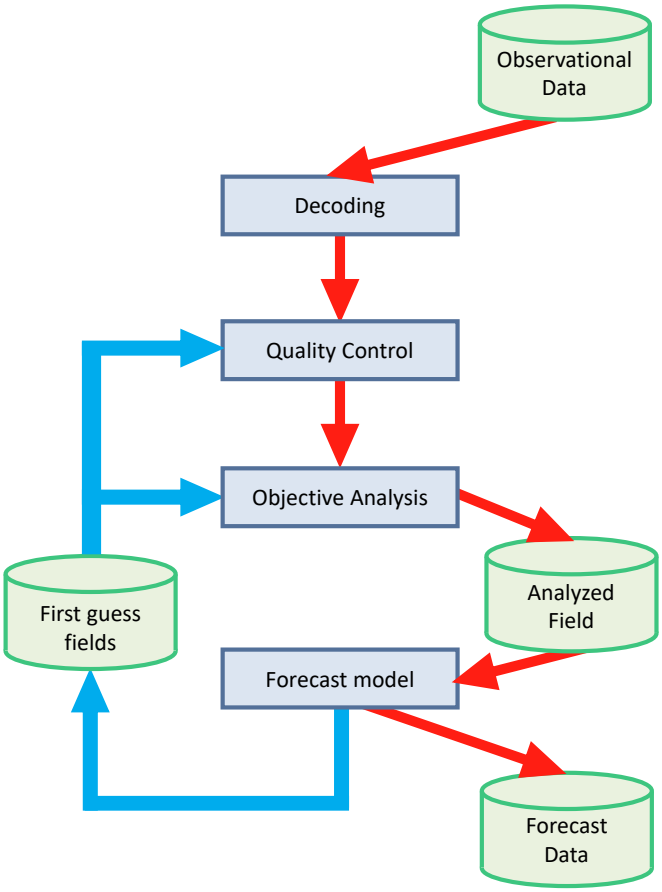


Figure 2.1.1: Major functional components and data flow in the JMA data assimilation system.

Table 2.1.1: Specifications of 4D-Var in Global Analysis (GA)

Analysis time	00, 06, 12 and 18 UTC
Analysis scheme	Incremental hybrid 4D-Var using LETKF
Data cut-off time	2 hours and 20 minutes for early run analysis at 00, 06, 12 and 18 UTC 11 hours and 50 minutes for cycle run analysis at 00 and 12 UTC 7 hours and 50 minutes for cycle run analysis at 06 and 18 UTC
First guess	6-hour forecast by the GSM
Domain configuration (Outer step)	Globe TQ959, Reduced Gaussian grid, roughly equivalent to 0.125° (13 km) [2880 (tropic) – 64 (polar)] × 1440
(Inner step)	TL319, Reduced Gaussian grid, roughly equivalent to 0.5625° (55 km) [640 (tropic) – 60 (polar)] × 320
Vertical coordinate	σ-p hybrid
Vertical levels	128 forecast model levels up to 0.01 hPa + surface
Analysis variables	Wind, surface pressure, specific humidity and temperature
Observation (as of 31 March 2023)	SYNOP, METAR, SHIP, BUOY, TEMP, PILOT, Wind Profiler, AIREP, AMDAR, Typhoon Bogus; atmospheric motion vectors (AMVs) from Himawari-9, GOES-16, Meteosat-9 and Metop-[B, C]; polar AMVs from Aqua/MODIS, NOAA/AVHRR, Metop/AVHRR, Suomi-NPP/VIIRS and NOAA-20/VIIRS; LEO-GEO AMVs; ocean surface wind from Metop-[B, C]/ASCAT; radiances from NOAA-[15, 18, 19]/AMSU-A, Metop-[B, C]/AMSU-A, DMSP-F[17, 18]/SSMIS, Suomi-NPP/ATMS, NOAA-20/ATMS, GCOM-W/AMSR2, GPM-core/GMI, Metop-[B, C]/IASI, Suomi-NPP/CrIS, NOAA-20/CrIS, FY-3C/MWHS2, NOAA-19/MHS and Metop-[B, C]/MHS; clear sky radiances from the water vapor channels (WV-CSRs) of Himawari-9, GOES-16 and Meteosat-9; GNSS RO bending angle data from Metop-B/GRAS and TerraSAR-X/IGOR; zenith total delay data from ground-based GNSS
Assimilation window	6 hours

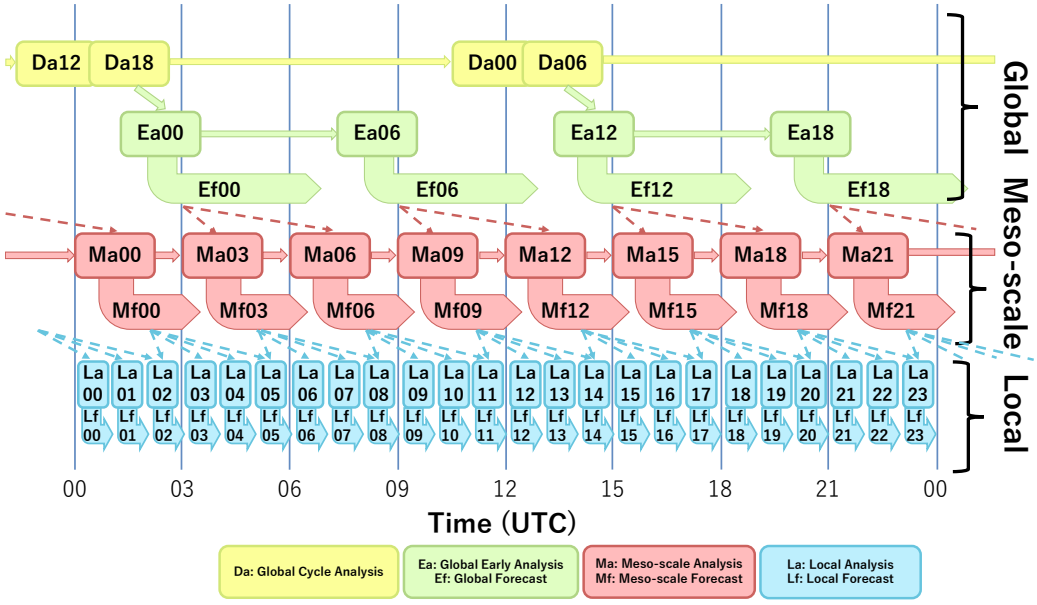


Figure 2.1.2: Main flow of JMA data assimilation systems. The first-guess and boundary conditions for Local Analysis are obtained from the latest MSM output.

Table 2.1.2: Specifications of the Mesoscale Analysis (MA)

Analysis time	00, 03, 06, 09, 12, 15, 18 and 21 UTC
Analysis scheme	Incremental 4D-Var using a nonlinear forward model in the inner step with low resolution
Data cut-off time	50 minutes for analysis at 00, 03, 06, 09, 12, 15, 18 and 21 UTC
First guess	3-hour forecast produced by ASUCA
Domain configuration (Outer step) (Inner step)	Japan and its surrounding area Lambert projection: 5 km at 60°N and 30°N, 817 × 661 Grid point (1, 1) is at the northwest corner of the domain. Grid point (565, 445) is at 140°E, 30°N Lambert projection: 15 km at 60°N and 30°N, 273 × 221 Grid point (1, 1) is at the northwest corner of the domain. Grid point (189, 149) is at 140°E, 30°N
Vertical coordinate	z-$z^\star$ hybrid
Vertical levels	(Outer step) 96 levels up to 37.5 km (Inner step) 48 levels up to 37.5 km
Analysis variables	Wind, potential temperature, surface pressure, pseudo-relative humidity, soil temperature and soil volumetric water content
Observations (as of 31 March 2023)	SYNOP, SHIP, BUOY, TEMP, PILOT, Wind Profiler, Weather Doppler radar (radial velocity, reflectivity), AIREP, AMDAR, Typhoon Bogus; AMVs from Himawari-9; ocean surface wind from Metop-[B, C]/ASCAT; radiances from NOAA-[15, 18, 19]/AMSU-A, NOAA-19/MHS, Metop-[B, C]/AMSU-A, Metop-[B, C]/MHS, DMSP-F[17, 18]/SSMIS, Suomi-NPP/ATMS, NOAA-20/ATMS, FY-3C/MWHS2, GCOM-W/AMSR2, GPM-core/GMI, Metop-[B, C]/IASI, Suomi-NPP/CrIS and NOAA-20/CrIS; clear sky radiances from the water vapor channels (WV-CSRs) of Himawari-9; Radar/Raingauge-Analyzed Precipitation; precipitation retrievals from DMSP-F[17, 18]/SSMIS, GCOM-W/AMSR2 and GPM-core/GMI; GPM-core/DPR; GNSS RO refractivity data from Metop-B/GRAS, TerraSAR-X/IGOR and TanDEM-X/IGOR; Total Precipitable Water Vapor from ground-based and shipborne GNSS
Assimilation window	3 hours

Table 2.1.3: Specifications of the Local Analysis (LA)

Analysis time	00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 and 23 UTC
Analysis scheme	The three-hour analysis cycle repeats hourly assimilation with hybrid 3D-Var and one-hour forecasts
Data cut-off time	30 minutes for analysis at 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 and 23 UTC
First guess	Initial fields produced by the latest MSM
Domain configuration	Japan and its surrounding area Lambert projection: 5 km at 60°N and 30°N, 633 × 521 Grid point (1, 1) is at the northwest corner of the domain. Grid point (449, 361) is at 140°E, 30°N
Vertical coordinate	z-$z^\star$ hybrid
Vertical levels	48 levels up to 21.8 km
Analysis variables	Wind, potential temperature, surface pressure, pseudo-relative humidity, skin temperature, ground temperature and soil moisture
Observations (as of 31 March 2023)	SYNOP, SHIP, BUOY, AMeDAS, TEMP, PILOT, Wind Profiler, Weather Doppler radar (radial velocity, reflectivity), AIREP, AMDAR; AMVs from Himawari-9; ocean surface wind from Metop-[B, C]/ASCAT; radiances from NOAA-[15, 18, 19]/AMSU-A, NOAA-19/MHS, Metop-[B, C]/AMSU-A, Metop-[B, C]/MHS, DMSP-F[17, 18]/SSMIS, Suomi-NPP/ATMS, NOAA-20/ATMS, FY-3C/MWHS-2, GCOM-W/AMSR2, GPM-core/GMI, Metop-[B, C]/IASI and NOAA-20/CrIS; clear sky radiances from the water vapor channels (WV-CSRs) of Himawari-9; soil moisture from GCOM-W/AMSR2 and Metop-B/ASCAT; Total Precipitable Water Vapor from ground-based and shipborne GNSS

2.2 Observation Data

2.2.1 Summary of Observation Data Used in Analysis

A variety of observations are utilized in JMA's current NWP systems. Table 2.2.1 summarizes the types used and the input parameters for the objective analysis systems, as of 31 March, 2023. Additional information on each observation type is provided in the following subsection.

2.2.2 Supplemental Information for Used Observation

2.2.2.1 SYNOP

SYNOP is a numerical code used for reporting surface observations at land stations. About 27,000 reports are produced every six hours.

2.2.2.2 AMeDAS

AMeDAS (the Automated Meteorological Data Acquisition System) is a JMA land surface automated network to observe near surface temperature and wind at about 930 stations in Japan at 10 minute intervals. JMA is proceeding with a plan for an upgrade of AMeDAS stations to attach humidity sensors up to FY 2026. Humidity observations started at 274 AMeDAS stations by FY 2022.

2.2.2.3 METAR

METAR is a numerical code used for reporting aerodrome weather information. Around 45,000 reports are produced every six hours.

2.2.2.4 SHIP

SHIP is a numerical code used for reporting surface observations performed at sea stations such as ships, oil rigs and moored buoys anchored at fixed locations. Around 8,000 reports are produced every six hours.

2.2.2.5 BUOY

BUOY is a numerical code used for reporting surface observations performed by drifting buoys. Around 14,000 reports are produced every six hours.

2.2.2.6 TEMP

TEMP is a numerical code used for reporting upper-level pressure, temperature, humidity and wind observations performed by radiosondes. Upper air observations are usually taken at the same time each day (00 and/or 12 UTC). Around 590 reports are produced at these times respectively.

2.2.2.7 PILOT

PILOT is a numerical code used for reporting upper-level wind observations performed by rawins or pilot balloons. Around 230, 210 and 70 reports are produced at 00, 12 and 06/18 UTC, respectively.

2.2.2.8 Aircraft

Aircraft observations are reported via Aircraft Report (AIREP) and Aircraft Meteorological Data Relay (AMDAR). While vertical profile data can be obtained at the vicinity of airports, only flight level data can be collected along the other airways. The numerous reports received from the U.S. are thinned to 1/50th over the continental U.S. in analysis pre-processing. Even after this processing, Around 85,000 reports are produced every six hours covering areas around the world. Aircraft data availability is improving from the reduced volume caused by the COVID-19 pandemic that began around March 2020.

Table 2.2.1: Summary of the observation types and parameters used in objective analysis. Third column: P: surface pressure; u: zonal wind; v: meridional wind; T: temperature; Rh: relative humidity; T_B: radiance in brightness temperature; R_1: precipitation amount; Pwv: precipitable water vapor;Vr: radial velocity, Smc: soil moisture content, Sd: snow depth. Fourth column: GA: global analysis; MA: meso-scale analysis; LA: local analysis.

Observation type (or code name used for reporting observation)	Brief description	Parameters used in analysis	Analysis type which observations are used
SYNOP	Land surface observations from world weather stations	P, u, v, T, Rh, Sd	GA, MA, LA
AMeDAS	Land surface automated observation network in Japan	u, v, T, Rh	$MA(Rh), LA$
METAR	Routine weather reports from aerodromes	P, T, Rh	GA
SHIP	Sea surface observations from ships, oil rigs and moored buoys	P, u, v, T, Rh	GA, MA, LA
BUOY	Sea surface observations from drifting buoys	P, u, v, T, Rh	GA, MA, LA
TEMP	Upper-air observations from radiosondes	P, u, v, T, Rh	GA, MA, LA
PILOT	Upper-air wind observations from rawins or pilot balloons	u, v	GA, MA, LA
Aircraft	Upper-air observations from aircraft (mainly commercial)	u, v, T	GA, MA, LA
Wind Profiler	Upper-air wind profile observations from Japan, Hong Kong and Europe	u, v	GA, MA, LA
AMV	Atmospheric motion vector (AMV) wind data from geostationary (GEO) satellites, low earth orbit (LEO) satellites and a combination of LEO and GEO	u, v	GA, MA, LA
Scatterometer	Ocean surface wind vector data from scatterometers on LEO satellites	u, v	GA, MA
MW Sounder	Radiance data from microwave (MW) sounders on LEO satellites	T_B	GA, MA, LA
MW Imager	Radiance data from MW imagers on LEO satellites and precipitation amounts estimated from MW imager radiance data	T_B, R_1(MA only)	GA, MA, LA
CSR	Clear sky radiance (CSR) data from water vapor channels on GEO satellite infrared imagers	T_B	GA, MA, LA
Hyperspectral IR Sounder	Radiance data from infrared(IR) sounders on polar orbiting satellites	T_B	GA, MA, LA
GNSS-RO	Bending angle and refractivity profile data retrieved from radio occultation (RO) measurements of global navigation satellite system (GNSS) receivers on LEO satellites	Bending Angle Refractivity	GA(Bending Angle), MA(Refractivity)
Ground-based GNSS	Zenith total delay (ZTD) data and precipitable water vapor (PWV) data estimated from ground-based GNSS receivers	ZTD,Pwv	$GA(ZTD)$, $MA, LA(Pwv)$
Shipborne GNSS	ZTD and PWV data estimated from shipborne GNSS receivers	Pwv	MA, LA
Radar Reflectivity	Relative humidity data estimated using 3-dimensional reflectivity data from JMA weather (Doppler) radars and Dual-frequency Precipitation Radar (DPR) onboard GPM-core satellite.	Rh	MA, LA
Radial Velocity	Radial velocity data from JMA weather Doppler radars (WDRs) and Doppler radars for airport weather (DRAWs)	Vr	MA, LA
R/A	Radar estimated precipitation amounts calibrated using AMeDAS raingauge network data	R_1	MA
Soil Moisture Contents	Soil moisture data retrieved from microwave imager radiances and microwave scatterometer observations	Smc	LA
Typhoon Bogus	See Section 2.4.	P, u, v	GA, MA

2.2.2.9 Wind Profiler

Upper air wind speed and direction are monitored by wind profilers on the ground. A total of 33 wind profilers operated by JMA produce data every 10 minutes. The specifications are detailed in Ishihara *et al.* (2006). Wind profiler data from Europe and Hong Kong are also available.

2.2.2.10 AMVs

Atmospheric motion vector (AMV) wind data are derived by tracing the movement of cloud and water vapor patterns in successive satellite images. AMVs from geostationary (GEO) satellites (Meteosat-9, GOES-16 and Himawari-9), low-earth orbit (LEO) satellites (Aqua, NOAA and Metop), LEOGEO-AMVs and Dual-Metop AMVs are used. LEOGEO-AMV data are derived using imagery from a combination of polar-orbiting and geostationary satellites for 60°N and 60°S latitude areas. Dual-Metop AMVs are derived using tandem Metop-B and -C imagery, and have global coverage. AMVs from GEO satellites cover the area from 60°N to 60°S, and those from LEO satellites cover polar regions (i.e., latitudes higher than 60°).

2.2.2.11 Scatterometers

Ocean surface wind vectors from scatterometers polar-orbiting satellites are used. Data from ASCAT (the advanced scatterometer) onboard Europe's Metop-B, -C polar-orbiting satellites are currently utilized.

2.2.2.12 MW Sounders

Microwave (MW) sounder radiance data from AMSU-A (Advanced Microwave Sounding Unit-A) on NOAA-15, -18, -19, Metop-B and -C, MHS (Microwave Humidity Sounder) units on NOAA- 19, Metop-B and -C, ATMS (Advanced Technology Microwave Sounder) units on Suomi-NPP (the Suomi National Polar-orbiting Partnership) and NOAA-20, and MWHS-2 (Microwave Humidity Sounder 2) on FY-3C are used. AMSU-A is a sounder primarily monitoring atmospheric temperature profiles, while MHS and MWHS-2 are sounders primarily monitoring middle-to-upper tropospheric humidity profiles, and ATMS has temperature and humidity sounding channels. Related radiance data in meso-scale and local analysis are assimilated under clear-sky conditions (i.e., with no effect from cloud/rain). In global analysis, MHS, ATMS (humidity channels) and MWHS-2 radiances are assimilated under all-sky conditions, while other MW sounder radiances are assimilated under clear-sky conditions.

2.2.2.13 MW Imagers

Radiances from the following MW imagers are used; AMSR2 (Advanced Microwave Scanning Radiometer 2) on GCOM-W (Global Change Observation Mission 1st - Water), SSMIS (Special Sensor Microwave Imager Sounder) on DMSP (Defense Meteorological Satellite Program) -F17, -F18 and GMI (GPM (Global Precipitation Measurement mission) Microwave Imager) on GPM. Their radiances are sensitive to water vapor amounts in the lower troposphere. They are assimilated under all-sky conditions in global analysis, and assimilated under clear-sky conditions in meso-scale and local analysis. Precipitation amounts estimated from radiances using the MSC method (Takeuchi and Kurino 1997) are also used in meso-scale analysis. Radiances from AMSR2, SSMIS are used for snow-cover product as well (see Subsection 2.8.1).

2.2.2.14 CSR

Clear sky radiance (CSR) is a product providing averaged radiance over cloud-free pixels in GEO satellite imagers. CSR data from GEO satellites (Meteosat-9, GOES-16 and Himawari-9) are used. CSR data of water vapor channels/bands which are sensitive to water vapor amounts in the upper and middle troposphere are used.

2.2.2.15 Hyperspectral IR Sounders

Radiances under clear-sky conditions from hyperspectral IR sounders are used. The data adopted are from IASI on Metop and CrIS on Suomi-NPP and NOAA-20.

2.2.2.16 GNSS-RO

GNSS-RO (Global Navigation Satellite Systems - Radio Occultation) is a technique for measuring atmospheric profiles. With this approach, a set of atmospheric time delay data of GNSS radio signals received by a LEO satellite is obtained during each radio occultation event. Since the delay is a result of atmospheric radio refraction along the propagation path of the signal, the vertical profiles of refractivity (or the bending angle) of the atmosphere at a tangent point can be estimated from the delay data set. As refractivity is a function of temperature, humidity and pressure, it can be used to determine the profiles of these properties. The currently used LEO satellites and their GNSS receivers are IGOR (Integrated GPS Occultation Receiver) onboard TerraSAR-X and TanDEM-X, GRAS (GNSS Receiver for Atmospheric Sounding) onboard Metop-B.

2.2.2.17 Ground-based GNSS

Ground-based GNSS data are provided from atmospheric time delays of GNSS radio signals collected by ground-based GNSS receivers. JMA uses ground-based GNSS data collected from the global network along with GEONET GNSS receiver data (Ishikawa 2010). GEONET is a ground-based GNSS receiver network operated by the Geospatial Information Authority in Japan using around 1,300 receivers located throughout the country. GNSS-ZTD (GNSS - zenith total delay) data are estimated by averaging the delays of multiple GNSS satellite signals monitored using a single receiver, and are used in global analysis. GNSS-PWV (GNSS - Precipitable Water Vapor) data based on analysis of GEONET GNSS-ZTD data are used in meso-scale and local analysis.

2.2.2.18 Shipborne GNSS

Shipborne GNSS data are provided every 10 minutes from two JMA observation vessels, four Japan Coast Guard vessels and ten merchant vessels (as of July 2023). Shipborne GNSS-derived precipitable water vapor (PWV) data (Shoji *et al.* 2017) are used in meso-scale and local analysis. One JMA vessel mission is to conduct targeted observations toward improved forecast accuracy for heavy rainfall in NWP during Japan's rainy season.

2.2.2.19 Radar Reflectivity

By March 2023, 10 of JMA's 20 C-band radars had been upgraded to dual-polarization types. With this system, three-dimensional reflectivity data are produced every 5 minutes, and relative humidity profiles are estimated from reflectivity data and NWP grid point values using a technique based on Bayes theorem (Caumont *et al.* 2010). Relative humidity data are produced for areas within a 200-km radius of each radar site below freezing level and used in meso-scale and local analysis. Relative humidity profiles estimated from reflectivity data collected using the Dual-frequency Precipitation Radar (DPR) on the GPM-core satellite are used in meso-scale analysis.

2.2.2.20 Radial Velocity of Doppler Radar

By March 2023, 10 of JMA's 20 Doppler Weather Radars (DWRs) and 8 of its 9 Doppler Radars for Airport Weather (DRAWs) had been upgraded to C-band dual-polarization types. With this system, three-dimensional radial velocity data are produced every 5 minutes within a radius of 150 km for DWRs (range resolution: 250 m) and 120 km for DRAWs (150 m). The azimuth resolution is 0.703° for both.

2.2.2.21 Analysis of Precipitation(R/A)

R/A is a product providing composite precipitation data produced by JMA. These data are cumulative precipitation estimations based on weather radar data with a Z-R relationship ($Z = 200R^{1.6}$) calibrated using AMeDAS raingauge data in real time. The details are found in Subsection 4.4.1.

2.2.2.22 Soil Moisture Content

Soil moisture content data estimated from microwave imager (AMSR2/GCOM-W) radiances and microwave scatterometer (ASCAT/Metop-B) observations are used for areas over land in local analysis.

2.3 Quality Control and Related Procedures

Quality control (QC) is a series of procedures by which "bad" observations are screened out. It is a vital component of the objective analysis system because observations sometimes include large errors and erroneous data can significantly impair the quality of atmospheric analysis, leading to low levels of forecast skill. QC procedures in JMA's objective analysis systems are described in the following subsections.

2.3.1 SYNOP, AMeDAS, METAR, SHIP, BUOY, TEMP, PILOT, Aircraft and Wind Profilers

Direct observations (i.e. SYNOP, AMeDAS, METAR, SHIP, BUOY, TEMP, PILOT and aircraft) and wind profilers measure prognostic variables in NWP such as pressure, temperature, wind and humidity. The QC system for these observations consists of internal and external QCs.

2.3.1.1 Internal QC

Internal QC involves procedures to check and correct observation values using collocated data in reports and several external lists or tables. The checks are outlined below.

1. Blacklist check: The blacklist is a list of problematic stations or data, and is prepared in advance via non-real-time QC (see Section 2.10). Blacklisted observations are rejected in this step.

2. Climatological check: Climatological reasonability is checked in this step. The criteria are defined in advance based on WMO (1993).

3. Trajectory check: Consistency at consecutive locations is checked for reports from moving stations such as SHIP, BUOY and aircraft. The movement velocity and direction are checked in this step and checking is also performed to ensure that SHIP and BUOY locations are in the ocean.

4. Inter-element consistency check: The temporal continuity of consecutive reports from surface stations is checked along with consistency among observation elements within the report.

5. Vertical consistency check: Vertical consistency is checked in TEMP and PILOT data. The check items are (1) instrument icing, (2) temperature lapse rate, (3) hydrostatic relationship, (4) consistency among data at standard pressure levels and those at significant levels and (5) vertical wind shear.

6. Bias correction: Bias correction is applied to TEMP data reported without radiative heating correction or with apparent systematic biases. Correction constants are determined from one-month statistics for the previous month. Kalman Filter is utilized to correct aircraft temperature biases in global analysis. Bias correction values are updated at each analysis time (00, 06, 12, 18 UTC) and calculated independent of analysis using a covariance of the first-guess departure separately for each aircraft, each vertical level and each flight phase (ascent, cruising and descent). In mesoscale analysis, aircraft temperature biases are corrected for each aircraft and vertical level using bias correction values statistically calculated using the first-guess departure from global analysis for the previous month.

2.3.1.2 External QC

External QC involves procedures to check observation values with comparison to (external) first guess and neighboring observations. The checks are outlined below.

1. Gross error check: The departure ($D \equiv O - B$) of the observed value (O) from the first guess (B) is calculated for all observations. The absolute value of D is compared with the tolerance limits C_P (the pass criterion) and C_R (the failure criterion). Data satisfying $|D| \leq C_P$ pass the QC, and those characterized by $|D| > C_R$ are rejected. Data characterized by $C_P < |D| \leq C_R$ are regarded as suspect and sent for spatial consistency checking.

2. Spatial consistency check: The departure D in suspect observation data is compared with departures interpolated using the optimum interpolation method (D_{OI}) with neighboring observations. The absolute difference of D and D_{OI} is compared with the tolerance limit C_S (the criterion for suspect) for final judgment and the data satisfying $|D - D_{OI}| \leq C_S$ are accepted.

   Here, the tolerance limits C_P, C_R, and C_S vary with local atmospheric conditions in first guess fields. The limits are made small if the time tendency and horizontal gradient are small in the fields, and *vice versa*. This scheme is called Dynamic QC (Onogi 1998).

3. Duplication check: Duplication is often found in observation reports with data obtained through different communication lines. The most appropriate report is picked out from among duplicates after performance of the above checks in consideration of status.

2.3.2 AMVs

Blacklisted AMVs (Table 2.3.1) are rejected in the first step, as are those with low quality indicators (QI, Holmlund 1998). QI thresholds are defined for each satellite, domain, vertical level and image type. Thinning is then performed based on a distance of 200 km. Around Japan region from 120°E to 150°E and from 20°N to 45°N, super-observation (SPOB) data are derived from Himawari-9 AMVs to arrange at 100 km intervals. (Yamashita 2016). Climatological checking (see Subsection 2.3.1.1) and external QC (see Subsection 2.3.1.2) are then performed. The details of QC and detailed settings for AMVs are given on the NWP SAF AMV monitoring page[1].

2.3.3 Scatterometers

Level 2 ocean surface wind data are used in global, meso-scale and local analysis, with initial rejection of low-quality data for land/sea ice areas. The most likely wind directions are then identified from inherent ambiguity scatterometer data with application of NWP nudging and median filtering. To avoid undesirable rejection of correct wind data in subsequent gross error checking (see Subsection 2.3.1.2) for areas in and around severe weather systems such as cyclones and fronts where wind direction/speed vary sharply, quality control (Group-QC) is applied. In this step, spatial consistency among wind vectors is checked in relation to smooth transition in wind direction/speed, and valid data are excluded from rejection (see the NWP SAF scatterometer monitoring page for details[2].

2.3.4 Satellite Radiance

Satellite radiance data are used in global, meso-scale and local analysis as a form of brightness temperature. The RTTOV-13.0 fast radiative transfer model (Saunders *et al.* 2020) is employed for radiance assimilation[3]. The common QC procedures for radiance data are blacklist checking, thinning and external QC. The blacklist specifies problematic instruments, and is made in advance based on non-real-time QC (see Section 2.10). Blacklisted data are rejected in the first step. In the next step, data are thinned spatially in each time slot of the assimilation window (approximately one hour) to reduce computational costs. The subsequent external QC includes reduction of instrumental scan biases (except for CSR), cloud/rain contamination checking, location checking, channel selection and gross error checking (see Subsection 2.3.1.2). Data passing this QC

[1] https://nwp-saf.eumetsat.int/site/amv-usage-in-the-jma-nwp-model/

[2] https://nwp-saf.eumetsat.int/site/monitoring/winds-quality-evaluation/scatterometer-mon/use-in-nwp/jmamodel/

[3] Coefficients of RTTOV-10.2 are still used.

Table 2.3.1: AMV blacklist summary. IR: infrared; WV: water vapor; CSWV: clear sky water vapor; NH: Northern Hemisphere; SH: Southern Hemisphere; Polar AMV: AMVs from polar-orbiting satellites; GEO AMV: AMVs from geostationary satellites other than Himawari-9; LEOGEO AMV: AMVs from a combination of low earth orbiting and geostationary satellites.; Dual-Metop AMV: AMVs from a combination of Metop-B and C.

Kind	Blacklisting area
Polar AMV (IR) at NH	above 300 hPa or below 900 hPa
Polar AMV (WV/CSWV) at NH	above 300 hPa or below 550 hPa
Polar AMV (IR/WV) at SH	above 300 hPa or below 550 hPa
Polar AMV (CSWV) at SH	above 350 hPa or below 550 hPa
Polar AMV (All)	poleward of 88°N or 88°S
GEO AMV (All)	above 175 hPa or below 975 hPa, and between 400 hPa and 825 hPa
GEO AMV (IR)	above 275 hPa at poleward of 20°N or 20°S
GEO AMV (WV)	above 225 hPa at poleward of 20°N or 20°S
Himawari-9 (IR/VIS)	below 700 hPa over land
Meteosat-9 (IR/VIS)	below 700 hPa over land
LEOGEO AMV	above 300 hPa or below 900 hPa
LEOGEO AMV	between 600 hPa and 640 hPa
LEOGEO AMV	at poleward of 70°N or 70°S
LEOGEO AMV	between 640 hPa and 900 hPa at poleward of 60°S
Dual-Metop AMV	between 50°N and 50°S, and at poleward of 88°N or 88°S
Dual-Metop AMV	above 300 hPa, and below 700 hPa over land

Table 2.3.2: Summary of microwave sounder channel sets used for each condition. In meso-scale and local analysis, AMSU-A Ch. are used up to 11 and 8, respectively.

	AMSU-A	MHS	ATMS
Clear sky ocean	Ch. 4–14	Ch. 3–5	Ch. 6–9
Clear sky land/coast/sea-ice	Ch. 6–14	Ch. 3–5	Ch. 7–9
Cloudy ocean	Ch. 7–14	Ch. 3–5	Ch. 8–9
Rainy ocean	Ch. 9–14	n/a	n/a

are thinned again to reduce the observation error correlation, and the thinned data are output for use in data assimilation systems. Variational bias correction (VarBC, Derber and Wu 1998; Dee 2004) is used to reduce air-mass dependent biases. VarBC is an adaptive bias correction scheme in which a linear regression formula representing biases is embedded in the observation operator and regression coefficients are set as analysis variables. The formulations are described in Subsection 2.5.7.4. The satellite radiance data used in the analyses are from MW sounders, MW imagers, CSR and hyperspectral IR sounders. The specific procedures for each data type are described in the following subsections.

2.3.4.1 MW sounders

For radiances assimilated under clear-sky conditions (not affected by cloud/rain), the sets of channels used are defined in advance according to individual surface and atmospheric conditions at the observation point (Table 2.3.2). In global analysis, microwave humidity sounders' radiances (183 GHz) are assimilated under all-sky conditions by assigning observation errors depending on cloud amount (see Subsection 2.5.7.2).

2.3.4.2 MW imagers

Vertically polarized-channel radiances are assimilated over ice-free ocean areas. Radiances are assimilated under all-sky conditions in global analysis (see Subsection 2.5.7.2), and assimilated under clear-sky conditions in meso-scale and local analysis. In meso-scale analysis, precipitation retrieval is also assimilated for areas over the ocean surrounding Japan. Precipitation amount estimations are resampled onto inner model grids with spatial smoothing.

2.3.4.3 CSR

CSR data are horizontally thinned to divisions of 220 km for global analysis and 45 km for meso-scale analysis and local analysis. Hourly (or temporally thinned to the hourly slots) CSR data are used in these analysis types. Values with a low percentage of clear pixels and a large standard deviation of brightness temperature are excluded due to their low representation of the area. CSR data from high-altitude areas (above 4,000 m) are not used. In calculation with RTTOV-13.0, emissivity atlas data and retrieved surface temperatures from window channel radiance are used for areas over land.

2.3.4.4 Hyperspectral IR sounders

Clear-sky radiance data from hyperspectral IR sounders are used in global, meso-scale and local analysis. Data are horizontally thinned to divisions of 200 km for global analysis and 45 km for meso-scale and local analysis. Cloud screening with estimation of cloud top height are applied in quality control. Usage is based on CO2 slicing (Eyre and Menzel 1989).

2.3.5 GNSS-RO

Bending angle data for altitudes up to 60 km are used in global analysis with 500-m vertical intervals. Refractivity data are used in meso-scale analysis. No bias correction is applied for GNSS-RO data.

2.3.6 Ground-based GNSS

ZTD data are used in global analysis. Stations at an elevation higher than 5,000 m are not used, and those from which the absolute difference of elevation to the model surface exceeds 300 m are excluded. GNSS-ZTD values smaller than 1,000 mm or larger than 3,000 mm are rejected in climatological checking. Data with absolute differences of more than 50 mm from first guess are regarded as suspect. If the absolute difference of departure D of suspect data from averaged D of vicinity data exceeds 50 mm, the ZTD data are not used. PWV data are used in meso-scale and local analysis. As Japan is characterized by steep mountainous terrain, large differences are found between actual ground surface elevations and model surface elevations especially in mountainous areas. In meso-scale analysis, stations at an elevation of 700 m or more above mean sea level are not used, and those from which the absolute difference of elevation to the model surface exceeds 200 m are excluded. GNSS-PWV values smaller than 1 mm or larger than 90 mm are rejected in climatological checking. The first guess PWV is then interpolated or extrapolated to the actual terrain surface and compared to the GNSS-PWV. The criterion C_P, C_R and C_S referred in Subsection 2.3.1.2 are set to 5 mm, 8 mm and 5 mm respectively for GNSS-PWV. As there is a dense GNSS-PWV network for the analysis systems, data are thinned by 30 km for meso-scale analysis and 20 km for local analysis.

2.3.7 Shipborne-GNSS

Quality control and data thinning for shipborne GNSS-PWV data are almost the same as for ground-based GNSS-PWV data (which have been used in mesoscale analysis since 2009 as per Subsections 2.2.2.17 and 2.3.6), except for the use of Dynamic QC (see Subsection 2.3.1.2) in gross error checking for shipborne GNSS-PWV. Quality parameters, the number of satellites and analysis times (reported only by shipborne GNSS) are also checked.

2.3.8 Radar Reflectivity

To assimilate radar reflectivity data in meso-scale analysis and local analysis, an indirect assimilation technique called 1D+4DVAR (Ikuta and Honda 2011) is employed. This approach is based on Caumont *et al.* (2010). In 1D+4DVAR, radar reflectivity data are used to retrieve relative humidity (RH) values, which are assimilated as conventional observation data in 4D-Var. In this system, only retrieved RH values from below the melting layer are used because it is known that reflectivity inappropriately simulated in the ice phase with the operational MSM hydrometeor forecast, causing large biases in RH retrievals. In addition, data from around a height of 2000 m above sea level are also not used since these data are used for making R/A and are already assimilated in meso-scale analysis in another form (surface rainfall, see Subsection 2.3.10). For operation, reflectivity data from the JMA C-band radar network are used. Reflectivity data from space-based Dual frequency Precipitation Radar are used to retrieve RH values and assimilated in meso-scale analysis.

2.3.9 Radial Velocity of Doppler Radar

Hourly radial velocity data from WDRs and DRAWs are used in meso-scale analysis and local analysis. In pre-processing, the data are resampled into a 5 km range resolution and a 5.625° azimuthal resolution. The resampled data are checked with respect to the number of data sampled, radial velocity variance and the difference between maximum and minimum velocity. High elevation angle data ($\geq$ 5.9°) are not used to avoid the contamination of precipitation velocity values, and those from areas close to radar site ($<$ 10 km) are not used to avoid the influence of back scatter noise. Data showing wind speeds of less than 5 m/s are also not used to avoid ground clutter contamination.

2.3.10 R/A

Hourly R/A data are assimilated in meso-scale analysis. After quality control, R/A data (1 km grid) are resampled into inner-model grid boxes (15 km) and input for this type of analysis.

2.3.11 Soil Moisture Content

Variable transformation using the cumulative distribution function (CDF) matching method is applied for soil moisture content (SMC). The CDF matching method involves fitting the probability density function (PDF) of observation to the PDF of model variables. This pre-conditioning via CDF matching helps to minimize cost function because the innovation of SMC becomes Gaussian after the CDF matching. A variational bias correction method is used for SMC in local analysis.

2.3.12 CDA: Feedback Data Base

All information concerning the quality of observational data obtained during the quality control procedure is archived in the Comprehensive Database for Assimilation (CDA) format, which is extensively used for both real-time and non real-time data monitoring activities. All information contained in the CDA is managed in the form of integer two byte data. The database format is designed for flexible use so that information on observations can be archived easily, and is also user-friendly to facilitate data retrieval.

2.4 Typhoon Bogussing

For tropical cyclones (TCs) over the western North Pacific, typhoon bogus data are generated as a form of pseudo-observation information and assimilated for realistic TC structure analysis based on model resolutions. The data consist of pressures values at the mean sea level (P_{msl}) and vertical profiles of the wind (W_{prf}) around TCs. Wind profiles are placed at 850 and 300 hPa in global analysis, 1000, 925, 850, 800, 700, 600, 500, 400 and 300 hPa in meso-scale analysis. The generated bogus has an axially asymmetric structure in types of the analysis.

Firstly, symmetric bogus profiles are generated automatically from central pressure values and the 15m/s wind speed radius of a TC (R_{15}) as analyzed by forecasters. The surface pressure profile is defined using Fujita's formula (Fujita 1952). Gradient wind balance is assumed for calculation of surface pressure profile meeting the requirements at hand from the wind speed at the particular radius of R_{15}. Upper geopotential profiles are defined using an empirical formula based on the TC analysis described by Frank (1977). It is assumed that the temperature anomaly has its maximum at 250 hPa. The wind field on each level is derived from geopotential height profiles with gradient wind balance. The surface wind field is also derived from gradient wind balance, but is modified to include the effects of surface friction.

Secondly, asymmetric components are retrieved from the first guess fields and added to the symmetric bogus profile to generate the final asymmetric bogus structure. When the target area of bogussing is across the lateral boundary in the meso-scale analysis, asymmetric components are not added.

Finally, pseudo-observation data are generated from the resulting bogus structure at the TC center analyzed (P_{msl}), the TC center of the first guess (P_{msl}), and several points surrounding the TC center analyzed (P_{msl} and W_{prf}). The configuration for the surrounding point distribution is adaptable to the typhoon track error of the first guess.

2.5 Global Analysis

2.5.1 Introduction

A hybrid four-dimensional variational (4D-Var) data assimilation method using a Local Ensemble Transform Kalman Filter (LETKF, Hunt *et al.* 2007) is employed in analysis of the atmospheric state for the Global Spectral Model (GSM), and is performed at 00, 06, 12, 18 UTC. An early analysis with a short cut-off time is performed to prepare initial conditions for operational forecasting, and a cycle analysis with a long cut-off time is performed to maintain the quality of the global data assimilation system. An incremental method (Courtier *et al.* 1994) is adopted to improve computational efficiency using outer-loop iterations, with increments evaluated at a lower (inner) resolution (TL319L128: grid roughly equivalent to 0.5625° (55 km) and up to 0.01 hPa). The increment is then interpolated and used to update the model trajectory at the original resolution (TQ959L128: grid roughly equivalent to 0.125° (13 km)), and the updated trajectory is used to refine the cost function for subsequent inner-loop iterations. Specification of the JMA Global Analysis system is summarized in Section 2.1.

The three-dimensional variational (3D-Var) data assimilation system was operated as the first operational variational analysis system for GSM with the inner resolution of T106L40 (grid roughly equivalent to 110 km and up to 0.4 hPa) in September 2001 (Takeuchi and Tsuyuki 2002). In February 2005, the 4D-Var data assimilation system was operated with the inner resolution of T63L40 (grid roughly equivalent to 180 km) (Kadowaki 2005). The inner resolution was upgraded to T106L40 in March 2006 (Narui 2006), T159L60 (grid roughly equivalent to 75 km and up to 0.1 hPa) in November 2007, TL319L60 in October 2011 (Kadowaki and Yoshimoto 2012), TL319L100 in March 2014. The hybrid 4D-Var data assimilation system using the LETKF with two outer-loop iterations was operated in December 2019 (Kadowaki *et al.* 2020). In March 2021, the number of ensemble members in LETKF was increased from 50 to 100, weighting for ensemble background error covariances was increased from 0.15 to 0.5 (Yokota *et al.* 2021) and the inner resolution was upgraded to TL319L128.

2.5.2 Incremental 4D-Var Formulation and Procedural Description

In 4D-Var, 3–9-hour forecasts from the GSM are used as a first guess (background), and all observations passed Quality Control (QC) from within three hours of analysis time (6-hour) are organized in hourly time slots as defined in Figure 2.5.1. The cost function is used to measure the distance between the model trajectory and observations over a six-hour assimilation window. In an incremental method, the analysis increment is evaluated at a low-resolution, and determined by minimization of the cost function in the inner-loop iterations. The low-resolution analysis increment is interpolated to the high-resolution analysis increment. Adding this analysis increment to the guess field produces high-resolution analysis.

In the outer-loop iterations, procedures mentioned above are iterated as in Figure 2.5.2 :

1. The GSM forecast from "latest" high-resolution analysis is performed to update the guess field (or model trajectory).

2. QC processes of observations are reperformed.

3. The cost function is updated.

4. The low-resolution analysis increment is updated by minimization of the cost function in the inner-loop iterations and used to update the high-resolution analysis.

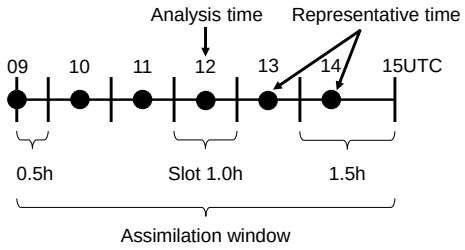


Figure 2.5.1: Schematic diagram of time slots for the analysis time 12 UTC. The black circles indicate the representative time of each time slot.

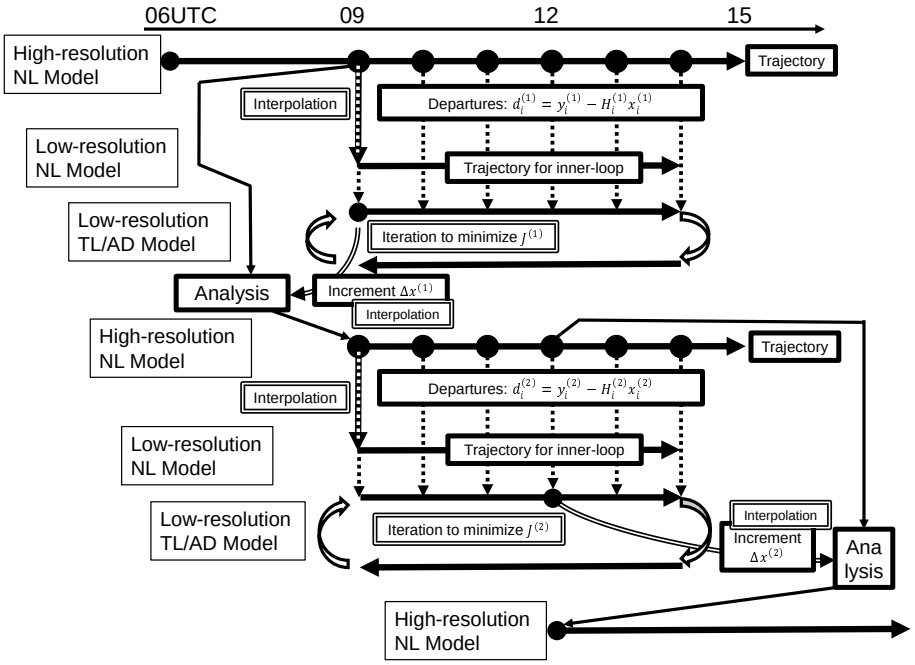


Figure 2.5.2: Flow of 4D-Var operation for the 12 UTC analysis time

In the $j(\geqq 1)$-th outer-loop, the cost function $J^{(j)}$ to determine the j-th low-resolution analysis increment update $\Delta x_i^{(j)}$ is defined by Eq. (2.5.1).

$$J^{(j)}\left(\Delta x_0^{(j)}\right) = \frac{1}{2}\Big(\sum_{l=1}^{j}\Delta x_0^{(l)}\Big)^{\mathrm{T}}\mathbf{B}^{-1}\Big(\sum_{l=1}^{j}\Delta x_0^{(l)}\Big) + \frac{1}{2}\sum_{i=1}^{n}\left(\mathbf{H}_i{}^{(j)}\Delta x_i^{(j)} - d_i^{(j)}\right)^{\mathrm{T}}\mathbf{R}_i^{(j)-1}\left(\mathbf{H}_i{}^{(j)}\Delta x_i^{(j)} - d_i^{(j)}\right) + J_C^{(j)}$$
$$\Delta x_{i+1}^{(j)} = \mathbf{M}_i^{(j)}\Delta x_i^{(j)} \qquad (i = 0, \ldots, n-1) \tag{2.5.1}$$

where the subscript $i(\geqq 1)$ is the index of time slot and n is the number of time slots. $\Delta x_0^{(j)}$ ($\Delta x_1^{(j)}$) is the low-resolution analysis increment at the initial time before (after) initialization, $\Delta x_{i\geqq 2}^{(j)}$ is the analysis increment evolved according to the tangent linear (TL) model to time of slot i. $\mathbf{M}_0^{(j)}$ is the TL operator of the nonlinear normal-mode initialization operator (Machenhauer 1977) and $\mathbf{M}_{i\geqq 1}^{(j)}$ is the TL model of the low-resolution nonlinear (NL) forecast model $M_i^{(j)}$ as detailed in Subsection 2.5.3 for integration from time of slot i to that of slot $i+1$. $\mathbf{R}_i^{(j)}$ denotes the covariance matrix of observation errors at time of slot i, and $\mathbf{B}$ is the covariance matrix of background errors as detailed in Subsection 2.5.5 and Subsection 2.5.7. $\mathbf{H}_i^{(j)}$ is the TL operator of the observation operator $H_i^{(j)}$. The innovation vector is given for each assimilation slot by $d_i^{(j)} = y_i^{(j)} - H_i x_i^{(j)}$, where $x_i^{(j)}$ is the j-th guess field evolved using the high-resolution NL model, and $y_i^{(j)}$ is the observation data passed QC at time of slot i. $J_C^{(j)}$ is a penalty term used to suppress gravity waves described in Subsection 2.5.4.

To minimize the cost function $J^{(j)}$, the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm (Liu and Nocedal 1989) with Veersé's preconditioner (Veersé *et al.* 2000) is applied. Here, the gradient of the cost function ∇J is required. This is determined via the adjoint procedures of Eq. (2.5.2)–Eq. (2.5.5) as computed in reverse time.

$$p_{n+1} = 0 \tag{2.5.2}$$
$$p_i = \mathbf{M}_i^{(j)\mathrm{T}}p_{i+1} + \mathbf{H}_i^{(j)\mathrm{T}}\mathbf{R}_i^{(j)-1}\left(\mathbf{H}_i^{(j)}\Delta x_i^{(j)} - d_i^{(j)}\right) \qquad (i = n, \ldots, 1) \tag{2.5.3}$$
$$p_0 = \mathbf{M}_0^{(j)\mathrm{T}}p_1 + \mathbf{B}^{-1}\Big(\sum_{l=1}^{j}\Delta x_0^{(l)}\Big) \tag{2.5.4}$$
$$\nabla J^{(j)}\left(\Delta x_0^{(j)}\right) = p_0 \tag{2.5.5}$$

where p_i is a auxiliary variable, $\mathbf{M}_i^{(j)\mathrm{T}}$ represents the adjoint (AD) model of the TL model $\mathbf{M}_i^{(j)}$, and $\mathbf{H}_i^{(j)\mathrm{T}}$ is the AD operator of $\mathbf{H}_i^{(j)}$. Note that Eq. (2.5.2)–Eq. (2.5.5) should contain additional terms for the penalty term in Eq. (2.5.1), which are omitted here for simplicity.

2.5.3 Inner Model

The inner NL model is based on the GSM, but many processes are based on those of the older GSM for many reasons. In particular, moisture processes (the convection and cloud schemes) are based on those of GSM0103 (JMA 2002), mainly for the stability of inner TL model integration. Nonlinear normal-mode initialization (Machenhauer 1977) is also added.

The inner TL model includes the following simple processes, most of which are based on the inner NL model:

1. **Initialization:** To control gravity waves, the TL version of the nonlinear normal-mode initialization is adopted.

2. **Horizontal Diffusion:** Horizontal diffusion is enhanced over that of the inner NL model based on Buizza (1998).

3. **Surface Turbulent Fluxes:** Surface turbulent fluxes are formulated as Monin-Obukhov bulk formulae based on the inner NL model. Sensible and latent heat flux are perturbed only over the sea.

4. **Vertical Turbulent Transports:** The vertical turbulent transports of momentum, heat, and moisture are formulated as the hybrid downgradient-type scheme based on the inner NL model (Subsection 3.2.7). Those diffusion coefficients are not perturbed.

5. **Orographic Gravity Wave Drag:** The parameterization for orographic gravity wave drag consists of two components: one for long waves (wavelength > 100 km) and the other for short waves (wavelength ≈ 10 km) based on the inner NL model which is almost same as in GSM1705 (JMA 2018). The Richardson number is not perturbed in some parts for long waves for the stability of inner TL model integration.

6. **Long-wave Radiation:** Long-wave radiation in the TL model is based on Mahfouf (1999). The tendency of the perturbed temperature T' is given by

$$\frac{\partial T'}{\partial t} = -\alpha \frac{g}{C_p} \frac{\partial}{\partial p}\left(4F\frac{T'}{T}\right) \tag{2.5.6}$$

   where $\alpha = 1/\{1 + (p_r/p)^{10}\}$, $p_r = 300$ hPa, F represents the net radiation fluxes calculated in the inner NL model, and g and C_p denote the gravitational constant and isobaric specific heat, respectively.

7. **Clouds and Large-scale Precipitation:** Clouds and large-scale precipitation are based on the inner NL model, in which the former are prognostically determined in a way similar to that proposed by Smith (1990). A simple statistical approach proposed by Sommeria and Deardorff (1977) is employed to compute cloud amounts and cloud water content. The parameterization of the conversion rate from cloud ice to precipitation follows the scheme proposed by Sundqvist (1978). These considerations are much simplified in the TL model. The cloud fraction, the amount of falling cloud ice and the dependence on water vapor of isobaric specific heat are not perturbed. Only certain variables are perturbed in computing the conversion from cloud water to precipitation and the evaporation of precipitation.

8. **Cumulus Convection:** Cumulus convection is formulated as the prognostic Arakawa-Schubert scheme (Arakawa and Schubert 1974; Moorthi and Suarez 1992; Randall and Pan 1993) based on the inner NL model, but is highly simplified. Vertical wind shear and the planetary mixing length are not perturbed. The magnitude of mass-flux perturbation is set bounds for the stability of inner TL model integration, and as a result this model is not exactly linear.

2.5.4 Penalty Term

The penalty term, which is the third term of Eq. (2.5.1), is given by

$$J_C^{(j)} = \frac{1}{2}\alpha\left(\left|\mathrm{N}_G \sum_{l=1}^{j} \Delta x_0^{(l)}\right|^2 + \sum_{i=2}^{\mathrm{n}} \left|\mathrm{N}_G \sum_{l=1}^{j} \Delta x_i^{(l)}\right|^2\right) \tag{2.5.7}$$

where N_G denotes an operator used to calculate the tendency of the gravity wave mode based on Machenhauer (1977). α is an empirically determined constant $3.0 \times 10^{-2}[\mathrm{s}^4/\mathrm{m}^2]$. Although this penalty term is primarily introduced to suppress gravity waves in the analysis increment, it is also effective in stabilizing calculation.

2.5.5 Background Term

The background term, which is the first term on the right side of Eq. (2.5.1), dominates how the difference between observation data and the first guess is converted into correction for the first guess in the 4D-Var analysis procedure. In the 4D-Var, the evolution of the initial background error covariances over the length of the assimilation window is taken into account, although the initial background error covariances are climatological and do not represent the day-to-day weather situation. One way to take into account the error of the day is to use ensemble forecast as part of background error covariances. In JMA, the hybrid method composed of the LETKF and the 4D-Var is employed as the operational system. Figure 2.5.3 shows a schematic diagram of the hybrid 4D-Var using the LETKF.

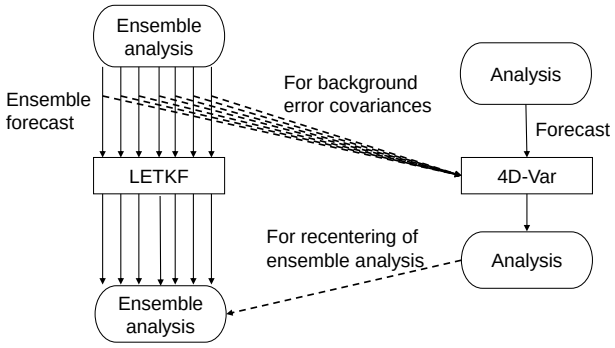


Figure 2.5.3: A schematic diagram of the hybrid 4D-Var using the LETKF.

Using the extended control variable method of Lorenc (2003), the background term in the hybrid 4D-Var is defined by Eq. (2.5.8).

$$\begin{aligned} \frac{1}{2}\Delta x^{\mathrm{T}}\mathbf{B}^{-1}\Delta x &= \frac{1}{2}\Delta x_{cl}^{\mathrm{T}}\mathbf{B}_{cl}^{-1}\Delta x_{cl} + \frac{1}{2}\Delta x_{en}^{\mathrm{T}}\mathbf{B}_{en}^{-1}\Delta x_{en} \\ \Delta x &= \beta_{cl}\Delta x_{cl} + \beta_{en}\Delta x_{en} \end{aligned} \tag{2.5.8}$$

where the subscript i and j in Eq. (2.5.1) are omitted here for simplicity. Δx is the low-resolution analysis increment. The subscript "cl" and "en" mean "climatological" and "ensemble". β_{cl} and β_{en} are the weight of "climatological" and "ensemble" background error covariances for making hybrid covariances. In JMA, both ${\beta_{cl}}^2$ and ${\beta_{en}}^2$ are 0.5 under 50 hPa, and they are approaching to 1 and 0 above that.

Two kinds of covariance inflations are applied for the ensemble background error covariances. One is an adaptive multiplicative covariance inflation used in the LETKF described in Subsection 2.5.6. The other is an additional covariance inflation to make vertical profiles of horizontal global mean of standard deviation from the ensemble covariances, consistent with those from the climatological background error covariances. A covariance localization is also applied for the ensemble background error covariances. The localization function is given as the Gaussian function of the distance between the analysis grid points. The localization scale l for which the localization function is $1/\sqrt{e}$ is set to 800 km for wind vectors, temperature and surface pressure and set to 400 km for specific humidity in the horizontal domain. The scale l is set to a 0.8 scale height in the vertical domain for all variables. A detailed description of the climatological background error covariances is provided in Subsection 2.5.5.1 - Subsection 2.5.5.5. A detailed description of the LETKF is provided in Subsection 2.5.6.

2.5.5.1 Climatological Background Error Covariance

For the climatological background error covariances, the multivariate couplings in the analysis variables are based on the geostrophic linear balance between mass and wind. Control variables are introduced to reduce the correlations among the analysis variables, and additional statistical relations are considered in the algorithm. These include the lower geostrophic balance on smaller horizontal and vertical scales, the almost complete lack of geostrophic balance near the equator, the dependency of geostrophy on the vertical level, and the weak coupling between divergence and vorticity as well as between divergence and mass.

The control variables are relative vorticity $\Delta\zeta$, unbalanced divergence $\Delta\eta_U$, unbalanced temperature and surface pressure (ΔT_U, ΔP_{sU}), and the logarithm of specific humidity $\Delta \ln q$ in the spectral space on the model layers. Δ denotes deviation from the first guess and subscript U means the term of "unbalanced". Autocovariances of the control variables are assumed to be homogeneous and isotropic. Correlation structures do not depend on geographical locations, but vertical correlations do depend on horizontal scale. The unbalanced variables $\Delta\eta_U$ and (ΔT_U, ΔP_{sU}) are defined as

$$\Delta\eta_U \equiv \Delta\eta - P\Delta\phi_B \qquad (\Delta\phi_B = \Delta\phi_B(\Delta\zeta)) \tag{2.5.9}$$

$$\begin{pmatrix}\Delta T_U \\ \Delta p_{sU}\end{pmatrix} \equiv \begin{pmatrix}\Delta T \\ \Delta p_s\end{pmatrix} - Q\Delta\phi_B - R\Delta\eta_U \tag{2.5.10}$$

where P, Q, and R are regression coefficients, $\Delta\phi_B$ is a modified balance mass variable derived from relative vorticity as described in Subsection 2.5.5.2. This formulation is similar to that previously used in ECMWF (Derber and Bouttier 1999). The regression coefficients are computed statistically using the NMC method (Parrish and Derber 1992) with 24/48-hour forecast differences to enable estimation of the total covariances for each total spectral coefficient.

2.5.5.2 Modified Balance Mass Variable

The geostrophic balance is well kept at mid-levels in the troposphere in the extratropics. In other areas, the balance is weak. To incorporate consideration of these relationships, a modified balance mass variable is introduced. The statistical relationships linking relative vorticity, divergence, temperature, and surface pressure are calculated. First, singular value decomposition of the linear balance operator L is conducted.

$$\Delta\tilde{\phi}_B = L\Delta\zeta = UWV^{\mathrm{T}}\Delta\zeta \tag{2.5.11}$$

where $\Delta\tilde{\phi}_B$ is the original balance mass variable, W is a positive semi-definite diagonal matrix, and U and V are orthogonal matrices. The decomposed modes depend on latitude (i.e., a singular mode with a small singular value has a large amplitude in the low latitudes). Each wave number component of L is denoted as

$$\begin{aligned}\Delta\tilde{\phi}_{Bn}^{\ m} &= c_n^m\Delta\zeta_{n-1}^m + c_{n+1}^m\Delta\zeta_{n+1}^m \qquad ((n,\,m) \neq (0,\,0),\, n = m,\, m+1,\ldots,\, N),\\ c_n^m &= -\frac{2\Omega a^2}{n^2}\sqrt{\frac{n^2-m^2}{4n^2-1}}, \qquad c_{N+1}^m = 0, \qquad \Delta\tilde{\phi}_{B0}^{\ 0} = 0\end{aligned} \tag{2.5.12}$$

here Ω is the angular velocity of the earth, a is the earth's radius, n is the total wavenumber, and m is the zonal wavenumber. Second, the coefficient of regression between mass variable $\Delta\Phi$ [4] (as derived from temperature and surface pressure) and balance mass variable is calculated as follows

$$D_j = \frac{\left\langle \left(U^{\mathrm{T}}\Delta\Phi\right)_j^m \left(U^{\mathrm{T}}\Delta\tilde{\phi}_B\right)_j^m \right\rangle}{\left\langle \left[\left(U^{\mathrm{T}}\Delta\tilde{\phi}_B\right)_j^m\right]^2 \right\rangle} \tag{2.5.13}$$

where $\langle\ \rangle$ denotes statistical, zonal-wavenumber, and vertical-level mean, D_j denotes a positive definite diagonal matrix, and j denotes an index of singular vectors in latitudinal wave numbers, respectively. The regression coefficients D_j $(0-1)$ indicate the extent to which geostrophic balance is satisfied. Modified balance mass variable $\Delta\phi_B$ is then constructed as follows:

$$\Delta\phi_B = UDU^{\mathrm{T}}\Delta\tilde{\phi}_B = UDWV^{\mathrm{T}}\Delta\zeta = \tilde{L}\Delta\zeta \tag{2.5.14}$$

Note that the modified balance operator $\tilde{L}$ is based on 1) conversion from the spectral space to the singular vector space, 2) the product of the regression coefficients D, and 3) conversion from the singular vector space to the spectral space. The correlation between the modified mass variable and the unbalanced mass variable (i.e. $\Delta\Phi - \Delta\phi_B$) can be neglected in all regions including the tropics.

[4] The mass variable $\Delta\Phi_k$ is defined as $\Delta\Phi_k = \Delta\phi_k + R_d\bar{T}_k\Delta p_k/\bar{p}$, where subscript k is the vertical level index, $\Delta\phi_k$ is the geopotential height, $\bar{T}_k$ is the reference (global mean) temperature, $\bar{p}$ is the reference (global mean) pressure at ground surface, Δp_k is the pressure, and R_d is the dry gas constant. In the calculation of $\Delta\phi_k$, $\bar{T}_k$ and $\bar{p}$ are also used and some approximation is done.

2.5.5.3 Regression Coefficients for $\Delta\eta_U$ and $(\Delta T_U,\ \Delta P_{sU})$

The regression coefficient matrices P, Q, and R are calculated for each total wavenumber n as follows:

$$P_n = \left\langle \Delta\eta_n^m \left(\Delta\phi_{B_n}{}^m\right)^{\mathrm{T}} \right\rangle \left\langle \Delta\phi_{B_n}^m \left(\Delta\phi_{B_n}{}^m\right)^{\mathrm{T}} \right\rangle^{-1} \tag{2.5.15}$$

$$Q_n = \left\langle \begin{pmatrix} \Delta T_n^m \\ \Delta p_{s_n}{}^m \end{pmatrix} \left(\Delta\phi_{B_n}{}^m\right)^{\mathrm{T}} \right\rangle \left\langle \Delta\phi_{B_n}^m \left(\Delta\phi_{B_n}{}^m\right)^{\mathrm{T}} \right\rangle^{-1} \tag{2.5.16}$$

$$R_n = \left\langle \left[\begin{pmatrix} \Delta T_n^m \\ \Delta p_{s_n}{}^m \end{pmatrix} - Q_n \Delta\phi_{B_n}^m \right] \left(\Delta\eta_{U_n}{}^m\right)^{\mathrm{T}} \right\rangle \left\langle \Delta\eta_{U_n}^m \left(\Delta\eta_{U_n}{}^m\right)^{\mathrm{T}} \right\rangle^{-1} \tag{2.5.17}$$

where $\langle\ \rangle$ denotes the statistical and zonal-wavenumber mean.

2.5.5.4 Background Error Covariance Matrix

The background error covariance matrices of the control variables are calculated for each total wavenumber n, and the matrix size is equivalent to the number of vertical levels for $\Delta\zeta$, $\Delta\eta_U$, and $\Delta \ln q$ or the number of vertical levels +1 for $(\Delta T_U,\ \Delta P_{sU})$.

$$B_{\zeta\, n} = \left\langle \Delta\zeta_n^m \overline{\Delta\zeta_n^m}^{\mathrm{T}} \right\rangle, \qquad B_{\eta_U\, n} = \left\langle \Delta\eta_{U_n}{}^m \overline{\Delta\eta_{U_n}{}^m}^{\mathrm{T}} \right\rangle \tag{2.5.18}$$

$$B_{\begin{pmatrix} T_U \\ p_{sU} \end{pmatrix}_n} = \left\langle \begin{pmatrix} \Delta T_{U_n}{}^m \\ \Delta p_{sU_n}{}^m \end{pmatrix} \overline{\begin{pmatrix} \Delta T_{U_n}{}^m \\ \Delta p_{sU_n}{}^m \end{pmatrix}}^{\mathrm{T}} \right\rangle, \qquad B_{\ln q\, n} = \left\langle \Delta \ln q_n^m \overline{\Delta \ln q_n^m}^{\mathrm{T}} \right\rangle \tag{2.5.19}$$

where $\langle\ \rangle$ denotes the statistical and zonal-wavenumber mean, and overline means complex conjugate. The total variances of the control variables are rescaled by a factor of 0.8836.

2.5.5.5 Cholesky Decomposition of Background Error Covariance Matrix

The background error covariance matrix mentioned above is decomposed using the Cholesky decomposition. This gives independent and normalized (i.e., preconditioned) control variables Δy_n^m as follows:

$$J_n^{(x)} = \sum_{m=-n}^{n} \frac{1}{2} \left(\overline{\Delta x_n^m}\right)^{\mathrm{T}} B_n^{-1} \Delta x_n^m = \sum_{m=-n}^{n} \frac{1}{2} \left(\overline{\Delta x_n^m}\right)^{\mathrm{T}} \left(L_n L_n^{\mathrm{T}}\right)^{-1} \Delta x_n^m = \sum_{m=-n}^{n} \frac{1}{2} \left(\overline{\Delta y_n^m}\right)^{\mathrm{T}} \Delta y_n^m \tag{2.5.20}$$

$$\Delta y_n^m \equiv L_n^{-1} \Delta x_n^m \tag{2.5.21}$$

where $J_n^{(x)}$ is a background error term for the control variable x at the total wavenumber n, B_n is a background covariance matrix for x, and L_n is a lower triangular matrix.

In summary, normalized control variables $\Delta y_n^m(k)$ are completely independent and normalized based on background error variance. The background term of the cost function is simplified as a summation of the square of the normalized control variables.

2.5.6 LETKF

The specifications of the LETKF approach used in the GA to provide flow-dependent background error covariances are listed in Table 2.5.1.

Observation datasets assimilated in the LETKF are the same as those of GA (Table 2.1.1) except for those of hyperspectral sounders (AIRS, IASI and CrIS).

Table 2.5.1: Specifications of the LETKF used in GA

Data cut-off time	same as in the 4D-Var
First guess	3–9-hour forecast of its own
Horizontal resolution	Same as for the Inner step of the 4D-Var
Vertical resolution (model top)	Same as for the 4D-Var
Ensemble size	100 members
Analysis variables	Same as in the 4D-Var
Observation	Same as in the 4D-Var except for AIRS, IASI and CrIS
Assimilation window	6 hours (± 3 hours of analysis time)
Covariance inflation	Adaptive multiplicative covariance inflation
Initialization	Horizontal divergence adjustment based on the analysis of surface pressure tendency (Hamrud *et al.* 2015)
Localization	Gaussian function. The localization scale for which the localization function is $1/\sqrt{e}$ is set to 400 km in the horizontal domain (300 km for humidity-sensitive observations), a 0.6 scale height in the vertical domain (1.2 for surface pressure and ground-based GNSS zenith total-delay observations) and three hours in the temporal domain. For satellite radiance observations, the maximum of the square of the weighting function divided by its peak value and the Gaussian function with a $0.6\sqrt{2}$ scale height centered at the peak of the weighting function is used as the vertical localization function.
Model ensemble method	Stochastic physics scheme
Other characteristics	A total of 100 analyses are re-centered so that their ensemble mean is consistent with the hybrid 4D-Var.

Observation localization is applied in the LETKF. Observation errors are multiplied by the inverse of the localization function to give less weight to data collected farther from the analysis grid point. The localization function is given as the Gaussian function of the distance between the analysis grid point and the observation location. The localization scale l for which the localization function is $1/\sqrt{e}$ is set to 400 km in the horizontal domain (300 km for humidity-sensitive observations), a 0.6 scale height in the vertical domain (1.2 for surface pressure and ground-based GNSS zenith total-delay observations) and 3 hours in the temporal domain. The tail of the localization function is set to 0 farther than $2\sqrt{10/3}l$. For satellite radiance observations, the maximum of a square of the weighting function divided by its peak value and the Gaussian function with a length scale of $0.6\sqrt{2}$ scale height centered at the peak of the weighting function is used as the vertical localization function.

Multiplicative covariance inflation (Anderson 2001) is applied to first-guess (6-hour forecast from the previous analysis) ensembles. Inflation coefficients are estimated for each analysis grid point so that the following relation (Desroziers *et al.* 2005) is observed using locally assimilated observations:

$$tr\left[\mathbf{d}_{A-B}\mathbf{d}^{\mathbf{T}}{}_{O-B}\right] = \alpha tr\left[\mathbf{HBH}^{\mathbf{T}}\right] \tag{2.5.22}$$

where $\mathbf{d}_{A-B}$, $\mathbf{d}_{O-B}$, $\mathbf{H}$ and $\mathbf{B}$ represent the analysis increment projected onto the observation space, innovations, observation operator and background error covariance, respectively. tr [] represents the trace of the matrix, and α represents the inflation coefficient. Based on (2.5.22), the raw inflation coefficient is estimated on analysis grid point j as

$$\alpha_{j,raw} = \frac{\sum_{m=1}^{N_{obs}} \rho_{m,j}\left(d_{A-B,m}d_{O-B,m}/\sigma_{o,m}^2\right)}{\sum_{m=1}^{N_{obs}}\left[\rho_{m,j}^2\left(\mathbf{H}_m\mathbf{X}_b\right)\left(\mathbf{H}_m\mathbf{X}_b\right)^{\mathbf{T}}/\left(K-1\right)/\sigma_{o,m}^2\right]} \tag{2.5.23}$$

where N_{obs} is the number of observations assimilated locally, K is the ensemble size, $\rho_{m,j}$ is the localization function of the mth observation and $\mathbf{H}_m\mathbf{X}_b$ is the first-guess perturbation projected onto the mth observation. The subscript *raw* represents a raw estimate. $d_{A-B,m}$ and $d_{O-B,m}$ are the observational increment and innovation of the mth observation, where $d_{A-B,m}$ is computed with a transformation matrix derived from LETKF analysis

on grid j. Temporal smoothing is applied to the estimated inflation coefficient as

$$\alpha_{i,j} = \frac{\alpha_{i-1,j}\sigma^2_{o,j} + \alpha_{i,j,raw}\sigma^2_b}{\sigma^2_{o,j} + \sigma^2_b} \tag{2.5.24}$$

where the subscript i represents the value at the ith analysis step, and $\sigma^2_{o,j}$ and σ^2_b are error variances of the estimated and prior coefficients and are set as

$$\sigma^2_{o,j} = 1/\sum_{m=1}^{N_{obs}} \rho_{m,j} \tag{2.5.25}$$

$$\sigma^2_b = 0.005 \tag{2.5.26}$$

Finally, the estimated coefficients are relaxed to the default values as the deviation from these defaults reaches $1/e$ at 10 days. The defaults are set to 1.21 from the surface to 45 hPa and linearly reduced with the logarithm of pressure to 1.0 at 0.85 hPa. Inflation to specific humidity is reduced by 30% of the estimated inflation coefficients below 230 hPa and linearly decreases with the logarithm of pressure to about 15% at 45 hPa.

Initialization based on analysis of surface pressure tendency (Hamrud *et al.* 2015) is applied after the LETKF analysis update. The surface pressure tendency of the first-guess ensembles $(\partial p_s/\partial t)_{guess}$ is diagnosed from the continuity equation and the hydrostatic balance as

$$\left(\frac{\partial p_s}{\partial t}\right)_{guess} = -\int_0^1 \nabla \cdot \left(\boldsymbol{v}_{guess}\frac{\partial p_{guess}}{\partial \eta}\right) d\eta \tag{2.5.27}$$

where p_s, $\boldsymbol{v}$, p and η are surface pressure, horizontal wind, pressure and the model vertical coordinates (1 at the bottom and 0 at the top), respectively. Surface pressure tendency is analyzed by adding the above variable to the first-guess state variables of the LETKF. The difference in surface pressure tendency diagnosed from horizontal wind and surface pressure analysis and that observed from LETKF analysis is distributed to each model layer so that the value is proportional to the global mean of the analysis spread of horizontal wind. Horizontal wind is adjusted using

$$\Delta\left[\nabla \cdot (\boldsymbol{v}_k dp_k)\right] = w_k\left[\left(\frac{\partial p_s}{\partial t}\right)_{diag} - \left(\frac{\partial p_s}{\partial t}\right)_{anl}\right] \tag{2.5.28}$$

where k is an index of the vertical model layer, w_k is the weight on the kth model layer and dp_k is the difference in half-level pressure adjacent to the kth full-level model layer. The horizontal wind increment is derived from (2.5.28) assuming $\Delta(dp_k) = 0$ and $\Delta(\nabla \times (\boldsymbol{v}_k dp_k)) = 0$.

The analysis ensemble is recentered so that the ensemble mean is consistent with global analysis.

The stochastic physics scheme (Subsection 3.3.4) is also applied to the forecast ensemble of the LETKF.

2.5.7 Observation Terms

2.5.7.1 Observation Data

The assimilated observation types are shown in Table 2.1.1, and brief explanations for each data type as well as the quality control procedures are found in Section 2.2 and Section 2.3.

Observational data and related departures (observation minus first guess) are given with the location and time via the pre-analysis procedure. Surface pressure data at the station height or sea level pressure data reported from surface observation station are assimilated after conversion to the pressure at model surface height. While scatterometer data are reported as winds at 10 m above sea level, assimilated as wind at the lowest model level. The zenith total delay from GNSS data are assimilated over land. Satellite radiance data from MW sounders, MW imagers, Hyperspectral IR sounders and CSRs are directly assimilated using the K matrix model of RTTOV-13.0 (Saunders *et al.* 2020). GNSS-RO data are assimilated in the form of bending angle at the tangent point using ROPP8 (Culverwell *et al.* 2015).

2.5.7.2 Observation Error

Observation errors (the diagonal part of the observation error covariance matrix) are estimated based on innovation statistics (Desroziers *et al.* 2005), and are summarized in Table 2.5.2 and Table 2.5.3. For ocean surface wind data from scatterometers, such errors are defined as those with values of 4 or 6 m/s. For GNSS-RO bending angle data, errors are defined as a function of height only. The observation error threshold is 1% of the observed bending angle above 10 km, and varies linearly from 20% at 0 km to 1% at 10 km. For ground-based GNSS-ZTD data, the observation error is 20 mm. The error at an arbitrary reported pressure level is linearly interpolated in the logarithm of pressure ($\log(p)$). The cross correlations of observation errors (off the diagonal part of the observation error covariance matrix) are not considered explicitly in 4D-Var. To eliminate the cross-correlation term in the cost function, horizontally or vertically dense observations are thinned spatially in pre-analysis, and observation errors are inflated with predefined factors.

For satellite radiance data of MW imagers and sounders assimilated under all-sky conditions, observation errors are given by an observation error model. Observation errors are assigned as a function of the symmetric (average of observation and first guess (FG)) cloud amount to handle the non-Gaussian distribution of observation errors. The symmetric cloud amount is based on Geer and Bauer (2011) using horizontal and vertical polarization difference at 37GHz for MW imagers, and Geer *et al.* (2014) using difference of scattering effect between 90GHz and 150GHz channels for MW sounders. The parameters of the observation error model are derived based on FG departure statistics as listed in Table 2.5.4. Observation error model for 183GHz channels is separately prepared for each surface types (ocean, land and sea ice). Only parameters for ocean are shown.

2.5.7.3 Observation Operator

In 4D-Var, observation data at a given location and time are simulated using forecast variables for the surrounding grids in the nearest forecast hour with spatial inter/extrapolation and variable conversion. Observation operator involves these consecutive procedures. A fast radiative transfer model known as RTTOV-13.0 is used as the observation operator for satellite radiance data assimilation. ROPP8 is used as the observation operator for the assimilation of GNSS-RO bending angle data. These operators are provided as external libraries from EUMETSAT NWP-SAF and ROM SAF, respectively.

2.5.7.4 Variational Bias Correction

As mentioned in Subsection 2.3.4, variational bias correction (VarBC) is applied to satellite radiance data biases. In 4D-Var with VarBC, the observation operators are extended to include bias correction terms and the control (analysis) variables are extended to include bias correction (regression) coefficients. These coefficients are optimized as control variables.

The extended form of the cost function Eq. (2.5.1) is defined as follows.

$$J(\Delta z_0) = \frac{1}{2}\Delta x_0^{\mathrm{T}}\mathbf{B}^{-1}\Delta x_0 + \frac{1}{2}\Delta\beta^{\mathrm{T}}\mathbf{B}_\beta{}^{-1}\Delta\beta + \frac{1}{2}\sum_{i=1}^{n}\left(\mathbf{H}_i\Delta x_i + \sum_{j=0}^{m}\Delta\beta_j p_{i,j} - d_i\right)^{\mathrm{T}}\mathbf{R}_i^{-1}\left(\mathbf{H}_i\Delta x_i + \sum_{j=0}^{m}\Delta\beta_j p_{i,j} - d_i\right) + J_C \tag{2.5.29}$$

where,

$$\Delta z_0 \equiv \left[\Delta x_0{}^{\mathrm{T}}, \Delta\beta^{\mathrm{T}}\right]^{\mathrm{T}}, \quad \mathbf{B}_\beta \equiv diag\left(\frac{{F_{inf}}^2}{N_{var}}, \ldots, \frac{{F_{inf}}^2}{N_{var}}\right), \quad N_{var} \equiv \begin{cases} \frac{N}{\log_{10}\frac{N}{N_0}+1} & (N \geq N_0) \\ N_0 & (N < N_0) \end{cases}$$

Δz_0 represents extended increments, consisting of low resolution model variable increments Δx_0 and bias correction coefficient increment $\Delta\beta$, $\mathbf{B}_\beta$ is the background error covariance matrix for the bias correction coefficient β, $p_{i,j}$ is predictors for bias correction, m is the number of predictors for all radiance observation types, F_{inf} is an arbitrarily defined inflation factor, N is the number of data and N_0 is the threshold for the valid number of data.

Table 2.5.2: Observation error tables used in operational global analysis for (a) conventional observation, (b) AMV, (c) AMSU-A, (d) ATMS and (e) CSR from four geostationary satellites. P_s, u, v, T, RH and T_B denote surface pressure, zonal and meridional wind components, temperature, relative humidity and brightness temperature, respectively. "x" in (c) denotes that the channel is not used.

(a) conventional observation

Level(hPa) \ Element	P_s(hPa)	u, v(m/s)	T(K)	RH(%)
Surface	0.7			
1,000		2.3	1.7	6.4
850		2.4	1.2	15.9
700		2.5	1.0	19.8
500		2.5	0.8	31.5
300		2.7	0.9	31.7
200		2.8	1.1	24.1
100		3.1	1.2	3.8
50		3.0	1.4	1.4
30		3.0	1.5	1.3
10		3.9	2.5	1.3
1		4.6	5.4	1.3
0.4		7.7	7.6	1.3
0.1		7.7	7.6	1.3

(b) AMV

Level(hPa) \ Element	u, v(m/s)
1,000	4.5
850	4.5
700	4.5
500	4.5
300	5.3
200	5.8
100	6.8
50	7.0
30	7.2
10	7.6
1	9.1
0.4	10.6
0.1	10.6

(c) AMSU-A T_B (K)

Channel \ Satellite	Metop-B	Metop-C	NOAA-15	NOAA-18	NOAA-19
4	x	0.45	0.45	0.45	0.45
5	x	0.3	0.3	x	0.3
6	0.3	0.3	x	0.3	0.3
7	0.3	0.3	0.3	0.3	0.3
8	0.3	0.3	0.3	x	x
9	0.3	0.3	0.3	x	0.3
10	0.3	0.3	0.45	0.3	0.3
11	0.45	0.45	x	0.45	0.45
12	0.64	0.64	0.64	0.64	0.64
13	1.02	1.02	0.85	1.02	1.02
14	2.83	2.83	x	2.63	2.63

(d) ATMS T_B (K)

Channel \ Satellite	Suomi-NPP	NOAA-20
6	0.3	0.3
7	0.3	0.3
8	0.3	0.3
9	0.3	0.3

(e) CSR T_B (K)

band (μm) \ Satellite	GOES-16	band (μm) \ Satellite	Meteosat-9	band (μm) \ Satellite	Meteosat-11	band (μm) \ Satellite	Himawari-9
8 (6.15)	1.5	5 (6.25)	1.5	5 (6.25)	1.5	8 (6.24)	1.5
9 (7.00)	1.5					9 (6.94)	1.5
10 (7.40)	1.5	6 (7.35)	1.5	6 (7.35)	1.5	10 (7.35)	1.5

Table 2.5.3: Observation error tables used in operational global analysis for hyperspectral IR sounders. (a) IASI and (b) CrIS.

(a) IASI T_B (K)						(b) CrIS T_B (K)		
Ch.	Metop-B	Metop-C	Ch.	Metop-B	Metop-C	Ch.	Suomi-NPP	NOAA-20
38	1.26	1.08	249	0.69	0.75	27	0.81	0.75
49	1.20	1.08	252	0.72	0.75	31	1.35	1.32
51	1.23	1.05	256	0.72	0.72	37	0.54	0.51
55	1.20	1.05	262	0.69	0.72	51	0.39	0.39
57	1.23	1.02	265	0.78	0.72	69	0.45	0.36
61	1.20	0.96	267	0.69	0.63	73	0.45	0.36
63	1.14	0.99	269	0.72	0.78	75	0.51	0.42
70	1.11	0.96	271	0.66	0.69	79	0.42	0.36
85	1.11	0.93	273	0.63	0.66	80	0.45	0.39
87	1.08	0.99	275	0.72	0.75	81	0.39	0.36
109	0.96	0.84	278	0.72	0.72	83	0.39	0.36
116	0.87	0.87	282	0.72	0.72	85	0.39	0.36
122	0.87	0.87	284	0.72	0.69	87	0.39	0.39
128	0.87	0.84	288	0.69	0.72	88	0.39	0.39
135	0.78	0.81	292	0.66	0.66	93	0.42	0.42
141	0.78	0.87	294	0.75	0.78	96	0.39	0.39
146	0.72	0.75	296	0.78	0.81	102	0.48	0.48
148	0.75	0.78	306	0.90	0.87	106	0.45	0.45
154	0.75	0.78	308	0.69	0.72	116	0.54	0.54
159	0.72	0.72	316	0.69	0.72	123	0.63	0.63
161	0.72	0.75	320	0.72	0.72	124	0.54	0.54
167	0.72	0.75	323	0.69	0.66	125	0.57	0.60
173	0.72	0.81	327	0.72	0.75	132	0.54	0.54
180	0.72	0.84	333	0.72	0.75	136	0.51	0.54
185	0.75	0.81	347	0.72	0.75	138	0.54	0.57
187	0.72	0.81	350	0.69	0.72	142	0.51	0.54
193	0.81	0.90	354	0.72	0.69	148	0.48	0.48
205	0.90	1.02	356	0.69	0.69	224	-	1.65
212	0.87	0.78	360	0.69	0.66	248	-	1.56
217	0.81	0.87	666	1.62	1.62	264	-	1.86
219	0.87	0.81	1110	1.65	1.65	279	-	1.50
224	0.78	0.84	2951	3.51	3.51	718	-	2.28
226	0.81	0.84	2993	4.32	4.32	978	-	3.81
230	0.75	0.78	3207	4.32	4.32	988	-	3.33
232	0.81	0.84	5130	3.03	3.03	1003	-	3.93
236	0.69	0.69	5315	1.74	1.74	1032	-	3.75
239	0.75	0.81	5397	3.09	3.09			
243	0.72	0.84	5492	3.51	3.51			
246	0.72	0.78						

Table 2.5.4: Observation error (S_{clr}, S_{cld}) and threshold of symmetric cloud amount for identifying clear-sky and cloudy samples (C_{clr}, C_{cld}) for assimilated channel under all-sky conditions.

Satellite/Sensor	Ch.	Frequency(GHz) and polarization	S_{clr}	S_{cld}	C_{clr}	C_{cld}
GCOM-W/AMSR2	7	19V	4.8	43.8	0.0	0.6
	9	23V	6.9	22.8	0.0	0.5
	11	37V	6.0	34.2	0.0	0.5
DMSP-F17/SSMIS	9	183 ± 6.6	9.4	82.2	0.0	40.0
	10	183 ± 3.0	8.0	56.1	0.0	40.0
	11	183 ± 1.0	7.9	29.2	0.0	40.0
DMSP-F17,18/SSMIS	13	19V	5.4	40.8	0.0	0.6
	14	23V	7.2	19.8	0.0	0.6
	16	37V	6.9	35.1	0.0	0.5
GPM-core/GMI	3	19V	5.1	47.1	0.0	0.6
	5	23V	6.9	26.1	0.0	0.5
	6	37V	6.6	36.6	0.0	0.5
	12	183 ± 3	6.8	75.8	0.0	40.0
	13	183 ± 7	7.1	105.0	0.0	30.0
NOAA-19,Metop-B,C/MHS	3	183 ± 1	6.1	44.0	0.0	45.0
	4	183 ± 3	5.4	71.4	0.0	45.0
	5	190V	4.9	104.1	0.0	45.0
FY-3C/MWHS-2	11	183 ± 1.0	6.9	50.1	0.0	50.0
	12	183 ± 1.8	6.6	71.0	0.0	50.0
	13	183 ± 3.0	6.5	91.6	0.0	50.0
	14	183 ± 4.5	6.9	110.0	0.0	50.0
	15	183 ± 7.0	6.6	128.2	0.0	50.0
Suomi-NPP,NOAA-20/ATMS	18	183 ± 7.0	8.2	117.0	0.0	50.0
	19	183 ± 4.5	8.1	99.9	0.0	50.0
	20	183 ± 3.0	7.7	81.1	0.0	50.0
	21	183 ± 1.8	7.3	43.6	0.0	50.0
	22	183 ± 1.0	7.4	43.6	0.0	50.0

The second term on the right hand side of Eq. (2.5.29) is the background term for the bias correction coefficients, and $\sum_{j=0}^{m} \Delta\beta_j p_{i,j}$ is the bias correction term. This equation is used instead of Eq. (2.5.1) in 4D-Var.

2.6 Meso-scale Analysis

2.6.1 Introduction

Meso-scale analysis (MA) produces initial conditions for the Meso-Scale Model (MSM, Subsection 3.5.1) every three hours with incorporation of information from observations into the model for optimal forecasting of weather phenomena with emphasis on high-impact events.

In March 2002, a four-dimensional variational (4D-Var) scheme was introduced as the data assimilation approach for MA (Ishikawa and Koizumi 2002) in place of the previous three-dimensional optimal interpolation (3D-OI) scheme to create the world's first operational limited-area 4D-Var system. Following the upgrade of the MSM forecast model to a non-hydrostatic type (JMA-NHM; Saito *et al.* 2006, 2007) in September 2004, the previous hydrostatic 4D-Var was replaced by a non-hydrostatic model-based 4D-Var system (known as JMA Nonhydrostatic model-based Variational Data Assimilation (JNoVA; Honda *et al.* 2005)) in April 2009. The domain of MA is the same as that of the MSM, covering Japan and its surrounding areas. Operation covering the current 4,080 × 3,300 km domain, extended from the previous 3,600 × 2,880 km, was introduced in March 2013 (Subsection 3.5.1). A further upgrade of the MSM forecast model in February 2017 replaced JMA-NHM with a newly developed non-hydrostatic model called ASUCA (Ishida *et al.* 2009, 2010; Hara *et al.* 2012) as described in Subsection 3.5.1. In March 2020, a 4D-Var system based on ASUCA (ASUCA-Var; Ikuta *et al.* 2021) was introduced (Ikuta *et al.* 2020). This development enabled MA to produce initial conditions more consistent with the upgraded MSM forecast model. In March 2022, the model top was upgraded from 21,801 m to 37,500 m, with the number of vertical layers increased from 76 to 96 in the outer model and from 38 to 48 in the inner model.

Various observational data are used to improve the accuracy of prediction for meso-scale weather events, including information from weather radars, satellite observations and ground-based GNSS (Table 2.1.2). Thanks to the advanced data assimilation scheme of 4D-Var utilized with these data, MA produces initial conditions highly consistent with the balance of model equations.

2.6.2 Operational System

MA is performed using ASUCA-Var system and produces initial conditions for MSM forecasts at 00, 03, 06, 09, 12, 15, 18 and 21 UTC. The data assimilation window is set as three hours, with the end of the window corresponding to the analysis time. Figure 2.6.1 shows the MA process:

1. With the initial condition produced in the previous MA, the high-resolution (5 km) forecast model is run within the three-hour data assimilation window to obtain the first guess.

2. Quality-control for observation (Section 2.3) and calculation of related deviations from the first guess are conducted as preparation for input to the next step.

3. 4D-Var analysis is performed to determine the optimal model state by assimilating observations in a low-resolution (15 km) space.

4. The low-resolution (15 km) analysis increment is added to the high-resolution (5 km) first guess via interpolation to determine the initial condition for the next step.

5. Based on this initial condition, the high-resolution (5 km) forecast model is run within the data assimilation window to determine the initial condition for the MSM.

In MA, the first and last steps in which the high-resolution forecast model is run are called the outer steps, and that in which ASUCA-Var in the low-resolution space is executed is called the inner step. ASUCA is used as the forecast model for both steps. The analysis domain is shown in Figure 2.6.2 with a topographic map (5-km resolution) as used by MA. The lateral boundary conditions are from Global Spectral Model (GSM, Section 3.2) forecasts, while the initial conditions of the first guess are taken from the previous MA (the three-hour forecast in the last outer step). Thus, MA frames the cycle analysis nested into the GSM.

The cut-off time of input observation data for MA is 50 minutes after each analysis time. Observation data received by this time are distributed into four time slots with the observation time rounded to the nearest hour (as represented by the four stars in Figure 2.6.1). Accordingly, data observed from 4 hours before to 0.5 hours after analysis are assimilated in the inner step.

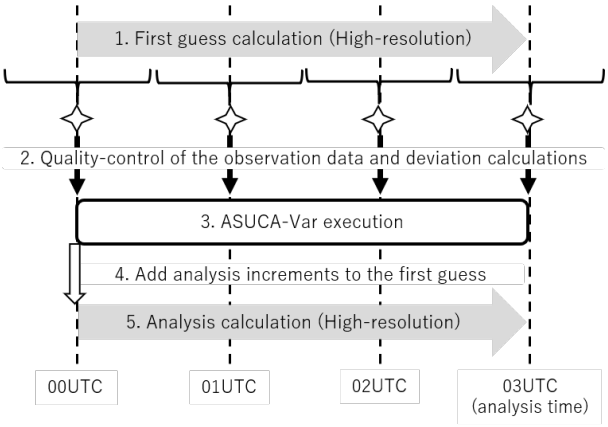


Figure 2.6.1: MA procedure (03 UTC analysis)

Figure 2.6.2: MA domain and topography

As described previously, ASUCA-Var in the inner step involves a data assimilation system based on a four-dimensional variational (4D-Var) method as detailed in Subsection 2.6.3. This approach is based on maximum likelihood estimation, with optimal values (i.e. analysis fields) determined by minimizing the cost function (Subsection 2.6.3.1 for details). This minimization requires iterative calculation of the cost function and its gradient (50 times at maximum), which takes considerable computational resources. To reduce this burden, operational ASUCA-Var involves an incremental approach (Courtier *et al.* 1994) in which a model with a low-resolution relative to that used in the outer step is adopted in minimization. The basic field in this approach is updated with the low-resolution nonlinear model twice (20 and 35 steps). Operational formulation of minimization with this approach is detailed in Subsection 2.6.3.1. The model used in minimization for ASUCA-Var is called the inner model, with the specifications described in Subsection 2.6.3.3. The horizontal grid spacing is 5 km (817 × 661 grid points) with 96 vertical layers for atmosphere in the outer steps, and 15 km (273 × 221 grid points) with 48 vertical layers in the inner step.

2.6.3 Basic Formulation

2.6.3.1 Cost Function

In the MA system, 4D-Var data assimilation is used to determine the optimal model trajectory in a phase space by minimizing its deviation from observations and the first guess. Deviation is measured using the cost function J, defined as

$$\begin{aligned} J(\boldsymbol{x}_0) &= J_b + J_o + J_{vbc} + J_p \\ &= \frac{1}{2}\left(\boldsymbol{x}_0 - \boldsymbol{x}_0^b\right)^T \mathbf{B}_0^{-1}\left(\boldsymbol{x}_0 - \boldsymbol{x}_0^b\right) + \sum_{t=0}^{N} \frac{1}{2}\left(H_t(\boldsymbol{x}_t) - \boldsymbol{y}_t + P(\boldsymbol{b})\right)^T \mathbf{R}_t^{-1}\left(H_t(\boldsymbol{x}_t) - \boldsymbol{y}_t + P(\boldsymbol{b})\right) \\ &\quad + J_{vbc} + J_p, \end{aligned} \tag{2.6.1}$$

where the superscript T represents transpose.

The first and second terms of Eq. (2.6.1) are background and observation terms, representing deviations from the first guess and observation, respectively. $\boldsymbol{x}_0$ is the model state at the beginning of the data assimilation

window (time level $t = 0$) to be optimized[5], $\boldsymbol{x}_0^b$ is the first guess of the model state at $t = 0$, $\boldsymbol{y}_t$ is a column vector consisting of observational data available at t ($t = 0, ..., N$), and $\boldsymbol{x}_t$ is the model state at t as forecast from the initial condition $\boldsymbol{x}_0$ as

$$\boldsymbol{x}_t = M_t(\boldsymbol{x}_0) \,, \tag{2.6.2}$$

where M_t denotes the forecast operator. H_t is an observation operator used to convert model state variables to observations, typically consisting of conversion from model variables to observation parameters and interpolation from model grid points to observation points. $P(\boldsymbol{b})$ is observation bias. The error covariance matrixes $\mathbf{B}_0$ and $\mathbf{R}_t$ specify the error profiles (uncertainty and error correlation) of $\boldsymbol{x}_0^b$ and $\boldsymbol{y}_t$, respectively (Subsection 2.6.3.2 and Subsection 2.6.4.2). Observation terms are modified via the introduction of variational quality control (Subsection 2.6.4.5).

The third term of Eq. (2.6.1) J_{vbc} is for variational bias correction (VarBC; Dee 2004) to estimate observation bias. It is given as

$$J_{vbc} = \frac{1}{2}\left(\boldsymbol{b} - \boldsymbol{b}^b\right)^T \mathbf{B}_{\text{vbc}}^{-1}\left(\boldsymbol{b} - \boldsymbol{b}^b\right), \tag{2.6.3}$$

where $\boldsymbol{b}$ is a control variable for bias correction, $\boldsymbol{b}^b$ is the first guess of $\boldsymbol{b}$, and $\mathbf{B}_{\text{vbc}}$ is the background error covariance matrix for VarBC, which is used to correct for satellite brightness temperature. VarBC background error is defined from Cameron and Bell (2018) as

$$B_{\text{vbc}} = \frac{\sigma_o^2}{N_b}, \tag{2.6.4}$$

$$N_b = \max(m_{\text{avg}}, m_{\text{min}})\left(\frac{1}{2^{1/n} - 1}\right), \tag{2.6.5}$$

where σ_o is the observation error standard deviation, m_{avg} is the average number of observations including a particular channel assimilated over the last three days, m_{min} is the lower limit of observation numbers (500), and n is the bias halving time (set as 8).

The fourth term of Eq. (2.6.1) J_p is a penalty term based on digital filter to suppress high-frequency noise mostly caused by gravity waves (Gauthier and Thépaut 2001), given as

$$J_p = \frac{1}{2}\left(\boldsymbol{x}_{N/2} - \overline{\boldsymbol{x}}_{N/2}\right)^T \mathbf{B}_{\text{df}}^{-1}\left(\boldsymbol{x}_{N/2} - \overline{\boldsymbol{x}}_{N/2}\right) \tag{2.6.6}$$

where $\boldsymbol{x}_{N/2}$ is the model state analysis increment at the center of the data assimilation window ($t = N/2$), $\overline{\boldsymbol{x}}_{N/2}$ is the digitally filtered analysis increment at $t = N/2$, and $\mathbf{B}_{\text{df}}$ is the background error covariance matrix for digital filtering. Based on Wee and Kuo (2004), $\mathbf{B}_{\text{df}}$ is given as

$$\mathbf{B}_{\text{df}}^{-1} = \frac{\lambda}{\text{diag}(\mathbf{B}_0)}, \tag{2.6.7}$$

where λ is the factor $1.0{\times}10^{-6}$. For a time series of model states over the data assimilation window $\{\boldsymbol{x}_0, ..., \boldsymbol{x}_N\}$, the digitally filtered state $\overline{\boldsymbol{x}}$ at $t = N/2$ is given as

$$\overline{\boldsymbol{x}}_{N/2} \quad = \quad \sum_{k=0}^{N} h_{N/2-k} W_k \boldsymbol{x}_k \,, \tag{2.6.8}$$

where

$$h_k \quad = \quad \frac{\sin k\theta_c}{k\pi} \,, \tag{2.6.9}$$

is a low-pass filter that removes time oscillations exceeding the cutoff frequency θ_c. The Dolph-Chebyshev window function W_k (Lynch 1997) is also used to suppress the noise from the Fourier truncation (Gibbs oscillation).

[5] Lateral boundary conditions over the data assimilation window can be included in the vector to be optimized, $\boldsymbol{x}_0$, but this is not adopted in operational MA.

In MA, an incremental approach is adopted to reduce computational cost. Optimization is performed using the inner model (Subsection 2.6.3.3) to determine an analysis increment for the low-resolution model space (inner step). The cost function is expanded around the basic field, which can be expressed as

$$\begin{aligned} J(\delta \boldsymbol{x}_0) = & \frac{1}{2} \delta \boldsymbol{x}_0{}^T \mathbf{B}_0^{-1} \delta \boldsymbol{x}_0 \\ & + \sum_{t=0}^{N} \frac{1}{2} \Big(\mathbf{H_t} \mathbf{M_t} \delta \boldsymbol{x}_0 + \mathbf{P} \delta \boldsymbol{b} - \mathbf{d_t} \Big)^T \mathbf{R}_t^{-1} \Big(\mathbf{H_t} \mathbf{M_t} \delta \boldsymbol{x}_0 + \mathbf{P} \delta \boldsymbol{b} - \mathbf{d_t} \Big) \\ & + J_{vbc} + J_p, \end{aligned} \tag{2.6.10}$$

where

$$\delta \boldsymbol{x}_0 = \boldsymbol{x}_0 - \boldsymbol{x}_0^b, \tag{2.6.11}$$

$$\delta \boldsymbol{b} = \boldsymbol{b} - \boldsymbol{b}_0^b, \tag{2.6.12}$$

$$\boldsymbol{d}_t = \boldsymbol{y}_t - H_t(M_t(\boldsymbol{x}_0^b)) - P(\boldsymbol{b}^b), \tag{2.6.13}$$

and $\mathbf{H_t}$, $\mathbf{M_t}$, and $\mathbf{P}$ are tangent-linear operators of H_t, M_t, and P, respectively.

The effect of the non-linear process was insufficiently incorporated in analysis values with linear optimization alone. To address this problem, a basic-field update (Trémolet 2008) was introduced to incorporate non-linear effects into optimization. Details of the update in ASUCA-Var are provided in *Section 2.1* of Ikuta *et al.* (2021). The field is updated twice during optimization.

The final analysis $\boldsymbol{x}_N$ is determined from forecasting with the high-resolution model over the data assimilation window (outer step).

$$\boldsymbol{x}_N = M_N(\boldsymbol{x}_0) \ . \tag{2.6.14}$$

2.6.3.2 Background Error Covariance

As detailed previously, the background error covariance $\mathbf{B}$ indicates the error profile of the first guess (Subsection 2.6.3.1). However, calculation using the complete form of $\mathbf{B}$ is impractical due to the extremely large dimensions of the model state space. In practice, drastic simplification is applied to $\mathbf{B}$ to make the problem tractable.

A group of parameters are defined as analysis variables, with background errors considered uncorrelated. The variables used in MA are

- u: x component of horizontal wind
- v: y component of horizontal wind
- $\big(T_g, p_s, \theta\big)$: underground and skin temperature, surface pressure and potential temperature
- $\big(W_g, \mu_p\big)$: soil moisture and pseudo-relative humidity ($\mu_p = q_v / q_{sat}^b$, where q_v is the mixing ratio of water vapour and q_{sat}^b is the saturated water vapour of the first guess).

T_g is perturbed only at land grids, as that at ocean grids is not a prognostic variable in the MSM.

$\delta \boldsymbol{x}_0$ is transformed from uncorrelated control variables χ as

$$\delta \boldsymbol{x}_0 = \boldsymbol{x}_0 - \boldsymbol{x}_0^b = \mathbf{B}^{1/2} \chi, \tag{2.6.15}$$

$\mathbf{B}^{1/2}$ is the square root of $\mathbf{B}$, and has the form

$$\mathbf{B}^{1/2} = \mathbf{K}_p \begin{pmatrix} \mathbf{B}_u^{1/2} & 0 & 0 & 0 \\ 0 & \mathbf{B}_v^{1/2} & 0 & 0 \\ 0 & 0 & \mathbf{B}_{T_g, p_s, \theta}^{1/2} & 0 \\ 0 & 0 & 0 & \mathbf{B}_{W_g, \mu_p}^{1/2} \end{pmatrix}. \tag{2.6.16}$$

Here, $\mathbf{K}_p$ is a linearized parameter transform from control variables to model variables, and $\mathbf{B}_{\chi_i}^{1/2}$ is the spatial transform for the i-th sub-group χ_i of the parameters in χ. The spatial structure of background error covariance $\mathbf{B}_{\chi_i}$ is modeled as

$$\mathbf{B}_{\chi_i}^{1/2} = \mathbf{C}_v \mathbf{B}_{h,\chi_i}^{1/2} \mathbf{B}_{v,\chi_i}^{1/2}. \tag{2.6.17}$$

Here, $\mathbf{C}_v$ denotes a vertical coordinate transformation to limit the terrain effect of the vertical coordinate within the lower troposphere.

The square root of the vertical background error covariance matrix $\mathbf{B}_{v,\chi_i}$ is given by

$$\mathbf{B}_{v,\chi_i}^{1/2} = \mathrm{diag}\left(\mathbf{B}_{v,\chi_i}\right)^{1/2} \mathbf{V}\mathbf{\Lambda}^{1/2}\mathbf{V}^T \,, \tag{2.6.18}$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues of $\left[\mathrm{diag}\left(\mathbf{B}_{v,\chi_i}\right)^{-1/2} \mathbf{B}_{v,\chi_i} \mathrm{diag}\left(\mathbf{B}_{v,\chi_i}\right)^{-1/2}\right]$, and $\mathbf{V}$ is the orthogonal matrix ($\mathbf{V}^T\mathbf{V} = \mathbf{I}$) whose columns are the related eigenvectors.

The horizontal background error correlation $\mathbf{B}_{h,\chi_i}$ is defined on the vertical level of the model. In the calculation of $\mathbf{B}_{h,\chi_i}^{1/2}$, the recursive filter technique (Purser *et al.* 2003) is used in each of the x- and y-directions.

The background error statistics $\mathbf{B}_{v,\chi_i}$ and $\mathbf{B}_{h,\chi_i}$ are estimated using the NMC method (Parrish and Derber 1992). Vertical background error covariance is calculated separately for sea and land grids (except u and v) and classified by local time.

Variational optimization is performed with respect to χ. The transform (2.6.15), called preconditioning, simplifies the background term of the cost function J (see Eq. (2.6.1) and Eq. (2.6.10); for simplicity, the present discussion does not deal with J_{vbc} and J_p). The cost function and its gradient after the transform are given as

$$J(\chi) = \frac{1}{2}\chi^T\chi + \sum_{t=0}^{N} \frac{1}{2}\left(\mathbf{H_t}\mathbf{M_t}\mathbf{B}^{1/2}\chi - \boldsymbol{d}_t\right)^T \mathbf{R}_t^{-1}\left(\mathbf{H_t}\mathbf{M_t}\mathbf{B}^{1/2}\chi - \boldsymbol{d}_t\right), \tag{2.6.19}$$

$$\nabla_\chi J = \chi + \sum_{t=0}^{N} \mathbf{B}^{T/2}\mathbf{M}_t^T\mathbf{H}_t^T\mathbf{R}_t^{-1}\left(\mathbf{H}_t\mathbf{M}_t\mathbf{B}^{1/2}\chi - \boldsymbol{d}_t\right), \tag{2.6.20}$$

where $\mathbf{M}_t^T$ and $\mathbf{H}_t^T$ are the adjoint model and the adjoint of the observation operator.

2.6.3.3 Inner Model

In ASUCA 4D-Var, a high-resolution with the same configuration as the MSM (5 km horizontal grid spacing, 96 vertical layers), is used in the outer step. The inner step is executed using the nonlinear model (NLM), tangent-linear model (TLM) and adjoint model (ADM), with a lower resolution (15 km horizontal grid spacing, 48 vertical layers) to reduce computational cost. The specifications of the inner and outer models are listed in Table 2.6.1.

The NLM is essentially the same as the outer model except for the resolution and the configuration of convective parameterization. The dynamics of the TLM are fully tangent-linearized, although some physical processes are simplified or not implemented to avoid errors associated with non-linearity.

In regard to cloud microphysics, saturation adjustment is linearized but other processes are not. For convective parameterization, the NLM adopts the modified Kain-Fritsch scheme (Kain and Fritsch 1990) but the TLM does not. For the boundary layer scheme, Mellor-Yamada-Nakanishi-Niino level 3 (Nakanishi and Niino 2006) is employed, and is linearized for the TLM except for diffusion coefficients and partial condensation. The surface flux scheme based on Beljaars and Holtslag (1991) is linearized for the TLM except for bulk coefficients. The TLM Radiation scheme is simplified based on Mahfouf (1999).

2.6.4 Observation Terms

2.6.4.1 Observation Data

Assimilated observation types are shown in Table 2.1.2, and brief outlines of each data type and related quality control procedures are given in Section 2.2 and Section 2.3.

Table 2.6.1: Specifications of the outer and inner models employed in MA

	Outer model	Inner NLM	Inner TLM/ADM
Resolution	5 km, 96 layers	15 km, 48 layers	15 km, 48 layers
Cloud microphysics	6-class 3-ice bulk scheme, Subsection 3.5.4	6-class 3-ice bulk scheme, Subsection 3.5.4	Only saturation adjustment process tangent-linearized
Convection	Modified Kain-Fritsch	Modified Kain-Fritsch	None
Boundary layer	Mellor-Yamada-Nakanishi-Niino level-3	Mellor-Yamada-Nakanishi-Niino level-3	Tangent-linearized except for diffusion coefficients and partial condensation
Surface flux	Beljaars and Holtslag	Beljaars and Holtslag	Tangent-linearized except for bulk coefficients
Radiation	Subsection 3.5.6	Subsection 3.5.6	Mahfouf (1999)

2.6.4.2 Observation Error

The observation error covariance matrix **R** in Eq. (2.6.1) is assumed to be diagonal. Estimation of observation errors (diagonal components of **R**) is based on innovation statistics (Desroziers *et al.* 2005). Errors for conventional observations, wind profiler data and AMVs are summarized in Table 2.6.2. Errors for satellite radiance are the same as those in global analysis (Table 2.5.2(c) - (j)). Errors for GNSS-PWV and radial velocity are 3 mm and 3.3 m/s, respectively. Errors for relative humidity data from DPR are 5%, and those from ground-based radar are approximately 30 %. Errors for ocean surface wind data from scatterometers are 3 m/s. For GNSS-RO refractivity data, observation errors are defined as a function of height only. Observation error is calculated using linear interpolation to be 10% of refractivity at 0 km, 0.5% at 10 km, 0.18% at 20 km and 0.04% at 30 km. Errors for Radar/Raingauge Analyzed Precipitation (R/A) are based on the precipitation amount (Koizumi *et al.* 2005). The error at an arbitrary reported pressure level is linearly interpolated in the logarithm of pressure ($\log(p)$). The cross-correlations of errors between different observations are not considered explicitly in 4D-Var. To eliminate consideration of cross-correlation terms in the cost function, dense observations are thinned spatially and observation errors are inflated in pre-analysis.

2.6.4.3 Observation Operator

The RTTOV-13.0 fast radiative transfer model is used as the observation operator for satellite radiance data assimilation, while ROPP8 for the assimilation of refractivity data from GNSS-RO. These operators are provided as external libraries from EUMETSAT NWP-SAF and ROM SAF, respectively.

2.6.4.4 Special Treatment for Precipitation Data

For the observation terms of the cost function in Eq. (2.6.1), a Gaussian probability density function (PDF) for observation errors is assumed. However, as the PDF for precipitation-amount data is not Gaussian, the following observation term is used for one-hour precipitation-amounts (Koizumi *et al.* 2005):

$$J_o{}^{PREC}(x) = \sum_{j\ (\text{where } r_j^o \geq 0.5)}^{n} \frac{\left(H_j(x) - r_j^o\right)^2}{2\sigma_o (r_j^o)^2}. \tag{2.6.21}$$

Here, $H_j(x)$ is an observation operator used to convert the state variable x to one-hour cumulative precipitation values at the j-th grid point, r_j^o is precipitation observed at the grid point, and n is the number of grid points in

Table 2.6.2: Observation error tables used in the operational meso-scale analysis for (a) conventional observation and wind profiler data and (b) AMV. P_s, u, v, T and RH denote surface pressure, x and y wind components in an MSM Lambert projection space, temperature and relative humidity respectively.

(a) Conventional observation and wind profiler data					
Element Level (hPa)	P_s (hPa)	u (m/s)	v (m/s)	T (K)	RH (%)
Surface	0.7				
	0.6 for SYNOP in Japan				
1,000		2.1	1.9	1.3	9.8
925		2.0	1.9	0.9	10.3
850		2.0	2.0	0.9	12.7
700		2.0	1.9	0.9	12.8
500		1.9	1.9	0.7	12.9
400		2.2	2.2	0.7	13.3
300		2.6	2.6	0.9	13.5
250		2.7	2.6	1.0	14.4
200		2.7	2.6	1.1	13.7
150		2.6	2.6	1.1	16.6
100		3.2	3.0	1.5	15.1
70		3.7	3.1	1.9	13.6
50		3.2	2.8	1.9	12.1
30		3.2	2.8	1.9	11.8
10		3.2	2.8	1.9	12.2

(b) AMV		
Element Level (hPa)	u (m/s)	v (m/s)
1,000	4.1	3.3
850	2.9	2.3
700	3.2	2.6
500	3.7	3.0
300	4.6	3.7
200	3.8	4.9
100	4.4	6.0
50	3.5	5.1
30	5.1	6.2
10	6.2	7.2

the inner model[6]. $\sigma_o(r_j^o)$ is the observation error standard deviation, defined as

$$\sigma_o(r_j^o) \equiv \begin{cases} C_{sat}\max(r_{min}, r_j^o) & \left(H_j(x) \le r_j^o\right) \\ C_{sat}C_a\max(r_{min}, r_j^o) & \left(H_j(x) > r_j^o\right) \end{cases}, \begin{cases} C_a = 3,\ C_{sat} = 1 & \text{for R/A} \\ C_a = 5,\ C_{sat} = 2 & \text{for satellite retrievals} \end{cases}, \quad r_{min} \equiv 1\text{mm/h}. \tag{2.6.22}$$

Here, C_{sat} is an observation error inflation factor for satellite retrievals, and C_a is a tuning factor for the asymmetric structure of the departure frequency distribution around 0.

One-hour precipitation observation values less than 0.5 mm are not assimilated, since the quality of such data is rather poor for snowfall. The observation error of satellite retrievals is considered to be larger than that of R/A because retrieval is from instantaneous observation rather than from one-hour cumulative observation.

2.6.4.5 Variational Quality Control

Variational quality control (VarQC, Andersson and Järvinen 1999) is applied in 4D-Var for conventional observations. With VarQC, the observation error PDF is assumed to be a summation of a Gaussian function and a positive constant value within a certain range. This constant represents the probability of rough error within the range.

The following observation term and its gradient are used for conventional observations in the cost function Eq. (2.6.1) in 4D-Var with VarQC:

$$j_o^{VarQC} = -\log\left(\frac{\gamma + \exp(-j_o)}{\gamma + 1}\right), \quad \gamma \equiv \frac{A\sqrt{2\pi}}{(1-A)\,2d} \tag{2.6.23}$$

$$\nabla j_o^{VarQC} = W^{VarQC}\nabla j_o, \quad W^{VarQC} \equiv 1 - \frac{\gamma}{\gamma + \exp(-j_o)}. \tag{2.6.24}$$

[6] If one-hour cumulative precipitation P_{1h} is larger than 1 mm/h, $\left(3P_{1h}^{1/3} - 2\right)$ is used for $H_j(x)$ and r_j^o.

Here, A is the prior probability of rough error (e.g., 0.05 for SYNOP), d is the maximum standard deviation below which rough error is possible (e.g., 9 for SYNOP), j_o^{VarQC} is the observation term for a single observation component with VarQC, and j_o is the term without VarQC.

Eq. (2.6.24) shows that ∇j_o^{VarQC} is almost the same (effective) as ∇j_o when j_o is small ($W^{VarQC} \approx 1$) and ∇j_o^{VarQC} is almost 0 (not effective) when j_o is large ($W^{VarQC} << 1$). Observation values satisfying $W^{VarQC} < 0.25$ are regarded for rejection in VarQC.

2.7 Local Analysis

2.7.1 Introduction

Local Analysis (LA) produces initial conditions for the Local Forecast Model (LFM) (Section 3.7) at a horizontal resolution of 2 km. Its operation started in August 2012, with eight runs per day on an area of Japan measuring 2,200 × 2,500 km to initialize LFM forecasts over a domain covering the eastern part of the country. An enhancement in the operation of LA, along with the LFM, was implemented in May 2013, extending its domain to cover Japan and the surrounding areas (3,160 × 2,600 km) and increasing its daily operations to 24 runs per day.

To provide initial conditions for this high-resolution forecast model targeting small-scale severe weather events, LA is designed to allow rapid production and frequent updating of analysis at a resolution of 5 km (Subsection 2.7.2). In each LA run, an analysis cycle with hourly three-dimensional variational (3D-Var) data assimilation is executed for the previous three hours to incorporate information from newly received observational data in each case. The analysis cycle was originally based on JMA-NHM (Saito *et al.* 2006, 2007) and JNoVA 3D-Var (the 3D-Var version of JNoVA (Honda *et al.* 2005)), which was replaced by the new-generation version based on ASUCA (Ishida *et al.* 2009, 2010; Hara *et al.* 2012) and ASUCA-3DVar in January 2015 (Aranami *et al.* 2015). In March 2022, the analysis cycle was updated to hybrid 3D-Var utilizing ensemble perturbation of Meso-Scale Ensemble Prediction System (MEPS) data (Section 3.6) (Yokota *et al.* 2022).

As with MA, high-density remote sensing data (including information from weather radars and ground-based GNSS) are assimilated on an hourly basis in LA as important sources of detailed information contributing to better forecasting of high-impact phenomena (see Table 2.1.3). LA was also adopted in January 2017 to make extensive use of satellite observations , with application of the variational bias correction technique. The capacity of high-resolution NWP to capture small-scale variations in topography is expected to help reduce representativeness errors in the assimilation of surface observations. In association, LA also assimilates automated surface station (AMeDAS) data ahead of other operational data assimilation systems with lower resolutions in order to appropriately reflect the effects of local-scale environments near the surface.

Based on these features, LA is characterized as a data assimilation system for high-resolution and high-frequency NWP.

2.7.2 Operational System

To satisfy the requirements outlined in Subsection 2.7.1, operational LA incorporates an analysis cycle with 3D-Var, which can meet demand for prompt and frequent product updates within a limited time frame (all processes, including data quality control, are completed within around 15 minutes) with far fewer computational resource requirements than 4D-Var. In 3D-Var, the weighted average of flow-dependent and climatological background error covariances (hybrid covariance) is used (hybrid 3D-Var), and the former is created from ensemble perturbations of MEPS.

LA involves the running of an analysis cycle on a domain identical to that of the LFM (see Figure 2.7.1), following the flow chart shown in Figure 2.7.2. The cycle consists of four successive instances of hybrid 3D-Var analysis (3, 2, 1 and 0 hours prior to the initial time of the LFM; (a), (b), (c) and (d) in Figure 2.7.2), where quality controlled observation data rounded to the nearest hour are assimilated. After each instance of analysis except the last one ((d) in Figure 2.7.2), a one-hour forecast (LF1) with a horizontal resolution of 5 km is executed using analysis for initial conditions, thereby providing the first guess for the next 3D-Var analysis.

In the first 3D-Var ((a) in Figure 2.7.2), the first guess is supplied from the MSM (Section 3.5), which also provides lateral and upper boundary conditions for LF1 throughout the three-hour data assimilation period. Sea surface temperature is given by HIMSST (see Subsection 5.2.2). The analyzed field in the last 3D-Var ((d) in Figure 2.7.2) is used as the initial condition of the LFM. Further LA specifications are provided in Table 2.1.3.

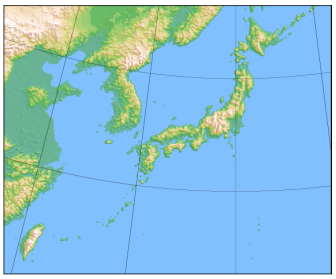


Figure 2.7.1: domains of LA and LFM.

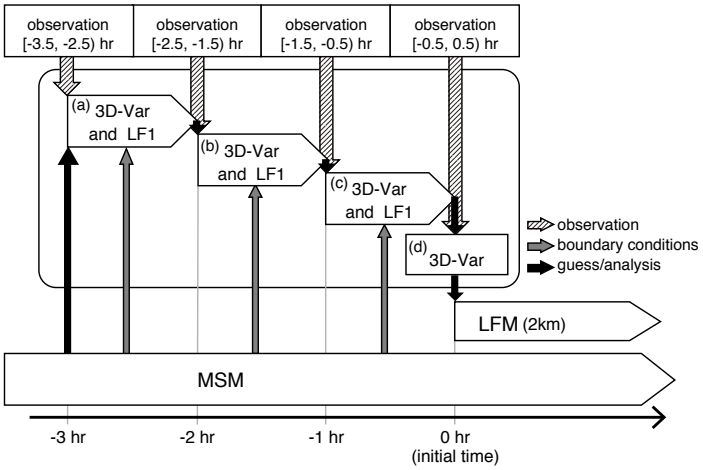


Figure 2.7.2: Schematic representation of LA analysis.

2.7.3 Basic Formulation

2.7.3.1 Cost Function

LA involves four time slots for the assimilation of observations using hybrid 3D-Var, in which the extended control variable method (Lorenc 2003) is used to take into account flow-dependent error covariance. The related cost function in the i-th time slot is defined as

$$J(\delta \boldsymbol{x}_i) = \frac{1}{2}\delta \boldsymbol{x}_{i,cl}^T \mathbf{B}_{cl}^{-1} \delta \boldsymbol{x}_{i,cl} + \frac{1}{2}\delta \boldsymbol{x}_{i,en}^T \mathbf{B}_{en}^{-1} \delta \boldsymbol{x}_{i,en} \tag{2.7.1}$$

$$+\frac{1}{2}\left(\mathbf{H}\delta \boldsymbol{x}_i - \boldsymbol{d}_i + \mathbf{P}\delta \boldsymbol{b}_i\right)^T \mathbf{R}^{-1} \left(\mathbf{H}\delta \boldsymbol{x}_i - \boldsymbol{d}_i + \mathbf{P}\delta \boldsymbol{b}_i\right) \tag{2.7.2}$$

$$+\frac{1}{2}\delta \boldsymbol{b}_i^T \mathbf{S}_i^{-1} \delta \boldsymbol{b}_i, \tag{2.7.3}$$

$$\delta \boldsymbol{x}_i = \beta_{cl}\delta \boldsymbol{x}_{i,cl} + \beta_{en}\delta \boldsymbol{x}_{i,en}. \tag{2.7.4}$$

Here, $\delta \boldsymbol{x}_i$ is the increment, $\delta \boldsymbol{b}_i$ is the increment of the observation bias vector, and the subscript i is the hourly time-index. The innovation vector $\boldsymbol{d}_i$ is given by

$$\boldsymbol{d}_i = \boldsymbol{y}_i - \mathcal{H}\left(\boldsymbol{x}_i^b\right) - \mathcal{P}\left(\boldsymbol{b}_i^b\right), \tag{2.7.5}$$

where $\boldsymbol{y}_i$ is the vector of observations, $\boldsymbol{x}_i^b$ is the background state and $\boldsymbol{b}_i^b$ is the background state of observation bias in the i-th time slot. $\mathcal{H}$ is the nonlinear observation operator, $\mathbf{H}$ is the tangent linearized observation operator, $\mathcal{P}$ is the nonlinear operator of predictors for VarBC, $\mathbf{P}$ is the matrix of predictors for VarBC , $\mathbf{B}$ is the covariance matrix of background error, $\mathbf{R}$ is the diagonal matrix of observation error and $\mathbf{S}_i$ is the diagonal matrix of parameters used to control the adaptivity of observation bias vector estimation in VarBC. The subscripts

"*cl*" and "*en*" mean "climatological" and "ensemble", respectively. β_{cl} and β_{en} are the weights for the hybrid covariances, set as $(\beta_{cl}^2, \beta_{en}^2) = (0.5, 0.5)$.

2.7.3.2 Solution Procedure

The analysis $\boldsymbol{x}^a$ at the initial time t of LFM, corresponding to $i = 4$, is calculated by repeating hybrid 3D-Var and one-hour forecasting. The one-hour forecast operator $\mathcal{M}_{i,i+1}$ is configured specifically for LA. In this configuration, the horizontal resolution is set to 5 km as in the MSM, but the physics schemes differ from those used in the MSM and the LFM.

The LA cycle is conducted as outlined below.

1. $\mathbf{B}_{en}$ valid at $t-3\text{h}, t-2\text{h}, t-1\text{h}, t$ are created from MEPS ensemble perturbations.

2. $\boldsymbol{x}_{i=1}^b$ valid at $t-3\text{h}$ is provided by the MSM, and $\boldsymbol{b}_{i=1}^b$ is equal to the analysis variables of observation bias $\boldsymbol{b}^a$ in the previous LA.

3. $\delta\boldsymbol{x}_i$ is optimized by minimization of the cost function, $J(\delta\boldsymbol{x}_i)$.

4. The i-th background state is updated with the one-hour forecast:

$$\boldsymbol{x}_{i+1}^b = \mathcal{M}_{i,i+1}\left(\boldsymbol{x}_i^b + \delta\boldsymbol{x}_i\right), \tag{2.7.6}$$

   and the observation bias is given by

$$\boldsymbol{b}_{i+1}^b = \boldsymbol{b}_i^b + \delta\boldsymbol{b}_i. \tag{2.7.7}$$

5. Steps 3 and 4 are repeated three times.

6. For the initial time, the cost function $J(\delta\boldsymbol{x}_{i=4})$ is minimized and the analysis increments are added to the background state. The low-resolution (5 km) analysis variables are given by

$$\boldsymbol{x}^a = \boldsymbol{x}_4^b + \delta\boldsymbol{x}_4, \tag{2.7.8}$$

   and the analysis variables of observation bias $\boldsymbol{b}^a$ are given by

$$\boldsymbol{b}^a = \boldsymbol{b}_4^b + \delta\boldsymbol{b}_4. \tag{2.7.9}$$

Finally, the low-resolution analysis $\boldsymbol{x}^a$ is interpolated to the high resolution of 2 km with consideration of ancillaries (i.e., topography, soil type and land usage) to be used as the initial condition with the LFM.

2.7.3.3 Analysis Variables

The analysis increment is defined as $\delta\boldsymbol{x} = \left(\delta u, \delta v, \left(\delta T_g, \delta p_s, \delta\theta\right), \left(\delta W_g, \delta\mu_p\right)\right)^T$.

- u: x-component of horizontal wind.
- v: y-component of horizontal wind.
- $\left(T_g, p_s, \theta\right)$: underground and skin temperature, surface pressure and potential temperature.
- $\left(W_g, \mu_p\right)$: soil moisture and pseudo-relative humidity.

T_g is perturbed only at land grids, as that at ocean grids is not a prognostic variable in the LFM.

2.7.3.4 Background Error Covariance

$\delta \boldsymbol{x}$ is transformed from uncorrelated control variables χ, as follows:

$$\delta \boldsymbol{x} = \mathbf{B}^{1/2}\chi = \begin{bmatrix} \beta_{cl}\mathbf{B}_{cl}^{1/2} & \beta_{en}\mathbf{B}_{en}^{1/2} \end{bmatrix} \begin{bmatrix} \chi_{cl} \\ \chi_{en} \end{bmatrix}, \tag{2.7.10}$$

$\mathbf{B}^{1/2}$ is the square root of $\mathbf{B}$.

The climatological background error covariance $\mathbf{B}_{cl}$ is modeled as described in Subsection 2.6.3.2, and statistics are estimated using the NMC method (Parrish and Derber 1992). However, the error profiles in the lower levels are modified artificially to localize spatial correlations so that surface observations are assimilated more appropriately. The seasonal variation of background error statistics is not taken into account.

The ensemble background error covariance $\mathbf{B}_{en}$ is created from 100 ensemble perturbations using 5 lagged MEPS forecasts with 20 members, and spatial localizations with Gaussian functions (scales of $1/\sqrt{e}$ are set as 100 km horizontally and 0.5 km vertically) to reduce sampling errors. $\mathbf{B}_{en}$ is inflated by multiplication with a factor, which is the ratio of $\mathbf{B}_{cl}$ and the horizontal mean of $\mathbf{B}_{en}$ for potential temperature at 5.5 km above ground level, meaning that error variance is comparable to the magnitude of $\mathbf{B}_{cl}$.

2.7.4 Observation Terms

2.7.4.1 Observation Data

Assimilated observation types and brief outlines of each data type are provided in Table 2.1.3.

2.7.4.2 Observation Error

The observation error covariance matrix is assumed to be diagonal, and cross-correlation between different observations is not considered as it is in MA.

2.7.4.3 Observation Operators

The observation operator for surface observations (i.e., wind at a height of 10 m and temperature, relative humidity at a height of 1.5 m) is based on the surface diagnostic scheme (Beljaars and Holtslag 1991) in the LFM. In the tangent linear operator for surface observation, perturbations of transfer coefficients for surface fluxes are not considered in diagnostic equations.

The observation operator for brightness temperature is RTTOV-13.0 (Saunders *et al.* 2020), but perturbations of surface elements and ozone are ignored as inputs to the RTTOV tangent-linear model.

The observation operator for soil moisture is provided using a linear regression equation with coefficients estimated from cumulative distribution function matching.

2.7.5 Parallelization

The data assimilation domain is two-dimensionally decomposed into blocks, each of which is assigned to an MPI process. The loops for the y-direction and certain fused horizontal loops are forked via OMP parallelization similar to that of the LFM (see Section 3.5.10). Observations are distributed to blocks according to their location, and innovations are calculated in each process. In addition, the recursive filter method is parallelized using decomposed blocks.

In calculation of minimization, the inner product of general vectors in the L-BFGS (Nocedal and Wright 2006) is determined for each block, and loops for observations and model variables are forked via OMP parallelization.

2.8 Snow Depth Analysis

2.8.1 Global Snow Depth Analysis

Global snow depth analysis is executed every six hours with data at the same resolution as Global Analysis with cycle runs and early runs (TQ959: grid approx. 0.125°(13 km); Table 2.1.1) . Analysis incorporates forecast snow depth fields, satellite snow-cover data and SYNOP snow depth data with two-dimensional optimal interpolation (2D-OI).

The satellite snow-cover product is retrieved in the 0.25° × 0.25° latitude/longitude grid from microwave imager radiances[7] with the same scheme as used for JRA-55 (Kobayashi *et al.* 2015). In the algorithm, regions where the observation satisfies the following criteria are considered to be covered by snow[8]:

$$T_b\,(19GHz, \mathrm{V}) < 265K$$

$$T_b\,(19GHz, \mathrm{H}) - T_b\,(37GHz, \mathrm{H}) > \mathrm{threshold}\,(\mathrm{H})$$

$$T_b\,(19GHz, \mathrm{V}) - T_b\,(37GHz, \mathrm{V}) > \mathrm{threshold}\,(\mathrm{V})$$

where T_b is the brightness temperature, H and V are the horizontal and vertical channels, respectively. Thresholds (H, V) should be 0 K for ideal snow but in practice need to be adjusted depending on the snow and surface condition. Accordingly, those values are determined for each region, vegetation, and month. For grid with no satellite radiance for not more than 3 days, past snow-cover product is used. For grid with no satellite radiance more than 3 days, missing value is used.

The first guess snow depth fields used in 2D-OI are estimated mainly from the satellite snow-cover product (SSCP) and the forecast snow depth as below.

1. The SSCP is interpolated into gridding for the GSM.
2. The SSCP is modified under certain conditions. Where SSCP is zero or missing, the climatological probability of snow (CPS) exceeds 80 % and snow is forecast, the CPS value is applied. Where SSCP is otherwise missing, the CPS value or zero is applied depending on the forecast land surface temperature.
3. If there is snow in both the forecast and the satellite, the forecast snow depth is used as the first guess.
4. If there is no snow in satellite data and the forecast snow depth exceeds 5 cm, the first guess is set to the forecast snow depth multiplied by exp(−0.1).
5. If there is no snow in satellite data and the forecast snow depth is equal to or less than 5 cm, the first guess is set to 0 cm.

This first guess and observations reported from SYNOP stations are handed over to the 2D-OI system for the snow depth analysis[9]. The spatial correlation coefficient for the background error, μ_{kl}, is based on Brasnett (1999) and given by

$$\mu_{kl} = \alpha(r_{kl})\beta(\Delta z_{kl})$$

with the horizontal and vertical separation, r_{kl} and Δz_{kl} between points k and l. $\alpha(r_{kl})$ and $\beta(\Delta z_{kl})$ are the horizontal and vertical structure functions:

$$\alpha(r_{kl}) = \left(1 + \frac{r_{kl}}{L}\right)\exp\left(-\frac{r_{kl}}{L}\right),$$

$$\beta(\Delta z_{kl}) = \exp\left\{-\left(\frac{\Delta z_{kl}}{h}\right)^2\right\},$$

where L and h are set to 100 km and 800 m, respectively. The standard deviation ratio of observation error to background error is set as 0.75. Grids on ocean are skipped in snow depth analysis. Grids where there is no

[7] SSMIS (DMSP F-17, F-18) and AMSR2 (GCOM-W1) are used.
[8] There are some typographical errors in Kobayashi *et al.* (2015).
[9] SYNOP reports from Japan are not used in 2D-OI. Analysis for grid points over the country is described later in this section.

snow in both the forecast and the satellite are also skipped and the analysis are set to 0 cm. Grids where the forecast ground surface temperature is above the freezing point and there is very few snow only in the forecast or the satellite are also skipped and the analysis are set to 0 cm. Grids covering land ice are also skipped and the analysis are set to the climatological values.

Grid point values for analysis over Japan are replaced by inverse distance-weighted interpolation of the latest observations reported by SYNOP and AMeDAS stations.

Snow depths are converted to snow water equivalents as an initial condition for land-surface processing (see Subsection 3.2.10) in the GSM.

2.8.2 Mesoscale Snow Depth Analysis

In the Meso Scale Model (MSM, Section 3.5), snow cover data are used to gauge the status of ground snow, with depths exceeding 5 cm in individual grid squares being classified as snow-covered. Snow depth analysis data are produced via a two-dimensional OI (2D-OI) in the high-resolution snow depth analysis system. The first guesses for 2D-OI are set using an offline version of the land surface model (LSM) with the same domain and grid spacing as the MSM. The offline LSM and the 2D-OI are outlined below.

The offline LSM, which includes a multi-layer snowpack model, simulates typical snow processes such as accumulation, compaction and ablation. The atmospheric forcing data necessary to drive the LSM are air temperature and wind velocity at the lowest atmospheric model level and radiative fluxes toward the surface as predicted by the MSM. Radar/Raingauge Analyzed Precipitation data (see Subsection 4.4.1) are used as rain and snowfall inputs to the LSM.

The model first guesses and observations reported from SYNOP and AMeDAS stations are handed over to the 2D-OI system in the snow depth analysis system. The methodology of OI is based on Brasnett (1999), where the correlation coefficient, μ_{kl}, is given by

$$\mu_{kl} = \alpha(r_{kl})\beta(\Delta z_{kl}) \tag{2.8.1}$$

with the horizontal and vertical separation, r_{kl} and Δz_{kl} between points k and l. $\alpha(r_{kl})$ and $\beta(\Delta z_{kl})$ are the horizontal and vertical structure functions:

$$\alpha(r_{kl}) = \left(1 + \frac{r_{kl}}{L}\right)\exp\left(-\frac{r_{kl}}{L}\right), \tag{2.8.2}$$

$$\beta(\Delta z_{kl}) = \exp\left\{-\left(\frac{\Delta z_{kl}}{h}\right)^2\right\}, \tag{2.8.3}$$

where L and h are set to 25 km and 500 m, respectively. The standard deviations of observation and background errors are set at 4 and 3 cm, respectively.

2.9 Soil Moisture Analysis

Global soil moisture analysis is performed as part of Global Analysis (Table 2.1.1) every six hours. Soil moisture fields for the first three soil layers are analyzed via a simplified Extended Kalman Filter (EKF) (Hess 2001, de Rosnay *et al.* 2013) using first-guess forecasts of soil moisture and analyzed screen-level temperature and relative humidity.

In this EKF, the analyzed state vector $\mathbf{x_a}$ at time t_i is calculated for each grid point as

$$\mathbf{x_a}(t_i) = \mathbf{x_b}(t_i) + \mathbf{K_i}\left[\mathbf{y_o}(t_i) - H_i(\mathbf{x_b})\right] \tag{2.9.1}$$

where $\mathbf{x_b}$, $\mathbf{y_o}$ and H_i are the first-guess state vector, observation state vector and nonlinear observation operator, respectively. In global soil moisture analysis, values (i.e., degrees of saturation) for the first three soil levels (from the surface to a depth of 0.19 m) correspond to $\mathbf{x_a}$ and $\mathbf{x_b}$. Analyzed screen-level temperature and relative humidity are treated as $\mathbf{y_o}$. The Kalman gain matrix $\mathbf{K_i}$ is calculated as

$$\mathbf{K_i} = \left[\mathbf{B^{-1}} + \mathbf{H_i}^{\mathbf{T}}\mathbf{R^{-1}}\mathbf{H_i}\right]^{-1}\mathbf{H_i}^{\mathbf{T}}\mathbf{R^{-1}} \tag{2.9.2}$$

where **B** and **R** are background-error and observation-error covariance matrices, respectively. Following de Rosnay *et al.* (2013), these are static and diagonal matrices composed of error variances for simplification. The terms are based on soil moisture standard deviation $\sigma_b = 0.02\,\mathrm{m}^3/\mathrm{m}^3$, and screen-level parameter standard deviations of $\sigma_T = 1\,\mathrm{K}$ for temperature and $\sigma_{RH} = 4\,\%$ for relative humidity. $\mathbf{H_i}$ is a linearized observation operator approximated from finite differences between perturbed and unperturbed forecasts based on a low-resolution (TL319L128) version of GSM2103 (Table 3.1.1).

The analyzed screen-level parameters are derived from two-dimensional optimal interpolation (2D-OI) assimilating SYNOP observations. The spatial correlation coefficient for background error, μ_{kl}, is based on Drusch and Viterbo (2007) as

$$\mu_{kl} = \alpha(r_{kl})\beta(\Delta z_{kl}) \tag{2.9.3}$$

with horizontal and vertical separation r_{kl} and Δz_{kl} between points k and l. $\alpha(r_{kl})$ and $\beta(\Delta z_{kl})$ are the horizontal and vertical structure functions:

$$\alpha(r_{kl}) = \exp\left\{-\frac{1}{2}\left(\frac{r_{kl}}{L}\right)^2\right\}, \tag{2.9.4}$$

$$\beta(\Delta z_{kl}) = \exp\left\{-\left(\frac{\Delta z_{kl}}{h}\right)^2\right\}, \tag{2.9.5}$$

where L and h are 200 km and 800 m, respectively. The standard deviations of background and observation errors are 1.5 and 2 K for temperature and 5 and 10 % for relative humidity. In Eq. (2.9.1), the analysis increments of screen-level parameters are used as background departures $\mathbf{y_o}(t_i) - H_i(\mathbf{x_b})$.

2.10 Non-real-time Quality Control

2.10.1 GDPFS-RSMC Operational Activities

JMA is designated as a Regional Specialized Meteorological Center (RSMC) of the World Meteorological Organization (WMO) Global Data-processing and Forecast System (GDPFS), and is known in this role as RSMC Tokyo. In March 1991, WMO Commission for Basic Systems (CBS) designated RSMC Tokyo as a lead center for monitoring the quality of land surface observations in Region II (Asia). As a part of its operational activities, JMA produces a six-monthly report containing a consolidated list of stations suspected of reporting low-quality observation data on station level pressure, mean sea level pressure and geopotential height during the six-month periods ending June and December. This report is available on JMA's website[10].

RSMC Tokyo also produces monthly statistics on the quality of all observations received in time for use in its final global analyses. Copies of these reports are provided to major GDPFS centers and to the WMO Secretariat. The reports are also available on JMA's website[11].

Data quality evaluation is based on differences between observations and first guess fields (three to nine-hour forecasts) from the global model. Standard procedures and formats for the exchange of monitoring results are given in the Manual on GDPFS (WMO-No.485).

2.10.2 WDQMS Operational Activities

The WMO Integrated Global Observing System (WIGOS) is a framework for all WMO observing systems and WMO contributions to co-sponsored observing systems in support of all WMO Programmes and activities. At WIGOS workshops on Quality Monitoring and Incident Management held in December 2014 and December 2015, plans were developed for a WIGOS Data Quality Monitoring System (WDQMS). The Task Team on WDQMS (TT-WDQMS) is working to develop the WDQMS under the Inter-Commission Coordination Group on WIGOS. Four NWP centers (European Centre for Medium-Range Weather Forecasts (ECMWF), National

[10] http://qc.kishou.go.jp/clsf.html
[11] http://qc.kishou.go.jp/mmr.html

Centers for Environmental Prediction (NCEP), JMA and Deutscher Wetterdienst (DWD)) contribute to the NWP Quality Monitoring Pilot Project on WDQMS, providing monitoring output in near-real time to the WMO Secretariat. Contributions began with surface pressure data, and now also include information on surface humidity, wind, temperature and upper-air soundings.

2.10.3 Blacklist Management

As mentioned in Section 2.3, low quality observational data can result in significant forecast degradation. The cause of low quality may be instrumental failure, which can continue for a long time. Such observation data should be excluded in the first step of QC, and a blacklist is kept to meet this need. Blacklist management is one of the most important activities in QC. The quality of all observations is evaluated based on differences between observations and first guess fields from the global model (three to nine-hour forecasts), the meso-scale model (zero to three-hour forecasts) and the local forecast model (one-hour forecasts). Providers of problematic observation data are added to the blacklist.

2.11 Climate Data Assimilation System

2.11.1 Introduction

Climate system monitoring and seasonal prediction require accurate and consistent comprehension based on high-quality, spatio-temporally homogeneous long-term data. To this end, JMA has to date developed the Japanese 25-year Reanalysis (JRA-25; Onogi *et al.* 2007) dataset in conjunction with the Central Research Institute of Electric Power Industry, the Japanese 55-year Reanalysis (JRA-55; Kobayashi *et al.* 2015) dataset and the Japanese Reanalysis for Three Quarters of a Century (JRA-3Q; Kosaka *et al.* 2024) dataset. At present, JRA-3Q data are used for real-time climate monitoring, seasonal prediction and various R&D activities at JMA and in wider communities (e.g., fundamental academic fields such as meteorology, climatology and oceanography) and for practical application in fields such as agricultural meteorology and renewable energy.

2.11.2 JRA-55

JMA's second reanalysis (JRA-55, completed in 2013) covers the period from 1958, when regular global radiosonde observation began. Production is continued to the present on a near-real-time basis for operational climate data assimilation as described in Kobayashi *et al.* (2015). The system is based on the TL319 version of the JMA global NWP system as of December 2009 (JMA 2007, 2013), and features numerous improvements made since JRA-25 (Table 2.11.1). These include a revised longwave radiation scheme, a 4D-Var data assimilation system and a variational bias correction scheme for satellite radiances. The system also incorporates new observational datasets produced from efforts to improve the quality of past observation data, including homogenization of radiosonde temperature observations (Haimberger *et al.* 2008, 2012) and reprocessing of satellite data from major meteorological satellite centers (e.g., van de Berg *et al.* 2002; Oyama 2010).

These enhancements make JRA-55 data considerably superior to those of JRA-25. The major JRA-25 issues of a cold bias in the lower stratosphere and a dry bias in the Amazon basin are mitigated, and the temporal consistency of temperature analysis is considerably better than in previous reanalysis products.

2.11.3 JRA-3Q

The latest Japanese reanalysis (JRA-3Q, completed in 2022) covers the period from September 1947, when Typhoon Kathleen caused extreme flooding in Japan. Production is continued to the present on a near-real-time basis for operational climate data assimilation. The system is based on the TL479 version of the JMA global NWP system as of December 2018 (JMA 2019), and incorporates developments in the operational NWP system, boundary conditions and forcing fields achieved at JMA since JRA-55 (Table 2.11.1). In JMA's GSM, for example, biases in radiation budgets, surface sensible/latent heat fluxes and precipitation are significantly reduced due to extensive improvement of parameterizations for physical processes. Boundary conditions over

the ocean applied from June 1985 onward have also been replaced with satellite-based high-resolution (0.25-degree) sea surface temperature (SST) data (MGDSST; Kurihara *et al.* 2006, Subsection 5.2.1) to improve the representation of atmospheric processes around western boundary currents and associated SST fronts. JRA-3Q also incorporates rescued historical observations to temporally extend analysis to around 10 years earlier than JRA-55.

The significant JRA-55 issue of global energy imbalance with excess upward net energy flux at the top of the atmosphere and at the surface is significantly reduced in JRA-3Q. The former's trend toward artificial weakening of TCs is also resolved via the generation of TC bogus based on JMA's operational system.

Table 2.11.1: Data assimilation in JRA-55 and JRA-3Q

	JRA-55	JRA-3Q
Analysis period	1958 – 2023/early 2024	September 1947 onward
Base system	JMA operational system as of December 2009 (JMA 2007, 2013)	JMA operational system as of December 2018 (JMA 2019)
Horizontal resolution	TL319 (– 55 km)	TL479 (– 40 km)
Vertical levels	Surface and 60 levels up to 0.1 hPa (Iwamura and Kitagawa 2008; Nakagawa 2009)	Surface and 100 levels up to 0.01 hPa (Kawai *et al.* 2013)
Analysis scheme	4D-Var with T106 inner resolution	4D-Var with TL319 resolution (Kadowaki and Yoshimoto 2012)
Radiosonde temperature bias correction	Until 2006: RAOBCORE V1.4 (Haimberger *et al.* 2008) From 2007 onward: RAOBCORE V1.5 (Haimberger *et al.* 2012) • Based on comparison with ERA	RISE (RICH with solar elevation dependence) v1.7.2 (Haimberger *et al.* 2012) • Based on comparison with surrounding stations • Seasonally dependent (from 1979 onward)
Radiative transfer model for satellite radiances	RTTOV-9.3 (Saunders 2008)	RTTOV-10.2 (Saunders *et al.* 2012) • Improved accuracy • Inclusion of GHGs variations
Land surface analysis	Offline SiB	Cycle of land surface forecasts from modeling
SST and sea ice	COBE-SST (1 degree) (Ishii *et al.* 2005, Subsection 5.2.3)	Until May 1985: COBE-SST2 (1 degree) (Hirahara *et al.* 2014, Subsection 5.2.3) From June 1985 onward: MGDSST (0.25 degrees) (Kurihara *et al.* 2006, Subsection 5.2.1)
Ozone	Until 1978: Climatology From 1979 onward: MRI-CCM1 (T42L68) (Shibata *et al.* 2005)	MRI-CCM2.1 (TL159L64) (Deushi and Shibata 2011) • Produced with the new model for the whole period

Chapter 3

Numerical Weather Prediction Models

3.1 Summary

JMA operates NWP models to meet various kinds of requirements on weather forecasting. The suite of the NWP models covers a wide temporal range of forecast periods from a few hours to two seasons providing a seamless sequence of products for the public. The following is a brief description of the major NWP models.

1. The Global Spectral Model (GSM) produces 132-hour forecasts four times a day (00, 06, 12 and 18 UTC) to support short-range forecasting (up to three days ahead) and tropical cyclone forecasts and to provide lateral boundary conditions for the Meso-Scale Model (MSM). The GSM forecasts at 00 and 12 UTC are extended to 264 hours (11 days) to support one-week forecasting. The specifications of the GSM are shown in Table 3.1.1 and a description is given in Section 3.2.

2. The Global Ensemble Prediction System (GEPS) produces seamless short to extended-range forecasts up to 34 days ahead with uncertainty for applications such as tropical cyclone forecasts, one-week forecasts and one-month forecasts as well as issuance of Early Warning Information on Extreme Weather. In addition, the Seasonal EPS system produces seven-month forecasts to support three-month forecasts, warm- and cold-season outlooks and El Niño outlooks. The specifications and other details of the GEPS are outlined in Section 3.3, and the coupled atmosphere-ocean general circulation model used in the Seasonal EPS system is described in Section 3.4.

3. The MSM produces 78-hour forecasts (00 and 12 UTC) and 39-hour forecasts (03, 06, 09, 15, 18 and 21 UTC) to support disaster prevention, very short-range precipitation forecasts and aviation forecasts, and also provides lateral boundary conditions for the Local Forecast Model (LFM). The specifications of the MSM are shown in Table 3.1.2, and a description is given in Section 3.5.

4. The Meso-scale Ensemble Prediction System (MEPS) produces 39-hour forecasts four times a day (00, 06, 12 and 18 UTC) to provide uncertainty of the MSM prediction. The specifications and other details of the MEPS are outlined in Section 3.6.

5. The LFM produces ten-hour forecasts 24 times a day on the hour to support aviation forecasts, disaster prevention and very short-range precipitation forecasts. Its specifications are shown in Table 3.1.3, and a description is given in Section 3.7.

JMA operates a global atmospheric transport and dispersion model (Section 3.8) to support its RSMC activities for nuclear environmental emergency response. The model is executed on request in coordination with the World Meteorological Organization (WMO).

JMA also operates three Chemical Transport Models (CTMs; Section 3.9 and Section 3.10). The aerosol CTM produces 96-hour forecasts of atmospheric distribution for airborne Kosa (Aeolian dust), which arrives seasonally from the Eurasian Continent. The global CTM produces 120-hour forecasts of Ultra Violet (UV)

index distribution, and the regional CTM produces 51-hour forecasts of photochemical oxidants. These models are run at 12 UTC daily.

The operational verification procedure is outlined in Section 3.11. An atmospheric transport model (Section 3.12) is operated to create information on the extent and movement of volcanic ash.

Table 3.1.1: Specifications of Global Spectral Model (GSM)

System	
Model (version)	Global Spectral Model (GSM2303)
Date of implementation	December 1987 (Latest version: 14 March 2023)
Configuration	
Horizontal resolution (Grid spacing)	Spectral triangular truncation 959 with a quadratic and reduced Gaussian grid system (TQ959), roughly equivalent to 0.125 ×0.125° (13 km) in latitude and longitude
Vertical resolution (model top)	128 stretched sigma pressure hybrid levels (0.01 hPa)
Forecast length (initial time)	132 hours (06 and 18 UTC) and 264 hours (00 and 12 UTC)
Coupling to ocean / wave / sea ice models	–
Integration time step	300 seconds
Initial conditions	
Data assimilation	Hybrid Four-dimensional variational (4D-Var) method using the Local Ensemble Transform Kalman Filter (LETKF)
Surface boundary conditions	
Treatment of sea surface	Climatological sea surface temperature with daily analysis anomaly Climatological sea ice concentration with daily analysis anomaly
Land surface analysis	Snow depth: two-dimensional optimal interpolation scheme using the first guess estimated from the satellite observation and the model forecast Temperature: first guess Soil moisture: simplified extended Kalman filter scheme
Other details	
Land surface and soil	GSM land model based on the Simple Biosphere (SiB) scheme
Radiation	Two-stream with delta-Eddington approximation for short wave (hourly) Two-stream absorption approximation method for long wave (hourly)
Numerical techniques	Spectral (spherical harmonics) in horizontal, finite differences in vertical Two-time-level, semi-Lagrangian, semi-implicit time integration scheme Hydrostatic approximation
Planetary boundary layer	Hybrid scheme combining Mellor and Yamada level-2 turbulence closure with local eddy diffusivity model Similarity theory in bulk formulae for surface layer
Convection	Prognostic Arakawa-Schubert cumulus parameterization
Cloud	PDF-based cloud parameterization
Subgrid orography	Low-level blocked-flow drag, gravity wave drag and turbulent orographic form drag schemes
Non-orographic gravity wave drag	Spectral gravity wave forcing scheme

Table 3.1.2: Specifications of Meso-Scale Model (MSM)

System	
Model (version)	Meso-Scale Model (forecast model: ASUCA)
Date of implementation	1 March 2001 (Last version: 16 June 2022)
Configuration	
Domain	Japan and its surroundings, Lambert projection, 817 × 661 grid points
Horizontal resolution (Grid spacing)	5 km
Vertical resolution (model top)	Stretched hybrid terrain-following coordinate grid with 96 levels (37.5 km)
Forecast length (initial time)	78 hours (00 and 12 UTC) and 39 hours (03, 06, 09, 15, 18 and 21 UTC)
Coupling to ocean / wave / sea ice models	–
Integration time step	100/3 seconds (3-stage Runge-Kutta method)
Initial conditions	
Data assimilation	4D-Var analysis
Surface boundary conditions	
Sea surface temperature	SST analysis (with variations due to atmospheric forcing considered via an ocean mixed layer model) and sea-ice distribution (fixed during integration)
Land surface analysis	Climatological soil moisture values Snow cover analysis for Japan using a land surface model
Lateral boundary conditions	
Model providing lateral boundary conditions	GSM
Lateral boundary condition update frequency	4 times a day, 0 – 45 hours (GSM00 – MSM03/06, GSM06 – MSM09, GSM12 – MSM15/18, GSM18 – MSM21) and 0 – 84 hours (GSM06 – MSM12, GSM18 – MSM00)
Other details	
Soil scheme	Ground temperature prediction using an eight-layer model Soil moisture prediction using force-restore application
Radiation	Two-stream with delta-Eddington approximation for short wave (every 15 minutes) Two-stream absorption approximation for long wave (every 15 minutes)
Large-scale dynamics	Finite volume method with Arakawa-C-type staggered coordinates, horizontally explicit and vertically implicit time integration scheme, combined third- and first-order upwind finite difference schemes in flux form with a limiter proposed by Koren (1993) in advection treatment for monotonicity, time-splitting of vertical advection, fully compressible non-hydrostatic equations
Planetary boundary layer	Mellor-Yamada-Nakanishi-Niino Level-3 scheme Similarity theory in bulk formulae for surface layer
Convection	Kain-Fritsch convection scheme
Cloud/microphysics	Three ice-phase bulk cloud microphysics Consideration of PDF-based cloud distribution in microphysics Time-split treatment for rain and graupel sedimentation Cloud water and cloud cover diagnosis using a partial condensation scheme
Orography	Mean orography smoothed to eliminate shortest-wave components
Horizontal diffusion	–
Gravity wave drag	–

Table 3.1.3: Specifications of Local Forecast Model (LFM)

System	
Model (version)	Local Forecast Model (forecast model: ASUCA)
Date of implementation	30 August 2012 (Last version: 28 March 2023)
Configuration	
Domain	Japan and its surroundings, Lambert projection, 1,581 × 1,301 grid points
Horizontal resolution (Grid spacing)	2 km
Vertical resolution (model top)	Stretched hybrid terrain-following coordinate grid with 96 levels (21.8 km)
Forecast length (initial time)	10 hours (00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 and 23 UTC)
Coupling to ocean / wave / sea ice models	–
Integration time step	12 seconds
Initial conditions	
Data assimilation	Three-hour assimilation window based on hybrid 3D-Var and three one-hour forecasts
Surface boundary conditions	
Sea surface temperature	SST analysis and sea-ice distribution (both fixed during time integration)
Land surface analysis	Climatological values of evaporability, roughness length and albedo Snow cover analysis from MSM
Lateral boundary conditions	
Model providing lateral boundary conditions	MSM
Lateral boundary condition update frequency	8 times a day, 0 – 13-hour forecasts using the latest MSM information
Other details	
Soil scheme	Ground temperature prediction using an eight-layer model Evaporability prediction initialized using climatological values depending on location and season
Radiation	Two-stream with delta-Eddington approximation for short wave (every 15 minutes) Two-stream absorption approximation for long wave (every 15 minutes)
Large-scale dynamics	Finite volume method with Arakawa-C-type staggered coordinates, horizontally explicit and vertically implicit time integration scheme, combined third- and first-order upwind finite difference schemes in flux form with a limiter proposed by Koren (1993) in advection treatment for monotonicity, time-splitting of vertical advection, fully compressible non-hydrostatic equations
Planetary boundary layer	Mellor-Yamada-Nakanishi-Niino Level-3 scheme Similarity theory in bulk formulae for surface layer
Convection	Convective initiation
Cloud/microphysics	Three ice-phase bulk cloud microphysics Time-split treatment for rain and graupel sedimentation Cloud water and cloud cover diagnosis using a partial condensation scheme
Orography	Mean orography smoothed to eliminate shortest-wave components
Horizontal diffusion	–
Subgrid orography	Turbulent orographic form drag scheme

3.2 Global Spectral Model (JMA-GSM2303)

3.2.1 Introduction

The Global Spectral Model (GSM) employs primitive equations to express resolvable motions and states of the atmosphere. It also incorporates sophisticated parameterization schemes for physical processes. In the horizontal direction, prognostic variables are spectrally discretized using triangular truncation at wave number 959 with a quadratic reduced Gaussian grid system (TQ959[1]), and the corresponding transform grids cover around 0.125° in both longitude and latitude. In the vertical direction, the model has 128 layers up to 0.01 hPa.

JMA has operated GSM since March 1988. The model originally had a horizontal resolution of T63 and 16 vertical layers up to 10 hPa with a sigma coordinate system.

In a model upgrade implemented in November 1989, the truncation wave number and the number of vertical layers were increased to T106 and 21, respectively, and a hybrid η vertical coordinate system was adopted.

In March 1996, the horizontal resolution was doubled to T213 and the number of vertical layers was increased to 30. The cumulus parameterization was changed from a Kuo scheme to a prognostic Arakawa-Schubert scheme.

In December 1999, the parameterization schemes underwent extensive refinement. Treatment of cloud water content as a prognostic variable was introduced, and the moist convection process was improved.

In March 2001, the number of vertical layers was increased to 40 and the vertical domain was extended up to 0.4 hPa. The model was highly parallelized to suit massively distributed-memory parallel computer operation.

In February 2005, the Eulerian advection scheme was replaced with a semi-Lagrangian one, and the spectral resolution was increased from T213 (quadratic grid) to TL319 (linear grid). Incremental non-linear normal mode initialization and vertical mode initialization were also introduced.

In March 2006, operations at 06 and 18 UTC were begun with a forecast range of 36 hours in addition to those conducted at 00 UTC with a forecast range of 90 hours and 12 UTC with a forecast range of 216 hours.

In November 2007, the horizontal resolution of GSM was enhanced to TL959, while the number of vertical layers was increased to 60 and the vertical domain was extended up to 0.1 hPa (Iwamura and Kitagawa 2008; Nakagawa 2009). The numerical integration scheme was upgraded from the three-time-level leap-frog scheme to a two-time-level scheme. The forecasts run at 00, 06 and 18 UTC were altered to each cover a uniform period of 84 hours. At the same time, the 20-km-resolution Regional Spectral Model (RSM) and the 24-km-resolution Typhoon Model (TYM) were retired from operational use.

In August 2008, a reduced Gaussian grid was incorporated into GSM as a new dynamical core. This removed redundant grid points at higher latitudes, thereby saving on computational resources (Miyamoto 2006). Incremental non-linear normal mode initialization and vertical mode initialization were eliminated.

In December 2012, a relative humidity threshold was introduced to the diagnostic stratocumulus scheme (Shimokobe 2012).

In March 2013, the coverage period of the forecast run at 12 UTC was extended from 216 to 264 hours.

In April 2013, the radiation scheme was improved by updating the coefficients used for the short-wave parameterization of water vapor.

In March 2014, the number of vertical layers was increased to 100 and the vertical domain was extended up to 0.01 hPa. The parameterization schemes for variables such as the boundary layer, radiation, non-orographic gravity waves and deep convection were also revised (Yonehara *et al.* 2014).

In March 2016, various parameterization schemes such as deep convection, cloud, radiation, land model, and sea surface were substantially revised (Yonehara *et al.* 2017).

In May 2017, the parameterization schemes underwent extensive refinement (Yonehara *et al.* 2018).

In June 2018, the coverage period of forecasts run at 00, 06 and 18 UTC was extended from 84 to 132 hours.

In March 2020, various parameterization schemes such as surface drag processes, land surface processes, and surface albedo and stratocumulus on sea ice were refined (Yonehara *et al.* 2020).

[1] In older GSM versions with Eulerian advection, quadratic grid denotations were indicated by 'T'. The latest GSM with semi-Lagrangian advection re-employed quadratic grid to increase effective resolution (the smallest spatial scale not contaminated by numerical errors or noises). In this section, 'TQ' denotes quadratic grids employed for effective resolution enhancement.

In February 2021, the coverage period of forecasts run at 00 UTC was extended from 132 to 264 hours.

In March 2021, the number of vertical layers was increased to 128 (Ujiie *et al.* 2021).

In March 2023, the horizontal resolution of GSM was enhanced to TQ959 with refinements to parametrized surface drag, non-orographic sub-grid gravity waves, and radiation (Yonehara *et al.* 2023). The source datasets used were updated for application of orographic ancillary files (Kanehama *et al.* 2023a).

3.2.2 Dynamics

The GSM is based on the framework of a semi-implicit semi-Lagrangian global model. In order to reduce the general shortcomings of semi-Lagrangian models (such as the lack of conservation properties and the high computational cost of three-dimensional interpolations), a vertically conservative semi-Lagrangian scheme (Yukimoto *et al.* 2011) is adopted for the GSM.

3.2.2.1 Governing Equations

The GSM is run on an η vertical coordinate system, which is a hybrid between pressure p and σ ($\sigma = p/p_S$, where p_S is surface pressure), implicitly defined as $p = A(\eta) + B(\eta)p_S$. η ranges from 0 to 1; $\eta = 1$ corresponds to the lower boundary (ground surface) and $\eta = 0$ corresponds to the upper boundary. The prognostic variables (wind vector $\boldsymbol{u} = (u, v)$, temperature T, pressure p, specific humidity q and cloud water content q_c) follow the system of primitive equations in the η-coordinate system as follows:

$$\frac{d\mathbf{u}}{dt} = -f\mathbf{z} \times \mathbf{u} - (\nabla\Phi + R_d T_V \nabla \ln p) + \mathbf{F}_{\boldsymbol{u}} \tag{3.2.1}$$

$$\frac{dT}{dt} = \frac{\kappa T_V \omega}{\left[1 + \left(C_{pv}/C_{pd} - 1\right) q\right] p} + F_T \tag{3.2.2}$$

$$\frac{dq}{dt} = F_q \tag{3.2.3}$$

$$\frac{dq_c}{dt} = F_c \tag{3.2.4}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial \eta}\right) + \nabla \cdot \left(\mathbf{u}\frac{\partial p}{\partial \eta}\right) + \frac{\partial}{\partial \eta}\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right) = 0 \tag{3.2.5}$$

Here, d/dt is a total derivative defined as $d/dt = \partial/\partial t + \mathbf{u} \cdot \nabla + \dot{\eta}\partial/\partial\eta$, and ∇ is a horizontal gradient operator. The other notations used above are conventional: $\mathbf{z}$ is the unit vertical vector, T_V is the virtual temperature, f is the Coriolis parameter, R_d is the gas constant for dry air, and $\kappa = R_d/C_{pd}$. C_{pd} is the specific heat capacity at the constant pressure of dry air and C_{pv} is the specific heat capacity at the constant pressure of water vapor. $\mathbf{F}_{\boldsymbol{u}}$, F_T, F_q and F_c are tendencies relating to parameterized processes. In addition, $\mathbf{F}_{\boldsymbol{u}}$ and F_T include the effects of horizontal diffusion (to be described later). Integrating Eq. (3.2.5) with respect to η using the boundary conditions of $\dot{\eta} = 0$ at $\eta = 0$ and $\eta = 1$, η-velocity and ω are found:

$$\dot{\eta}\frac{\partial p}{\partial \eta} = -\frac{\partial p}{\partial t} - \int_0^\eta \nabla \cdot \left(\mathbf{u}\frac{\partial p}{\partial \eta'}\right) d\eta' \tag{3.2.6}$$

$$\omega \equiv \frac{dp}{dt} = -\int_0^\eta \nabla \cdot \left(\mathbf{u}\frac{\partial p}{\partial \eta'}\right) d\eta' + \mathbf{u} \cdot \nabla p \tag{3.2.7}$$

The geopotential Φ is given by the following hydrostatic relation:

$$\frac{\partial \Phi}{\partial \eta} = -R_d T_V \frac{\partial \ln p}{\partial \eta} \tag{3.2.8}$$

3.2.2.2 Vertical Finite Difference Scheme

The vertical finite difference scheme is coded by following Simmons and Burridge (1981). The prognostic variables $\mathbf{u}$, T, q and q_c are defined on the full levels, while η (including vertical fluxes) is defined on half-integer levels. Pressure on half-integer levels are expressed as

$$p_{k-1/2} = A_{k-1/2} + B_{k-1/2} p_S \qquad (k = 1, 2, \ldots, \mathrm{kmax}) \tag{3.2.9}$$

Here, the level index k increases with height, *kmax* is the index of the highest model level, $A_{k-1/2} = A(\eta_{k-1/2})$ and $B_{k-1/2} = B(\eta_{k-1/2})$. The profiles of $A_{k-1/2}$ and $B_{k-1/2}$ are determined by following Kawai *et al.* (2013). $A_{1/2}$ is set to zero so that the lowest level coincides with the ground surface, and values of $B_{k-1/2}$ above 60hPa are set to zero so that these levels coincide with constant pressure surfaces. For intermediate levels, $A_{k-1/2}$ and $B_{k-1/2}$ vary smoothly with k.

From the hydrostatic relation given by Eq. (3.2.8) the finite difference form of geopotential on the full level is chosen as

$$\Phi_k = \Phi_S + \sum_{k'=1}^{k-1} R_d T_{Vk'} \ln\left(\frac{p_{k'-1/2}}{p_{k'+1/2}}\right) + \alpha_k R_d T_{Vk} \tag{3.2.10}$$

$$\alpha_k = \begin{cases} 1 - \dfrac{p_{k+1/2}}{\delta p_k} \ln\left(\dfrac{p_{k-1/2}}{p_{k+1/2}}\right) & (1 \le k < \mathrm{kmax}) \\ \ln 2 & (k = \mathrm{kmax}) \end{cases} \tag{3.2.11}$$

Here, Φ_S is the geopotential at the surface, and $\delta p_k = p_{k-1/2} - p_{k+1/2}$. The pressure gradient force term in Eq. (3.2.1) and the adiabatic heating rate term in Eq. (3.2.2) can then be written in discretized form as

$$(\nabla\Phi + R_d T_V \nabla \ln p)_k = \nabla\Phi_k + \frac{R_d T_{Vk}}{\delta p_k}\left[\ln\left(\frac{p_{k-1/2}}{p_{k+1/2}}\right)\nabla p_{k+1/2} + \alpha_k \nabla(\delta p_k)\right] \tag{3.2.12}$$

and

$$\left[\frac{\kappa T_V}{C_p/C_{pd}}\frac{\omega}{p}\right]_k = \frac{\kappa T_{Vk}}{C_{pk}/C_{pd}}\frac{1}{\delta p_k}\left[\left(\ln\frac{p_{k^-1/2}}{p_{k+1/2}}\right)\left(B_{k+1/2}\mathbf{u}_k \cdot \nabla p_S - \sum_{l=k+1}^{\mathrm{kmax}} \nabla\cdot(\mathbf{u}_l \delta p_l)\right) - \alpha_k (\nabla\cdot\mathbf{u}_k)\delta p_k\right] \tag{3.2.13}$$

respectively, where C_p is the specific heat capacity at the constant pressure of moist air, that defined as $C_p = [1 + (C_{pv}/C_{pd} - 1)q]C_{pd}$. The vertical mass flux in Eq. (3.2.6) is discretized as

$$\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2} = -B_{k-1/2}\frac{\partial p_S}{\partial t} - \sum_{l=k}^{\mathrm{kmax}} \nabla\cdot(\mathbf{u}_l \delta p_l) = B_{k-1/2}\sum_{l=1}^{\mathrm{kmax}} \nabla\cdot(\mathbf{u}_l \delta p_l) - \sum_{l=k}^{\mathrm{kmax}} \nabla\cdot(\mathbf{u}_l \delta p_l) \tag{3.2.14}$$

3.2.2.3 Horizontal Grid

To mitigate the overconcentration of grid points at high latitudes and lower the computational cost, a reduced Gaussian grid is adopted for the GSM. The number of east-west grid points at each latitude is determined based on the magnitude of associated Legendre functions, which is negligibly small at high latitudes and in high orders. With this method, the computational cost of Legendre transformation can also be reduced (Juang 2004). The number of east-west grid points is in fact restricted by FFT package specifications, the number of east-west decompositions in parallelization (as described in 3.2.10) and the interval of coarser radiation grids (as shown in 3.2.3).

3.2.2.4 Semi-implicit Semi-Lagrangian Formulation

Prior to the time integration, the forecast equations (Eq. (3.2.1) - Eq. (3.2.5)) are rewritten in the form of $d_H X/dt = \partial X/\partial t + \mathbf{u} \cdot \nabla X = R$ with vertical advection terms incorporated into R on the right-hand side. These equations are integrated with respect to time along the trajectory of the parcel from the departure point D at time t to the arrival point A at time $t + \delta t$. The linear term L separated from the forcing term R is treated semi-implicitly (i.e. using a trapezoidal rule), and the remaining R, including vertical advection terms, are treated with spatial averaging (Tanguay *et al.* 1992).

The resulting linear terms are slightly amplified by the factor $\beta = 1.2$ for computational stability, and the following is obtained:

$$X^{A+} - X^{D0} = \delta t \frac{R^{A0} + R^{D(+)}}{2} + \delta t \beta \left[\frac{L^{A+} + L^{D-}}{2} - \frac{L^{A0} + L^{D0}}{2} \right] \tag{3.2.15}$$

Superscript A represents the arrival point $\mathbf{x}_{ij}$ assumed to be on the Gaussian grid, and D is the departure point $\mathbf{x}_{ij} - \boldsymbol{\alpha}$ (the displacement vector $\boldsymbol{\alpha}$, whose calculation will be described later). The abbreviations used above are the same as those for $X^{A+} = X(\mathbf{x}, t + \delta t)$, $X^{D0} = X(\mathbf{x} - \boldsymbol{\alpha}, t)$, $R^{A0} = R(\mathbf{x}, t)$, $R^{D(+)} = R(\mathbf{x} - \boldsymbol{\alpha}, t + \delta t)$ and others. $R^{D(+)}$ is calculated based on extrapolation with respect to time. Rearranging the terms of the above equations gives a system of linear equations for the unknown values X^{A+}:

$$X^{A+} - \frac{\beta \delta t}{2} L^{A+} = \left[X^0 + \frac{\delta t}{2} \left\{ R^{(+)} - \beta \left(L^0 - L^- \right) \right\} \right]^D + \frac{\delta t}{2} \left[R^0 - \beta L^0 \right]^A \tag{3.2.16}$$

3.2.2.5 Vertically Conservative Semi-Lagrangian Scheme

Yoshimura and Matsumura (2003, 2004) developed a vertically conservative semi-Lagrangian scheme in which vertical advection is treated separately from horizontal advection so that conserved quantities such as water vapor under non-dissipative conditions are preserved in the vertical direction. Processing advection separately in the horizontal and vertical directions also reduces the model's cost of interpolation.

Eq. (3.2.16) can be reformulated with flux forms appropriate for a scheme in which vertical advection can retain conservative properties. Beginning with Eq. (3.2.5) and Eq. (3.2.1) - Eq. (3.2.4), rewriting is performed as follows:

$$\frac{d_H}{dt} \frac{\partial p}{\partial \eta} = -D \frac{\partial p}{\partial \eta} - \frac{\partial}{\partial \eta} \left(\dot{\eta} \frac{\partial p}{\partial \eta} \right) \tag{3.2.17}$$

$$\frac{d_H}{dt} \left(X \frac{\partial p}{\partial \eta} \right) = -DX \frac{\partial p}{\partial \eta} - \frac{\partial}{\partial \eta} \left(\dot{\eta} X \frac{\partial p}{\partial \eta} \right) + R_X \frac{\partial p}{\partial \eta} \tag{3.2.18}$$

Here, X represents $\mathbf{u}$, T_V, q and q_c, and $R_X = dX/dt$. The parallel nature of these equations is easily recognizable. The first term on the right hand side of these equations represents the increase caused by horizontal convergence, and the second term is the increase caused by vertical flux convergence. With respect to the latter, where q and q_c being conservative when $R_X = 0$, devising a vertically integrated quantity that remains unchanged in vertical advection appears to be a promising approach. A simple outline of the procedure is given here for specific humidity q without R_q.

Vertical discretization and time integration during the period δt described earlier give the following equations with the omission of terms related to the semi-implicit method for reasons of simplicity:

$$\begin{aligned} (\delta p_k)^{A+} = & \left[(\delta p_k)^0 - \frac{1}{2} (D_k \delta p_k)^{(+)} \delta t + \frac{1}{2} \left\{ \left(\dot{\eta} \frac{\partial p}{\partial \eta} \right)_{k+1/2} - \left(\dot{\eta} \frac{\partial p}{\partial \eta} \right)_{k-1/2} \right\}^{(+)} \delta t \right]^D \\ & + \left[-\frac{1}{2} (D_k \delta p_k)^0 \delta t + \frac{1}{2} \left\{ \left(\dot{\eta} \frac{\partial p}{\partial \eta} \right)_{k+1/2} - \left(\dot{\eta} \frac{\partial p}{\partial \eta} \right)_{k-1/2} \right\}^0 \delta t \right]^A \end{aligned} \tag{3.2.19}$$

$$(q_k\delta p_k)^{A+} = \left[(q_k\delta p_k)^0 - \frac{1}{2}q_k^0\,(D_k\delta p_k)^{(+)}\,\delta t + \frac{1}{2}\left\{\left(q\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+1/2} - \left(q\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2}\right\}^{(+)}\delta t\right]^D$$
$$+\left[-\frac{1}{2}q_k^+\,(D_k\delta p_k)^0\,\delta t + \frac{1}{2}\left\{\left(q\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+1/2} - \left(q\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2}\right\}^{0}\delta t\right]^A \tag{3.2.20}$$

$$p_{k-1/2} = \sum_{k'=k}^{\mathrm{kmax}} \delta p_{k'}, \quad (k = 1, 2, \ldots, \mathrm{kmax}) \tag{3.2.21}$$

Here, the vertically cumulative quantity Q is defined as follows:

$$Q_{k-1/2} = \sum_{k'=k}^{\mathrm{kmax}} \delta Q_{k'}, \quad \delta Q_k = q_k\delta p_k, \quad (k = 1, 2, \ldots, \mathrm{kmax}+1) \tag{3.2.22}$$

Eq. (3.2.20) rewritten for δQ_k is found to be similar to Eq. (3.2.19) for δp_k, and there is a clear correspondence between Q and p. Computation of Q can therefore be carried out in the five steps outlined below in a fashion parallel to that of p. The first two steps concern the operations inside the square brackets $[\ldots]^D$ in the above equations. The third step involves the calculation of variables at departure points based on interpolation. The fourth and the fifth steps are similar to the first two, but for the operations in the square brackets $[\ldots]^A$.

1. First step: Horizontal divergence is calculated. As the mass of each layer δp_k varies to $\delta p'_k$, the half-level pressure values $p_{k-1/2}$ by which layers are bound also shift to $p'_{k-1/2}$, which can be computed using Eq. (3.2.21). The values of q_k remain constant under the horizontal convergence $q'_k = q_k$.

2. Second step: Vertical flux convergence is calculated using Eq. (3.2.14) as in the Eulerian scheme. In the same way as in the first step, $\delta p'_k$ varies to $\delta p''_k$, and the values of $p'_{k-1/2}$ shift to $p''_{k-1/2}$ except $k = 1$ ($p'_{1/2} = p''_{1/2}$). In this step, the shift in $Q'_{k-1/2}$ caused by the vertical flux convergence is computed based on interpolation from $Q'_{k-1/2}(p'_{k-1/2})$ using $Q''_{k-1/2} = Q'_{k-1/2}(p''_{k-1/2})$. This procedure ensures the conservation of the total mass-weighted integral $Q'_{1/2} = Q''_{1/2}$, because $p'_{1/2} = p''_{1/2}$ holds and the other values of $p''_{k-1/2}$ ($k = 2, 3, \ldots, kmax$) merely have their intervals changed in the vertical column. New values of q''_k are computed using $\delta Q''_k$ and $\delta p''_k$ with Eq. (3.2.22).

3. Third step: Horizontal advection is incorporated by computing $(\delta p_k)^D$ and q_k^D via quasi-cubic interpolation.

4. Forth step: Vertical flux convergence is calculated at the arrival point via the second step.

5. Fifth step: Horizontal divergence is calculated at the arrival point via the first step.

The time-integration of q and q_c is completed based on these five steps, and that of $\mathbf{u}$, T_V and p_S is followed by the semi-implicit calculation shown in Eq. (3.2.16).

3.2.2.6 Departure Point Determination

The displacement vector $\boldsymbol{\alpha}$ (as yet undetermined) obeys the implicit equation

$$\boldsymbol{\alpha} = \delta t\left\{\frac{\mathbf{u}_k(\mathbf{x}_{ij} - \boldsymbol{\alpha},\, t+\delta t) + \mathbf{u}_k(\mathbf{x}_{ij},\, t)}{2}\right\} \tag{3.2.23}$$

which expresses that the horizontal advection during the time interval δt is related to the average of future time-step wind value at the departure point and current time-step wind value at the arrival point (SETTLS; Hortal 2002). To improve stability, a method based on wind integrated in a semi-Lagrangian scheme rather than the time extrapolated wind is adopted (Yoshimura 2002). This implicit equation is solved by successive insertions of $\boldsymbol{\alpha}$. For the computation of these vector components, it is considered that the axes of the local coordinates (λ, φ) rotate due to the spherical metric as a parcel advances along a trajectory, as is the case whenever horizontal vector components are interpolated on a sphere. The wind at the departure point is computed from linear interpolation except for the last third of the iteration, for which a quasi-cubic approach is used.

3.2.2.7 Spectral Method and Horizontal Diffusion

Spectral variables (i.e. vorticity $\zeta(= \mathbf{z} \cdot \nabla \times \mathbf{u})$, divergence $D(= \nabla \cdot \mathbf{u})$, T_V and $\ln(p_S)$) are expanded in terms of spherical harmonics with triangular truncation. In accordance with the framework of the semi-Lagrangian scheme, a linear Gaussian transformation grid is used. Solutions of horizontal Helmholtz equations (derived when Eq. (3.2.16) is solved for D), horizontal diffusion and variables such as the differentials on the sphere are calculated using the spectral method (Bourke 1974; Hoskins and Simmons 1975). The remaining variables q and q_c are defined only on grid points.

To prevent the accumulation of small scale noise (spectral blocking), fourth-order linear horizontal diffusion is applied to ζ, D and T_V:

$$\left(\frac{\partial \zeta}{\partial t}\right)_{\text{hdiff,4th}} = -K_{4\text{th}}\left(\nabla^4 - \frac{4}{a^4}\right)\zeta \tag{3.2.24a}$$

$$\left(\frac{\partial D}{\partial t}\right)_{\text{hdiff,4th}} = -K_{4\text{th}}\nabla^4 D \tag{3.2.24b}$$

$$\left(\frac{\partial T_V}{\partial t}\right)_{\text{hdiff,4th}} = -K_{4\text{th}}\nabla^4\left[T_V - \frac{\partial \bar{T}_V}{\partial \bar{p}}p\right] = -K_{4\text{th}}\nabla^4\left[T_V - \frac{\partial \bar{T}_V}{\partial \bar{p}}B(\eta)p_S\right] \tag{3.2.24c}$$

Here, $K_{4\text{th}}$ is the diffusion coefficient for the fourth-order horizontal diffusion and a is the radius of the earth. Bars over variables indicate the global average on the η-surface. Angular momentum conservation does not allow the horizontal diffusion process to work on vorticity with total wave number 1 as shown by Eq. (3.2.24a). Diffusion for virtual temperature is modified to work on the constant pressure surface; otherwise, diffusion on a declining η-surface may produce spurious mixing along steep mountain slopes. $K_{4\text{th}}$ is chosen so that the power spectrum of enstrophy coincides with that expected based on the two-dimensional turbulence theory.

To provide a sponge layer that absorbs waves incident on the upper boundary, second-order linear horizontal diffusion is applied to the divergence term D in layers above 30hPa:

$$\left(\frac{\partial D}{\partial t}\right)_{\text{hdiff,2nd}} = -K_{2\text{nd}}\nabla^2 D \tag{3.2.25a}$$

$$K_{2\text{nd}} = K_0 \sin^2\left(\frac{\pi}{2}\frac{\ln p - \ln p_{\text{btm}}}{\ln p_{\text{top}} - \ln p_{\text{btm}}}\right) \tag{3.2.25b}$$

Here, K_0 is the base diffusion coefficient for second-order horizontal diffusion, p_{top} is the pressure at the highest model level (0.01hPa), and p_{btm} is the pressure at the altitude where the sponge layer begins (30hPa). To suppress wave reflection at the upper boundary, $K_{2\text{nd}}$ is gradually enhanced with height as shown by Eq. (3.2.25b). K_0 is determined experimentally so that spurious wave reflections at the upper boundary can be appropriately removed.

These fourth and second-order horizontal diffusion terms are calculated backward and implicitly in spectral forms as an independent step after semi-implicit time integration.

3.2.3 Radiation

The radiative heating rate is computed as the divergence of net radiation fluxes:

$$\left(\frac{\partial T}{\partial t}\right)_{\text{rad}} = \frac{g}{C_p}\frac{\partial F}{\partial p} \tag{3.2.26}$$

where F $(= F^{\uparrow} - F^{\downarrow})$ is the net radiation flux, $F^{\uparrow}$ $(F^{\downarrow})$ is the upwelling (downwelling) radiation flux, g is the acceleration of gravity and C_p is the specific heat at the constant pressure of moist air.

Solving the radiative transfer equation is computationally very expensive. To reduce this burden, full radiation computation is performed only once an hour for longwave and shortwave on a coarser (reduced radiation) grid. The radiative heating rates associated with longwave and shortwave radiation are corrected for every time step using the surface temperature and the solar zenith angle, respectively.

3.2.3.1 Longwave Radiation

The two-stream radiation transfer method involving the absorption approximation approach (Yabu 2013) is adopted for longwave flux and cooling rate computation. The spectrum in the longwave region is divided into 11 bands as shown in Table 3.2.1.

Table 3.2.1: Band configuration for the longwave radiation scheme. The calculation approaches for absorption associated with atmospheric molecules are C-k (correlated k-distribution method) and S-k (k-distribution method with scaling approximation). Notation in each parenthesis denotes gas overlap assumption (pf: perfectly correlated; no: perfectly uncorrelated; pt: partly correlated; cg: combined gas). The number of sub-bands is also shown in each parenthesis.

Band Number	1	2	3a	3b	3c	4	5	6	7	8	9
Wavenumber (/cm)	25-340	340-540	540-620	620-720	720-800	800-980	980-1100	1100-1215	1215-1380	1380-1900	1900-3000
Major absorption gas											
H2O(Line)	C-k(16)		C-k(16,cg)			S-k(6)	S-k(6,pf)	S-k(16,pf)	S-k(4)	C-k(16)	S-k(6)
CO2											
O3							C-k(16)	C-k(16)			
H2O(Continuum)	S-k(16,pt)					S-k(6,pf)		S-k(16,pf)	S-k(4,pf)	S-k(16,pt)	S-k(6,pf)
Minor absorption gas											
CO2						S-k(6,pf)					
N2O			S-k(6,pf)					S-k(16,pf)	S-k(2,no)		
CH4									S-k(2,no)		
CFC-11,CFC-12,HCFC-22						S-k(6,pf)					
Number of sub-bands	16	16	16	16	16	6	16	16	16(=4x2x2)	16	6

In the two-stream method, downwelling (upwelling) radiation fluxes are calculated sequentially from the model top (surface) as follows:

$$F^{\downarrow}_{k-1/2} = 0 \qquad (k = k_{max} + 1) \tag{3.2.27a}$$

$$F^{\downarrow}_{k-1/2} = F^{\downarrow}_{k+1/2} e^{-\tau_k/\mu} + B^{\downarrow}_k (1 - e^{-\tau_k/\mu}) \qquad (k = k_{max}, \cdots, 1) \tag{3.2.27b}$$

$$F^{\uparrow}_{k-1/2} = B_s \qquad (k = 1) \tag{3.2.27c}$$

$$F^{\uparrow}_{k-1/2} = F^{\uparrow}_{k-3/2} e^{-\tau_{k-1}/\mu} + B^{\uparrow}_{k-1} (1 - e^{-\tau_{k-1}/\mu}) \qquad (k = 2, \cdots, k_{max} + 1) \tag{3.2.27d}$$

where $F^{\downarrow}_{k-1/2}$ ($F^{\uparrow}_{k-1/2}$) is the downwelling (upwelling) radiation flux at the half-integer level $k - 1/2$, τ_k is the optical thickness of the model layer k, $B^{\downarrow}_k$ ($B^{\uparrow}_k$) is the downward (upward) effective Planck flux (Chou *et al.* 2001) at the model layer k, B_s is the Planck flux at the surface and $1/\mu$ (= 1.66) is the diffusivity factor.

Depending on the absorber and the spectral band, absorption associated with atmospheric molecules is evaluated using one of two k-distribution methods (see Table 3.2.1). The correlated k-distribution method (Fu and Liou 1992) is applied to absorption dominant in the middle atmosphere. Absorption coefficients at 51 pressure levels between 1000 and 0.01 hPa are tabulated in advance based on the Line-By-Line calculation, and gas absorption data are derived from HITRAN2000 (Rothman *et al.* 2003). The k-distribution method with scaling approximation (Chou *et al.* 2001) is applied to absorption with a Lorentzian line shape assumed, and only one absorption coefficient at a specified pressure level (500 hPa) is prepared. This method is also applied to H_2O continuum absorption based on the MT-CKD model with scaling parameters from Zhong and Haigh (1995).

To handle the overlapping of gas absorption within each band, one of three assumptions (perfectly correlated, perfectly uncorrelated and partly correlated) (Zhang *et al.* 2003) is made other than that for CO_2 and H_2O in the 540–800 cm^{-1} region, which is handled using the direct combined gas mapping approach of Li and Barker (2005) (see Table 3.2.1).

Maximum-random cloud overlapping (Geleyn and Hollingsworth 1979) is assumed in the longwave radiation scheme, and is implemented to the two-stream method as per Li (2002). Cloud optical thickness is parameterized as per Lindner and Li (2000) for liquid droplets and as per Ebert and Curry (1992) for ice particles.

3.2.3.2 Shortwave Radiation

Shortwave scattering and absorption are modeled in two-stream formulation using the delta-Eddington approximation (Joseph *et al.* 1976; Coakley *et al.* 1983). The spectrum in the shortwave region is divided into 16 bands (10 in ultraviolet, 5 in visible and 1 in near-infrared) based on Freidenreich and Ramaswamy (1999), while absorption by water vapor in the near-infrared region is calculated via exponential-sum fitting of transmissions method with seven sub-bands based on Collins *et al.* (2006).

Assuming a plane-parallel atmosphere, diffuse radiation fluxes are derived from the following simultaneous equations (Meador and Weaver 1980):

$$\frac{dF^{\uparrow}}{d\tau} = \gamma_1 F^{\uparrow} - \gamma_2 F^{\downarrow} - \gamma_3 \omega_0 S_0 e^{-\tau/\mu_0} \tag{3.2.28a}$$

$$\frac{dF^{\downarrow}}{d\tau} = \gamma_2 F^{\uparrow} - \gamma_1 F^{\downarrow} + (1-\gamma_3)\omega_0 S_0 e^{-\tau/\mu_0} \tag{3.2.28b}$$

where τ is the optical thickness, ω_0 is the single scattering albedo and S_0 is the incident solar irradiance in the direction μ_0 (the cosine of the solar zenith angle). The coefficients γ_i ($i = 1, \cdots, 3$) are given by

$$\gamma_1 = \frac{1}{4}[7 - \omega_0(4+3g)], \quad \gamma_2 = -\frac{1}{4}[1 - \omega_0(4-3g)], \quad \gamma_3 = \frac{1}{4}(2 - 3g\mu_0) \tag{3.2.29}$$

where g is the asymmetry factor. In the delta-Eddington method, solar optical properties such as τ, ω_0 and g are adjusted using the fraction of forward-scattering peak f:

$$\tau' = (1 - \omega_0 f)\tau, \quad \omega_0' = \frac{(1-f)\omega_0}{1-\omega_0 f}, \quad g' = \frac{g-f}{1-f} \tag{3.2.30}$$

The shortwave radiation flux in each column is calculated using the Practical Independent Column Approximation (PICA; Nagasawa 2012) method, which is a simplified ICA approach based on Collins (2001). The total shortwave radiation flux F is given as a weighted average of the fluxes in each subcolumn as follows:

$$F = \frac{1}{A_{tot}} \sum_{i=1}^{N_{max}} A_i F_i \tag{3.2.31}$$

where A_{tot} is the total area of the relevant subcolumns, A_i is the fractional area of each subcolumn, F_i is the flux in each subcolumn and N_{max} is the maximum number of subcolumns. The binary cloud configuration in the column is given by cloud cover assuming maximum-random cloud overlapping.

The reflectance and transmittance of the layer are calculated as functions of the total optical thickness τ_{total}, the total single scattering albedo $\omega_{0\,\text{total}}$ and the total asymmetry factor g_{total} of the layer:

$$\tau_{\text{total}} = \tau_R + \tau_g + \tau_a + \tau_c \tag{3.2.32a}$$

$$\omega_{0\,\text{total}} = \frac{\tau_R + \omega_{0a}\tau_a + \omega_{0c}\tau_c}{\tau_R + \tau_g + \tau_a + \tau_c} \tag{3.2.32b}$$

$$g_{\text{total}} = \frac{g_a\omega_{0a}\tau_a + g_c\omega_{0c}\tau_c}{\tau_R + \omega_{0a}\tau_a + \omega_{0c}\tau_c} \tag{3.2.32c}$$

where the subscripts R, g, a and c denote molecular Rayleigh scattering, gaseous absorption, and Mie scattering/absorption caused by aerosols and clouds, respectively.

The cloud optical properties are parameterized following Dobbie *et al.* (1999) for liquid droplets and Ebert and Curry (1992) for ice particles.

3.2.3.3 Gas Concentrations and Aerosol Climatology

The radiatively active gases considered in the radiation scheme are water vapor, ozone, carbon dioxide, oxygen, methane, nitrous oxide, CFC-11, CFC-12 and HCFC-22. Prognostic water vapor concentrations are used for the troposphere, while climatological distribution based on Randel *et al.* (1998) is used for areas above it. The

Table 3.2.2: Values of the globally uniform gas concentration (unit is ppmv).

CO_2	O_2	CH_4	N_2O	CFC-11	CFC-12	HCFC-22
396	209490	1.824	0.3259	0.0003	0.0005	0.0002

three-dimensional monthly mean climatology of ozone concentration is derived from the latest MRI CCM2 reanalysis (Deushi and Shibata 2011). Other radiatively active gases have globally uniform concentrations as shown in Table 3.2.2.

In the radiation scheme, five aerosol types (sulfate, black carbon, organic carbon, sea salt and mineral dust) are considered (Yabu *et al.* 2017). The three-dimensional monthly mean climatology of aerosol mass concentration is derived from aerosol CTM calculation (see Section 3.9), and the optical properties of these aerosols are pre-computed via Mie scattering calculation. The climatological distribution of total-column values for aerosol optical depth (ATOD) is used in combination. Monthly mean ATOD climatology is derived from Moderate Resolution Imaging Spectroradiometer (MODIS), Multi-angle Imaging Spectro-Radiometer (MISR) and Ozone Monitoring Instrument (OMI) observations.

3.2.3.4 Cloud Properties

Two types of cloud are considered in the radiation scheme. One is stratiform cloud, whose properties (such as cloud cover, cloud water content and cloud ice content) are provided by the cloud scheme. The other is convective cloud, whose properties are diagnosed using the upward convective mass flux calculated in the convection scheme. The effective radius of cloud liquid droplets $r_{e,liq}$ [μm] is parameterized based on Martin *et al.* (1994) as follows:

$$r_{e,liq} = 10^4 \left(\frac{3CWC}{4\pi\rho_w k N_{TOT}} \right)^{1/3} \tag{3.2.33}$$

where CWC is the cloud water content [gm^{-3}], N_{TOT} is the number concentration of water cloud droplets [cm^{-3}], ρ_w is the density of water [gm^{-3}] and k is a constant. N_{TOT} is set as 100 and 300 cm^{-3} over ocean and land areas, respectively. The effective radius of cloud ice particles $r_{e,ice}$ [μm] depends on temperature T [K] and cloud ice content IWC [gm^{-3}] as follows (Sun 2001):

$$\begin{aligned} f &= 1.2351 + 0.0105(T - 273.15) \\ a &= 45.8966\,\mathrm{IWC}^{0.2214} \\ b &= 0.7957\,\mathrm{IWC}^{0.2535} \\ r &= 0.838167 f\,[a + b(T - 83.15)] \\ r_{e,ice} &= \min(\max(r, 13), 130). \end{aligned} \tag{3.2.34}$$

3.2.4 Cumulus Convection

The GSM employs a spectral mass-flux convective parametrization scheme based on Arakawa and Schubert (1974) and Moorthi and Suarez (1992). Prognostic closure based on Randall and Pan (1993) is used, although many modifications are made to the original. In addition, a triggering mechanism based on the DCAPE concept (Xie and Zhang 2000) is adopted to suppress excessive convective activity. Convective downdraft, convective momentum transport and mid-level convection are also included in the scheme.

3.2.4.1 Convective Effect on Large-scale Variables

The convective effect on large-scale variables is estimated using the following equations:

$$\left(\rho \frac{\partial \overline{s}}{\partial t} \right)_{\mathrm{conv}} = \sum_n D_n^u (s_n^u - \overline{s}) + D^d (s^d - \overline{s}) + \left(\sum_n M_n^u - M^d \right) \frac{\partial \overline{s}}{\partial z} - L_i c - L_v e - \delta E_s \tag{3.2.35}$$

$$\left(\rho\frac{\partial\bar{h}}{\partial t}\right)_{\text{conv}} = \sum_n D_n^u(h_n^u - \bar{h}) + D^d(h^d - \bar{h}) + \left(\sum_n M_n^u - M^d\right)\frac{\partial\bar{h}}{\partial z} - L_i c - \delta E_h \tag{3.2.36}$$

where ρ is the density of air, s and h are dry and moist static energy, D is detrainment to the environment, L_v and L_i are the latent heat of vaporization and melting, e is the evaporation amount under the convective cloud base, and c is the melt amount. Details of δE_s and δE_h are provided in Subsection 3.2.4.2. The over-bar denotes an environmental value, the superscripts u and d indicate updraft and downdraft respectively, and the subscript n indicates each plume of an ensemble of convection. Multiple plumes are considered for updraft, and a single plume is calculated for downdraft.

The first terms on the right of Eqs. (3.2.35) and (3.2.36) represent detrainment from updraft to the environment, the second terms indicate detrainment from downdraft to the environment, the third represent compensating subsidence, and the fourth represent the melting effect below freezing level. The fifth term in Eq. (3.2.35) denotes the evaporation effect under the cloud base.

3.2.4.2 Cloud Model

Based on Arakawa and Schubert (1974), the ensemble effect of multiple cumulus types is considered.

The mass flux of each plume is expressed as

$$M_n = M_{Bn}(t)\eta_n(z) \tag{3.2.37}$$

where M_B is the mass flux at the cloud base and η is the normalized cloud mass flux, which is 1 at the cloud base. The details of M_B calculation are provided in Subsection 3.2.4.3.

The cloud base of each plume is fixed near 900 hPa in the model. Each type of cumulus is defined by the level of the cloud top, where the updraft cloud mass loses buoyancy and detrainment occurs. The vertical profile of the upward mass flux η is assumed to be a linear function of height z, as proposed by Moorthi and Suarez (1992), and can be expressed as

$$\eta_n = 1 + \lambda_n(z - z_b) \tag{3.2.38}$$

where λ is the entrainment rate and z_b is the cloud base height. λ is diagnosed using a condition in which each plume loses buoyancy at its cloud top.

Cloud water content in the updraft is converted to precipitation, and the conversion is formulated as an autoconversion scheme as proposed by Kessler (1969).

The mass flux below the cloud base is calculated based on Jakob and Siebesma (2003):

$$\frac{\partial\eta}{\partial z} = \frac{C}{z}\eta \tag{3.2.39}$$

where C is a constant set to 0.5.

The plume ascends with the entrainment rate obtained from Eq. (3.2.39) below the cloud base. Assuming the occurrence of convection to be associated with positive subgrid scale fluctuations of temperature and moisture, air with higher dry and moist static energy than the grid-mean environment is entrained into the plume. δE_s and δE_h in Eqs. (3.2.35) and (3.2.36) represent the entrainment of excess energy to the updraft below the cloud base.

3.2.4.3 Closure

Closure is based on Randall and Pan (1993), and numerous modifications are made to the original scheme. For deep convection, the following prognostic equation is used to calculate the upward mass flux at the cloud base M_B for each plume (the subscript n is omitted for simplicity):

$$\frac{dM_B}{dt} = \max\left(\frac{A - fA_0}{2\alpha}, 0\right)\min\left(\frac{\lambda}{\lambda_{\min}}, 1\right)\max\left(\lambda_{\max}, 0\right)\left(\frac{\Delta p}{\Delta p_{\text{eff}}}\right) - \frac{M_B}{2\tau_d} \tag{3.2.40}$$

where A denotes the cloud work function, A_0 is the average of observed cloud work functions as given by Lord and Arakawa (1980), Δp is the depth of model cloud top layer, Δp_{eff} is the effective depth of the cloud top, and τ_d is the time constant of cumulus kinetic energy decay. The parameter f is introduced to incorporate the effects of grid-scale vertical wind and convective inhibition. This is given by

$$f = \frac{\omega}{\omega_0} + \frac{A_i}{A_{i0}} + c \tag{3.2.41}$$

where ω denotes the vertical pressure velocity at the lowest level, A_i represents the work involved in lifting the parcel to the level of free convection, and ω_0, A_{i0} and c are empirically determined constants. The constraint $0 \leqq f \leqq 2$ is imposed to ensure realistic tendency of M_B. In order to suppress tall cumuli in dry conditions and incorporate the effects of turbulence in the planetary boundary layer, the parameter $\lambda_{\min}$ is defined as follows:

$$\lambda_{\min} = \max\left(\frac{0.9 - \text{RH}}{0.2},\ 10^{-3}\right)\frac{0.3}{5l_0} \tag{3.2.42}$$

where RH denotes the vertical mean of relative humidity between the cloud base and the cloud top, and l_0 represents the mixing length of the planetary boundary layer. The parameter $\lambda_{\max}$ is introduced to suppress tall cumuli with unnaturally large entrainment rates, and is defined as

$$\lambda_{\max} = \min\left(\frac{\lambda - \lambda_2}{\lambda_1 - \lambda_2},\ 1\right) \tag{3.2.43}$$

where $\lambda_1 = a_1/(z_t - z_b)$, $\lambda_2 = a_2/(z_t - z_b)$, z_t is the cloud top height, and a_1 and a_2 are empirically determined constants.

For shallow convection, a simplified version of Eq. (3.2.40) is adopted for closure.

3.2.4.4 Triggering Mechanism

The convective triggering mechanism proposed by Xie and Zhang (2000) known as the dynamic CAPE generation rate (DCAPE) is used in the cumulus parameterization. DCAPE is defined as follows:

$$\text{DCAPE} = \left(\text{CAPE}\left(T^*, q^*\right) - \text{CAPE}\left(T, q\right)\right)/\Delta t \tag{3.2.44}$$

where T is the temperature, q is the specific humidity, and (T^*, q^*) are (T, q) plus the change caused by overall large-scale advection over a certain time period Δt (the integration time step used in the model). These values are equivalent to (T, q) just after dynamics calculation. CAPE is defined as

$$\text{CAPE} = \int_{z_{\text{LFC}}}^{z_{\text{LNB}}} g\frac{T_v^u - T_v}{T_v}dz \tag{3.2.45}$$

where z_{LFC} and z_{LNB} are the height of the level of free convection and that of neutral buoyancy, respectively, g is the acceleration of gravity, and T_v is the virtual temperature. The superscript u denotes a lifted air parcel. Deep convection is prohibited when DCAPE does not exceed an empirically determined critical value.

3.2.4.5 Convective Downdraft

For reasons of economy, only one type of downdraft is assumed, while many types are considered in the updraft scheme.

Downdraft is initiated at the level where the net upward mass flux is reduced to half that at the cloud base. The downdraft mass flux M^d at the cloud base is given by,

$$M^d = 0.4M_B \tag{3.2.46}$$

where M_B is the net mass flux at the cloud base of updraft as calculated using Eq. (3.2.40).

Entrainment from the environment is assumed to occur above the cloud base, while detrainment is assumed to occur both above and below it. The entrainment and detrainment rates are set to the same constant value above the cloud base, so that the mass flux of downdraft is constant with height.

3.2.4.6 Mid-level Convection

A mid-level convection scheme is incorporated to represent cumulus convection with a cloud base on a frontal system in the extratropics. The height of the cloud base is given by the maximum moist static energy level in the vertical column, and the cloud top is defined as the level where an air mass rising from the cloud base with a constant entrainment rate loses buoyancy. Closure employs a simpler form of Eq. (3.2.40).

3.2.4.7 Convective Momentum Transport

Convective momentum transport is parameterized in a different way from heat and moisture transport. A multiple plume model is adopted both for updraft and downdraft. The entrainment and detrainment rates are set to the same value both for updraft and downdraft between the cloud base and the cloud top, making each mass flux constant with altitude. The magnitude of each updraft M_{cn}^{u} is set to the mass flux at the cloud base in the heat and moisture transport scheme as calculated using Eq. (3.2.40), and the magnitude of the downdraft is set to $0.4 \times M_{cn}^{u}$.

3.2.4.8 Melting and Re-evaporation of Precipitation

Melting of snow is calculated below freezing level, with formulation similar to that of the cloud scheme (Eq. (3.2.57)). Re-evaporation of precipitation is considered below the cloud base. Related calculation is based on the equation used in the cloud scheme (Eq. (3.2.59)) with a minor modification.

3.2.5 Clouds and Large-scale Precipitation

Clouds are prognostically determined in a fashion similar to that proposed by Smith (1990). The simple statistical approach proposed by Sommeria and Deardorff (1977) is adopted for the calculation of cloud amounts and their water content.

3.2.5.1 Cloud Scheme

Representing conserved quantities in phase transition between water vapor and cloud water, the total water content (water vapor and cloud water) q_w and the liquid water temperature T_L are defined as follows:

$$q_w = q_v + q_c \tag{3.2.47}$$

$$T_L = T - \frac{L}{C_p} q_c \tag{3.2.48}$$

where q_v is specific humidity, q_c is cloud water content, T is temperature, L is the latent heat of condensation and C_p is specific heat at a constant pressure. In each grid box, q_w is assumed to vary due to unresolved atmospheric fluctuations having a probability distribution function with a top-hat shape. The cloud fraction C is given by the part of the grid box where q_w exceeds the saturation specific humidity q_s, and cloud water content is given as the condensation amount in the grid box:

$$C = \frac{a_L \left(\overline{q_w} - q_s\left(T_L\right)\right) + \Delta q_w}{2\Delta q_w} \tag{3.2.49a}$$

$$q_c = C^2 \Delta q_w \tag{3.2.49b}$$

$$a_L = \frac{1}{1 + \frac{L}{C_p}\left(\frac{\partial q_s}{\partial T}\right)_{T=T_L}} \tag{3.2.49c}$$

where Δq_w is the maximum local deviation from the grid-box mean total water content $\overline{q_w}$, the overbar denotes an average over the grid box, and C is under the constraint $0 \leq C \leq 1$. Δq_w is calculated as follows:

$$\Delta q_w = \frac{a_L}{2}\left(\overline{q_w'^2} - 2b\,\overline{q_w' s_l'} + b^2 \overline{s_l'^2}\right)^{\frac{1}{2}} \tag{3.2.50}$$

where $s_l = C_p T_L + gz$ is dry static energy, g is acceleration under gravity, z is height above the surface, $b = \left(\frac{\partial q_s}{\partial T}\right)_{T=T_L} / C_p$, and the prime denotes a deviation from the grid average. $\overline{q_w'^2}$, $\overline{q_w' s_l'}$, and $\overline{s_l'^2}$ are calculated using the level 2 turbulence closure model of Mellor and Yamada (see Subsection 3.2.7) in the boundary layer scheme. Δq_w is limited by q_s:

$$0.2A\, a_L\, q_s\,(T_L) \;\leq\; \Delta q_w \;\leq\; 0.5A\, a_L\, q_s\,(T_L) \tag{3.2.51}$$

$$A \;=\; \min\left(\frac{p_s - p}{p_s - 850}, 1\right) \tag{3.2.52}$$

where p_s is surface pressure and p is pressure.

3.2.5.2 Stratocumulus Scheme

For the representation of subtropical marine stratocumulus clouds, the stratocumulus scheme proposed by Kawai and Inoue (2006) is adopted to diagnose the cloud fraction C and the cloud water content q_c instead of the cloud scheme:

$$C = 12.0\left(-\frac{\partial \theta}{\partial p} - 0.07\right) \tag{3.2.53a}$$

$$q_c = 0.05\, a_L\, C\, q_s \tag{3.2.53b}$$

where θ is the potential temperature. This scheme works when the following three conditions in each model layer and two conditions near the surface are satisfied:

[in each model layer]

- $-\frac{\partial \theta}{\partial p} > 0.07$ [K/hPa] (just above the model layer)
- relative humidity ≥ 80 [%]
- model layer height below the 924 hPa level

[near the surface]

- $-\frac{\partial \theta}{\partial p} < 0.01$ [K/hPa]
- not completely covered with sea ice

The first condition in each model layer represents the formation of marine stratocumulus clouds under a strong inversion layer, and the second and third prevent the false representation of stratocumulus clouds over dry area and shallow convection area, respectively. The conditions near the surface prevents the false representation of stratocumulus clouds over land or sea ice during the night.

In the cloud scheme and the stratocumulus scheme, liquid (ice) cloud is assumed to be present when the temperature is above 0°C (below −15°C). Between −15°C and 0°C, mixed-phase cloud is assumed to be present and the mixing ratio changes in a linear fashion with temperature.

3.2.5.3 Cloud Ice Fall and Conversion to Precipitation

The prognostic equation for cloud water content is as follows:

$$\frac{\partial q_c}{\partial t} = C_g + I - O \tag{3.2.54}$$

where C_g represents the generation rate of clouds in the grid box, I is the inflow from the upper layer, and O is the downward outflow. C_g is calculated using the results described in Subsection 3.2.5.1 or Subsection 3.2.5.2.

For ice-only cloud, I and O are calculated based on Kawai (2005). Small cloud particles ($\leq$100 μm) fall into the lower layer, while large ones (>100 μm) fall to the surface immediately as snow:

$$O = \frac{v_{cice} q_c}{\Delta z} + D_{I2S}\, q_c \tag{3.2.55}$$

where Δz is the layer thickness, v_{cice} is the terminal velocity of a small ice particle in cloud, and D_{I2S} is the conversion rate from ice cloud to snow. The first term on the right becomes I in the next layer down.

For mixed-phase cloud or water-only cloud, parameterization for the rate of conversion from cloud water to precipitation P follows the scheme proposed by Sundqvist (1978):

$$\begin{aligned} P &= \frac{1}{\tau_p} q_c \left[1 - \exp\left\{ -\left(\frac{q_c}{C q_c^{\text{crit}}} \right)^2 \right\} \right] \\ &= O \end{aligned} \tag{3.2.56}$$

where τ_p represents a characteristic time scale for the conversion of cloud droplets into raindrops and snowflakes, and q_c^{crit} is the critical cloud water content at which the release of precipitation becomes efficient. In this case, I is not considered because raindrops and snowflakes are assumed to fall to the surface immediately. The coalescence process (collection of cloud droplets by raindrops falling through a cloud) and the Bergeron-Findeisen effect (enhancement of precipitation release in clouds containing a mixture of droplets and ice crystals) are modeled following Sundqvist *et al.* (1989).

3.2.5.4 Melting and Evaporation

The snow melting rate M is parameterized using the same method as ECMWF (2014):

$$M = 0.5 \frac{C_p}{L_m} \frac{T_w - T_0}{\tau_m} \tag{3.2.57}$$

$$\tau_m = \frac{7200}{1 + 0.5\,(T_w - T_0)} \tag{3.2.58}$$

where T_w is the wet-bulb temperature, T_0 is the melting temperature, L_m is the latent heat of fusion, and τ_m is the relaxation time of melting. Based on Kessler (1969) and Tiedtke (1993), the evaporation rate E for large-scale precipitation is parameterized as

$$E = b \frac{1}{\tau_e} (q_s - q_v) \left\{ \left(\frac{p}{p_s} \right)^{1/2} \frac{1}{5.09 \times 10^{-3}} \frac{P_l}{b} \right\}^{0.577} \tag{3.2.59}$$

$$\frac{1}{\tau_e} = 5.44 \times 10^{-4} \tag{3.2.60}$$

where b is the clear-sky precipitation fraction (set to 0.5), τ_e is the relaxation time of evaporation, and P_l is the local precipitation rate.

3.2.6 Surface Turbulent Fluxes

Surface fluxes are represented with bulk formula based on Monin–Obukhov (M–O) similarity theory, with momentum flux F_m, heat flux F_h and specific humidity flux F_q expressed as follows:

$$\mathrm{F}_m = \overline{(w'\mathbf{v}')}_s = -C_m |\mathbf{v}_1| \mathbf{v}_1 \,, \tag{3.2.61}$$

$$F_h = \overline{(w'\theta')}_s = -C_h |\mathbf{v}_1| (\theta_1 - \theta_s) \,, \tag{3.2.62}$$

$$F_q = \overline{(w'q')}_s = -C_h |\mathbf{v}_1| (q_1 - q_s) \,. \tag{3.2.63}$$

Here $\mathbf{v} = (u, v)$ represents horizontal wind, θ is potential temperature and q is specific humidity, and the subscripts "1" and "s" indicate variables at the lowest level of the model grid and at the ground surface, respectively. C_m and C_h are exchange coefficients for momentum and heat.

According to M–O similarity theory, the exchange coefficients are defined as follows:

$$\frac{z_1}{L}\frac{\left[\log\left(\frac{z_1+z_{0m}}{z_{0h}}\right)-\Psi_h\left(\frac{z_1+z_{0m}}{L}\right)+\Psi_h\left(\frac{z_{0h}}{L}\right)\right]}{\left[\log\left(\frac{z_1+z_{0m}}{z_{0m}}\right)-\Psi_m\left(\frac{z_1+z_{0m}}{L}\right)+\Psi_m\left(\frac{z_{0m}}{L}\right)\right]^2}=\frac{gz_1}{|\mathbf{v}_1|^2}\frac{2(\theta_{v1}-\theta_{vs})}{(\theta_1+\theta_s)}\quad, \tag{3.2.64}$$

$$C_m=\frac{\kappa^2}{\left[\log\left(\frac{z_1+z_{0m}}{z_{0m}}\right)-\Psi_m\left(\frac{z_1+z_{0m}}{L}\right)+\Psi_m\left(\frac{z_{0m}}{L}\right)\right]^2}\quad, \tag{3.2.65}$$

$$C_h=\frac{\kappa^2}{\left[\log\left(\frac{z_1+z_{0m}}{z_{0m}}\right)-\Psi_m\left(\frac{z_1+z_{0m}}{L}\right)+\Psi_m\left(\frac{z_{0m}}{L}\right)\right]\left[\log\left(\frac{z_1+z_{0m}}{z_{0h}}\right)-\Psi_h\left(\frac{z_1+z_{0m}}{L}\right)+\Psi_h\left(\frac{z_{0h}}{L}\right)\right]}\quad, \tag{3.2.66}$$

where κ is von Kármán's constant (= 0.4), g is the standard acceleration due to gravity (= 9.80665 [m/s^2]), z_1 is the height of the lowest level of the model grid above the ground, and θ_v is the virtual potential temperature, while z_{0m} and z_{0h} are the surface momentum and heat roughness lengths, respectively. Eq. (3.2.64) gives Obukhov length L from the prognostic variables on right. The stability functions Ψ_m and Ψ_h are parameterized by Beljaars and Holtslag (1991) as follows:

$$x\equiv(1-16\xi)^{\frac{1}{4}} \tag{3.2.67}$$

$$\Psi_m(\xi)=\begin{cases}\dfrac{\pi}{2}-2\arctan(x)+\log\dfrac{(1+x)^2(1+x^2)}{8} & \xi<0\\ -\dfrac{2}{3}\left(\xi-\dfrac{5}{0.35}\right)e^{-0.35\xi}-\xi-\dfrac{2}{3}\dfrac{5}{0.35} & \xi\geq 0\end{cases}\quad, \tag{3.2.68}$$

$$\Psi_h(\xi)=\begin{cases}2\log\dfrac{1+x^2}{2} & \xi<0\\ -\dfrac{2}{3}\left(\xi-\dfrac{5}{0.35}\right)e^{-0.35\xi}-\left(1+\dfrac{2}{3}\xi\right)^{\frac{3}{2}}-\dfrac{2}{3}\dfrac{5}{0.35}+1 & \xi\geq 0\end{cases}\quad. \tag{3.2.69}$$

Over land grids, surface parameters are determined using the land model on each grid in consideration of vegetation type, soil conditions, and snow cover (Subsection 3.2.10). Over ocean grids, surface fluxes are calculated separately for the different subgrid surface fractions. Tiling between open water and sea ice is used with the coupling approach suggested by Best *et al.* (2004).

Surface wind stress depends on oceanic waves excited by surface winds. Roughness length and wind-induced stress are iteratively calculated in the model. Following the method of Beljaars (1995), surface roughness lengths over open sea (ice-free ocean) are determined from Charnock's relation (Charnock 1955):

$$\begin{aligned}z_{0m}&=\frac{0.11\nu}{u_*}+\frac{\alpha}{g}u_*^2\quad,\\ z_{0h}&=\frac{0.62\nu}{u_*}\quad,\end{aligned} \tag{3.2.70}$$

where $u_*\left(\equiv\sqrt{\left|\overline{(w'\mathbf{v}')_s}\right|}\right)$ is the friction velocity, ν is the kinematic viscosity of air (= 1.5×10^{-5}m^2/s) and α is the Charnock coefficient (= 0.020). The surface roughness lengths over sea ice are fixed at 0.001m for momentum and 0.0005m for heat.

3.2.7 Boundary layer (turbulent transport)

A hybrid approach between turbulence kinetic energy (TKE) closure and eddy diffusivity (ED) type scheme is employed to parameterize vertical turbulent transports of momentum, heat and moisture in the atmosphere. The TKE scheme used is the level 2 turbulence closure approach of Mellor and Yamada (1974, 1982), and the

ED–type scheme involves the use of stability functions based on Han and Pan (2011). Turbulent transports are expressed as

$$\overline{w'\mathbf{v}'} = -\max(K_m^{TKE}, K_m^{ED})\frac{\partial \mathbf{v}}{\partial z}\ , \tag{3.2.71}$$

$$\overline{w's_L'} = -\max(K_h^{TKE}, K_h^{ED})\frac{\partial s_L}{\partial z}\ , \tag{3.2.72}$$

$$\overline{w'q_w'} = -\max(K_h^{TKE}, K_h^{ED})\frac{\partial q_w}{\partial z}\ , \tag{3.2.73}$$

where $s_L(\equiv C_pT + gz - Lq_c)$ is the liquid water static energy, $q_w(\equiv q + q_c)$ is the total water content, the superscripts "TKE" and "ED" indicate the scheme types for calculating diffusion coefficients K, and the subscripts "m" and "h" indicate momentum and heat, respectively.

Following mixing-length theory, the diffusion coefficients of both schemes can be written as

$$K_m = l^2\left|\frac{\partial \mathbf{v}}{\partial z}\right| f_m\ , \tag{3.2.74}$$

$$K_h = l^2\left|\frac{\partial \mathbf{v}}{\partial z}\right| f_h\ , \tag{3.2.75}$$

where f_m and f_h are stability functions and the mixing length l is given according to Blackadar (1962) as

$$l = \frac{\kappa z}{1 + \kappa z/l_0}, \tag{3.2.76}$$

where κ is von Kármán's constant (= 0.4).

In the TKE scheme, the mixing length l_0 is calculated from sub-grid scale orographic variances and the planetary boundary layer depth. l_0 is a constant equal to 50 m in the ED–type scheme.

The stability functions in the ED type scheme are functions of the gradient Richardson number R_i, given as

$$f_m^{ED} = \begin{cases} \dfrac{1 + 2.1R_i}{(1 + 5R_i)^{1.5}} & R_i \geq 0 \\ 1 - \dfrac{8R_i}{1 + 1.746\sqrt{-R_i}} & R_i < 0 \end{cases}, \tag{3.2.77}$$

$$f_h^{ED} = \begin{cases} \dfrac{1}{(1 + 5R_i)^{1.5}} & R_i \geq 0 \\ 1 - \dfrac{8R_i}{1 + 1.286\sqrt{-R_i}} & R_i < 0 \end{cases}. \tag{3.2.78}$$

Those of the TKE scheme can be written as follows:

$$f_m^{TKE} = S_M\sqrt{B_1 S_M(1 - R_f)}, \tag{3.2.79}$$

$$f_h^{TKE} = S_H\sqrt{B_1 S_M(1 - R_f)}, \tag{3.2.80}$$

$$S_M = A_1F_1\frac{RF_1 - R_f}{A_2F_2(RF_2 - R_f)}S_H,$$

$$S_H = 3A_2\frac{(\gamma_1 + \gamma_2)(RF_c - R_f)}{1 - R_f},$$

$$R_f = RI_1\left(\hat{R}_i + RI_2 - \sqrt{\hat{R}_i(\hat{R}_i - RI_3) + RI_2^2}\right),$$

$$RF_1 = B_1\frac{\gamma_1 - C_1}{F_1}, \quad RF_2 = B_1\frac{\gamma_1}{F_2}, \quad RF_c = \frac{\gamma_1}{(\gamma_1 + \gamma_2)},$$

$$RI_1 = \frac{1}{2}\frac{A_2F_2}{A_1F_1}, \quad RI_2 = B_1\gamma\frac{(\gamma_1 - C_1)}{F_1}\frac{A_1F_1}{A_2F_2}, \quad RI_3 = 4B_1\frac{\gamma_1}{F_2}\frac{A_1F_1}{A_2F_2} - 2RI_2,$$

$$F_1 = B_1(\gamma_1 - C_1) + 2A_1(3 - 2C_2) + 3A_2(1 - C_2), \quad F_2 = B_1(\gamma_1 + \gamma_2) - 3A_1(1 - C_2),$$
$$\gamma_1 = \frac{1}{3} - 2\frac{A_1}{B_1}, \quad \gamma_2 = \frac{B_2}{B_1}(1 - C_3) + 2\frac{A_1}{B_1}(3 - 2C_2),$$

where $A_1(= 1.0)$, $A_2(= 0.58)$, $B_1(= 24.0)$, $B_2(= 11.0)$, $C_1(= 0.13)$, $C_2(= 0.6)$, and $C_3(= 0.14)$ are the closure constants of the TKE scheme. The modified gradient Richardson number $\hat{R}_i$ used instead of R_i is defined after the method of Smith (1990):

$$\hat{R}_i = g\left\{\tilde{\beta}_s \frac{\partial s_L}{\partial z} + \tilde{\beta}_Q \frac{\partial q_w}{\partial z}\right\} \Big/ \left|\frac{\partial \mathbf{v}}{\partial z}\right|^2 , \tag{3.2.81}$$

where $\tilde{\beta}_s$ and $\tilde{\beta}_Q$ are buoyancy parameters in terms of the cloud-conserved quantities s_L and q_w, respectively.

3.2.8 Sub-grid Orography

The effect of orographic drag from sub-grid orography on atmospheric flow is separately parametrized depending on scales of unresolved orography. Gravity wave drag (GWD) from sub-grid orography with scales larger than 5 km (the minimum scale for vertically propagating gravity waves under typical conditions), and blocked flow drag (BFD) associated with the sub-grid orography are based on Lott and Miller (1997), and turbulent orographic form drag (TOFD) caused by sub-grid orography with scales smaller than this is parametrized in line with Beljaars *et al.* (2004a). For numerical stability, BFD and TOFD tendencies are implicitly computed via a tridiagonal solver in the boundary layer scheme by passing the linear coefficients, and the GWD tendency is passed to the solver as a forcing term (Subsection 3.2.8.5).

3.2.8.1 Turbulent Orographic Form Drag

TOFD formulation is based on the spectra of topography over North America and related integrals in the wavenumber space. An approximated equation is used for simplicity, and TOFD involves explicit application of drag on model levels in the equation for horizontal wind u (m s^{-1})

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial z}\frac{\tau_x}{\rho} = -C_{\mathrm{tofd}} u|u|, \tag{3.2.82}$$
$$C_{\mathrm{tofd}} = \alpha\beta C_{\mathrm{md}} C_{\mathrm{corr}} 2.109 e^{-(z/1500)^{1.5}} a_2 z^{-1.2}$$

where τ_x is stress (N m^{-2}) in the x-direction, ρ is atmospheric density (kg m^{-3}), z is geopotential height (m). $\alpha = 24$ and $\beta = 1$ are a shear-dependent parameter and a shape factor, respectively. $C_{\mathrm{md}} = 0.005$ is a drag coefficient, $C_{\mathrm{corr}} = 0.6$ is a correction parameter, $a_2 = a_1 k^{n_1 - n_2}$, $a_1 = \sigma_{\mathrm{flt}}^2 (I_H k_{\mathrm{flt}}^{n_1})^{-1}$, $k_1 = 0.003$ (m^{-1}), $n_1 = -1.9$, $n_2 = -2.8$, $k_{\mathrm{flt}} = 0.00035$ (m^{-1}), $I_H = 0.00102$ (m^{-1}) and σ_{flt} (m) is the standard deviation of filtered sub-grid orography. The same equation also holds for meridional wind v (m s^{-1}). For numerical stability, the equation is solved implicitly and simultaneously in the vertical diffusion scheme (Subsection 3.2.8.5).

The standard deviation of filtered sub-grid orography for TOFD, σ_{flt}, is calculated from two filtered fields where different scales are removed from a global $30'' \times 30''$ orographic source dataset (Kanehama *et al.* 2023b) (see Subsection 3.2.12). A smoothing filter is defined in the grid-point space by

$$h(r) = \begin{cases} \frac{1}{\Delta} & \text{for} \quad r \le \Delta/2 - \delta \\ \frac{1}{2\Delta} + \frac{1}{2\Delta}\cos\left\{\frac{\pi(r - \frac{\Delta}{2} + \delta)}{2\delta}\right\} & \text{for} \quad \Delta/2 - \delta < r < \Delta/2 + \delta \\ 0 & \text{for} \quad r \ge \Delta/2 + \delta \end{cases} \tag{3.2.83}$$

where r is radial distance (m), Δ is filter width (m) and δ is filter edge width (m). The parameter Δ controls extracting scales, and δ is applied to smooth data around the filter edge. To extract the necessary scales, the filter is applied twice to the source dataset. Filtering with $\Delta_1 = 2$ and $\delta_1 = 1$ is applied first to remove smaller scales smaller than 2 km (h_2), and with $\Delta_2 = 20$, $\delta_2 = 1$ is applied to extract scales larger than 20 km (h_{20}). The

smoothing functions with these parameters are displayed in Figure 3.2.1. The standard deviation of filtered orography σ_{flt} is calculated as

$$\sigma_{\text{flt}} = \sqrt{\sum_i w_i (d_i - \overline{d})^2 \Big/ \sum_i w_i} \tag{3.2.84}$$

where $d = h_2 - h_{20}$ and $\overline{d} = \sum_i w_i d_i / \sum_i w_i$, $w_i = \cos\phi$ is latitudinal weight, with ϕ being latitude, and summation over i is taken over all points within a model grid box. A map of the filtered standard deviation for TOFD at TQ959 is shown in Figure 3.2.2

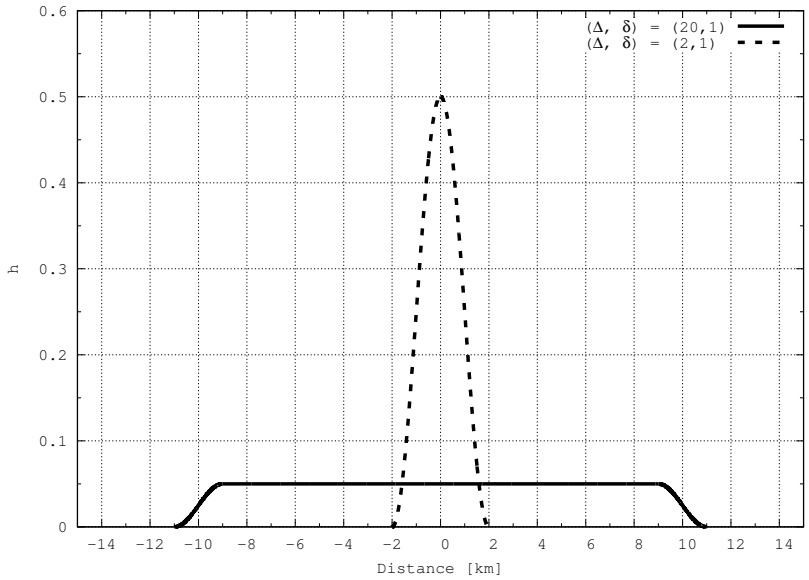


Figure 3.2.1: Shape of smoothing functions defined in Eq. (3.2.83) with the values of $(\Delta, \delta) = (20, 1)$ for the straight line and $(\Delta, \delta) = (2, 1)$ for the dotted line.

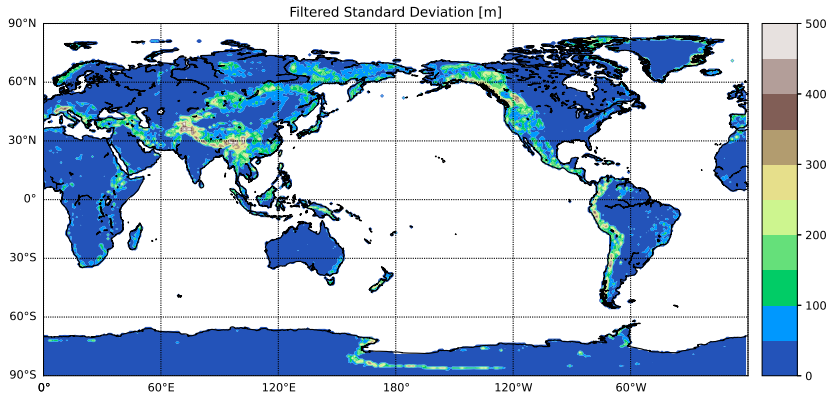


Figure 3.2.2: Filtered standard deviation [m] for TOFD at TQ959. Only land grid points are plotted.

3.2.8.2 Blocked Flow Drag

In the sub-grid scale orographic (SSO) drag scheme, it is assumed that the effect of sub-grid orography on atmospheric flow is equivalently represented by the effect of one elliptic mountain (Baines and Palmer 1990).

The scheme deals with two regimes, one is gravity wave drag exerted by topography, transporting momentum upwards and depositing momentum when breaking or filtering at a critical level. The other is a flow-blocking drag caused by flow going around, rather than over, orography. These two regimes are separated using non-dimensional mountain height as outlined below.

In the SSO scheme, it is assumed that flow is lateral when the non-dimensional mountain height is sufficiently high. This corresponds to high atmospheric stability, low incident flow speed, or mountain terrain high enough to suppress the flow crossing the elliptic mountain top. The maximum flow height associated with the blocking height Z_{blk} (m) is evaluated to satisfy inequality between non-dimensional mountain height and the critical value $H_{n_{\text{crit}}}$

$$\int_{Z_{\text{blk}}}^{H} \frac{N}{u_p} dz \geq H_{n_{\text{crit}}} \tag{3.2.85}$$

where N is Brunt-Vaisala frequency (s^{-1}), u_p is wind speed in the direction of incident flow (m s^{-1}), and z is geopotential height. H (m) is a statistical mountain height representing the peak sub-grid orography height, and is set as 3μ based on the standard deviation of sub-grid orography μ (m), and $H_{n_{\text{crit}}}$ is 0.5. Eq. (3.2.85) is evaluated from the top, and Z_{blk} is determined as the first level that satisfying this formula. The blocked flow drag D_{blk} is directly applied to model levels and formulated as

$$D_{\text{blk}}(z) = -C_d \max\left(5 - \frac{1}{r^3}, 0\right) \rho \frac{\sigma}{2\mu} \left(\frac{Z_{\text{blk}} - z}{z + \mu}\right)^{1/2} (B\cos^2\psi + C\sin^2\psi) \frac{u|u|}{2} \tag{3.2.86}$$

where C_d is a drag coefficient of 2. r is the aspect ratio of the elliptic mountain as seen from incident flow, ρ is air density, σ is the orographic slope along the short elliptic axis and ψ is the angle (rad) between incident flow with the angle φ and the principal axis of the elliptic with the angle θ, i.e., $\psi = \theta - \varphi$. Following Phillips (1984), B and C are

$$B = 1 - 0.18\gamma - 0.04\gamma^2, \; C = 0.48\gamma + 0.30\gamma^2 \tag{3.2.87}$$

where γ is the elliptic aspect ratio. The aspect ratio seen from the incident flow r is formulated as

$$r = \sqrt{\frac{\cos^2\psi + \gamma^2 \sin^2\psi}{\gamma^2 \cos^2\psi + \sin^2\psi}} \tag{3.2.88}$$

The tendency of blocked flow drag is defined by

$$\frac{\partial u}{\partial t} = \frac{D_{\text{blk}}}{\rho} \tag{3.2.89}$$

The same equation applies for meridional wind v as well as zonal wind u. For numerical stability, it is solved implicitly and simultaneously in the vertical diffusion scheme (Subsection 3.2.8.5).

3.2.8.3 Gravity Wave Drag

The gravity wave drag part of the SSO scheme in Lott and Miller (1997) is similar to that proposed by Palmer *et al.* (1986), but is modified to represent the effects of anisotropic mountains (Baines and Palmer 1990). As flow under the Z_{blk} level is blocked by sub-grid elliptic mountains and is assumed to not cross mountain tops, the amplitude of gravity wave drag exerted by elliptic mountains is equal to the effective mountain height $H_{\text{eff}} = H - Z_{\text{blk}}$. The factor $n_{\text{eff}} = 1$ is applied to adjust gravity wave stress, i.e., $H_{\text{eff}} = n_{\text{eff}}(H - Z_{\text{blk}})$. Following Phillips (1984), gravity wave surface stress (τ_1, τ_2) parallel and perpendicular to incident flow can be formulated as

$$(\tau_1, \; \tau_2) = \rho_l |u_l| N_l \left(\frac{H_{\text{eff}}}{3}\right)^2 \frac{\sigma}{\mu} G \left\{B\cos^2\psi_l + C\sin^2\psi_l, \; (B - C)\sin\psi_l \cos\psi_l\right\} \tag{3.2.90}$$

where $G = 0.6$ is a function of mountain sharpness and the subscript l denotes the low-level mean between μ to 2μ. Note that Eq. (3.2.90) involves division by 9 rather than 4 as originally proposed in Lott and Miller (1997) so as to scale Eq. (3.2.90) with H being able to reach 3μ at its maximum when Z_{blk} is zero.

τ_{sfc}, the magnitude of stress at the wave generation level, is

$$\begin{aligned}\tau_{\text{sfc}} &= \rho_l |u_l| N_l \left(\frac{H_{\text{eff}}}{3}\right)^2 \frac{\sigma}{\mu} G(D1^2 + D2^2)^{1/2} \\ D1 &= B\cos^2\psi_l + C\sin^2\psi_l \\ D2 &= (B - C)\sin\psi_l \cos\psi_l\end{aligned} \tag{3.2.91}$$

For momentum deposit, the GWD scheme incorporates critical level filtering and wave saturation in upward propagation. Unless gravity waves encounter these processes, waves are assumed to propagate upward conservatively. When the background wind changes sign in propagating upward from the wave generation level, waves are assumed to reach a critical level and overall wave stress is deposited to the background.

The wave saturation process is implemented with comparison of surface stress (Eq. (3.2.91)) and saturated stress as per Palmer *et al.* (1986) and Iwasaki *et al.* (1989). Saturated stress τ_{sat} is estimated using the background Richardson number $R_i = N^2 \Big/ \left(\frac{\partial u}{\partial z}\right)^2$ and the local Richardson number $R_{i,\text{loc}}$. Assuming wave phase difference, the relationship between R_i and $R_{i,\text{loc}}$ is

$$R_{i,\text{loc}} = R_i \left(\frac{1 - \alpha}{(1 + R_i^{1/2}\alpha)^2}\right) \tag{3.2.92}$$

where $\alpha = N|\delta z|/V$, $|\delta z|$ is gravity wave amplitude (m) and V is wind speed (m s^{-1}) at a particular level in the direction of gravity wave stress. $R_{i,\text{loc}} = 0.25$ when the amplitude of gravity waves is saturated, and solving Eq. (3.2.92) with the saturated amplitude $|\delta z|_{\text{sat}}$ gives

$$|\delta z|_{\text{sat}} = \frac{V}{N}\left\{2(R_i^{-\frac{1}{2}} + 2)^{\frac{1}{2}} - (R_i^{-\frac{1}{2}} + 2)\right\} \tag{3.2.93}$$

Thus saturated flux τ_{sat} is

$$\tau_{\text{sat}} = \rho |u| N \left(\frac{H_{\text{sat}}}{3}\right)^2 \frac{\sigma}{\mu} G(D1^2 + D2^2)^{1/2} \tag{3.2.94}$$

with the effective amplitude H_{sat} as the saturated amplitude $H_{\text{sat}} = |\delta z|_{\text{sat}}$. τ_{sfc} and τ_{sat} are compared at each half level $k + 1/2$, and the residual momentum flux $\Delta\tau = \tau_{\text{sfc}} - \tau_{\text{sat}}$ is considered to be deposited to the background at the full level k if the condition $\tau_{\text{sfc}} > \tau_{\text{sat}}$ is satisfied. τ_{sat} is set as a new value for gravity wave stress propagating vertically. Otherwise, τ_{sfc} is considered to conservatively propagate vertically. The deposited momentum flux $\Delta\tau$ is distributed over a width equivalent to the diagnosed vertical wavelength λ_z (m) of hydrostatic gravity waves (Vosper 2015; Epifanio and Qian 2008).

$$\lambda_z = 2\pi \frac{u_1}{N} \tag{3.2.95}$$

where u_1 (m s^{-1}) is the wind component parallel to low-level wind. The tendency is

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho}\frac{\partial \tau_\lambda}{\partial z} \tag{3.2.96}$$

where τ_λ is the sum of distributed deposited momentum flux. Distributing deposited momentum flux also increases numerical stability by inhibiting the GWD scheme to produce extremely large drag at a single level.

3.2.8.4 Ancillary Fields for Blocked Flow Drag and Gravity Wave Drag

In the Lott and Miller (1997) scheme, sub-grid orography is assumed to be represented by one elliptic mountain with standard deviation μ, orographic slope along the short axis σ, an aspect ratio γ and a long-axis angle θ.

As the scales included in SSO schemes are larger than 5 km (the minimum scale allowing vertical gravity wave propagation in typical conditions), scales smaller than this are removed by averaging the 30" (approx. 1 km) source dataset to a 2'30" resolution h_{org}, (approx. 5 km). The deviation of elevation (h) in the averaged 2'30" data from mean orography is defined as $h = h_{org} - h_m$, with h_m being grid-box mean orography, i.e., $h_m = \overline{h}_{org}$. Overlines indicate the grid-box mean, with orography linearly interpolated into 2'30" grid boxes in consideration of differences.

The parameters γ and θ are derived using variables K, L and M

$$K = \frac{1}{2}\left\{\overline{\left(\frac{\partial h}{\partial x}\right)^2} + \overline{\left(\frac{\partial h}{\partial y}\right)^2}\right\},\ L = \frac{1}{2}\left\{\overline{\left(\frac{\partial h}{\partial x}\right)^2} - \overline{\left(\frac{\partial h}{\partial y}\right)^2}\right\},\ M = \overline{\frac{\partial h}{\partial x}\frac{\partial h}{\partial y}} \tag{3.2.97}$$

Horizontal derivatives are computed in the grid-point space using adjacent grids in the latitudinal direction. However, in the longitudinal direction, equidistant points are computed via linear interpolation at each grid point in consideration of derivatives because the distance becomes smaller toward higher latitudes. The principal axis angle θ is computed using

$$\theta = \frac{1}{2}\arctan(\mathrm{M/L}) \tag{3.2.98}$$

Applying coordinate transform with respect to θ gives coordinates along the short and long elliptic axes as

$$x' = x\cos\theta + y\sin\theta,\ y' = y\cos\theta - x\sin\theta \tag{3.2.99}$$

and the transformed coordinates involves introduction of new K',L' and M' values

$$K' = K,\ L' = (L^2 + M^2)^{1/2} \text{ and } \mathrm{M}' = 0$$

The aspect ratio γ is defined as

$$\begin{aligned}\gamma^2 &= \overline{\left(\frac{\partial h}{\partial y'}\right)^2} \Big/ \overline{\left(\frac{\partial h}{\partial x'}\right)^2} \\ &= \frac{K' - L'}{K' + L'} = \frac{K - (L^2 + M^2)^{1/2}}{K + (L^2 + M^2)^{1/2}}\end{aligned} \tag{3.2.100}$$

The orographic slope along the short-axis σ is defined by

$$\sigma^2 = \overline{\left(\frac{\partial h}{\partial x'}\right)^2} \tag{3.2.101}$$

and the standard deviation of sub-grid orography μ is computed as

$$\mu = \sqrt{\sum_i w_i(h_i - \overline{h})^2 \Big/ \sum_i w_i} \tag{3.2.102}$$

where $\overline{h} = \sum_i w_i h_i / \sum_i w_i$. Maps of the four parameters at TQ959 are shown in the Figure 3.2.3.

3.2.8.5 Joint Implicit Calculation

For numerical stability, Eq. (3.2.82) and Eq. (3.2.89) are solved implicitly, meaning that the absolute wind speed $|u|$ is taken from the current time step and the u and v components are evaluated at the next time step. The coefficient for the component at the new time level is passed to a tridiagonal implicit solver in the boundary layer and solved implicitly to increase consistency among multiple processes involved and further enhance numerical stability via time integration. The tendency for gravity wave drag is also passed to the solver as a forcing term so that the fast processes can adjust among multiple processes. By solving simultaneously, the

tendencies calculated via the solver in the boundary layer scheme include the planetary boundary layer (pbl), blocked flow drag (bfd), turbulent orographic form drag (tofd) and gravity wave drag (gwd). Among these, the tendency of gravity wave drag as a forcing term can be subtracted from total tendency. Accordingly, the implicitly computed tendency can be separated as

$$\left(\frac{\partial u}{\partial t}\right)_{\text{pbl+bfd+tofd}} = \left(\frac{\partial u}{\partial t}\right)_{\text{pbl+gwd+bfd+tofd}} - \left(\frac{\partial u}{\partial t}\right)_{\text{gwd}} \tag{3.2.103}$$

3.2.9 Non-orographic Gravity Wave Drag

The parameterization of non-orographic gravity wave drag follows the scheme proposed by Scinocca (2003). Assuming hydrostatic gravity waves in the absence of rotation, the vertical flux of momentum directed into the ϕ azimuth $\tilde{F}$ is expressed as follows:

$$\rho\tilde{F} = \rho A \frac{\tilde{c} - \tilde{U}}{N} \left(\frac{\tilde{c} - \tilde{U}}{\tilde{c}}\right)^{2-p} \frac{1}{1 + \left[\frac{m^*(\tilde{c}-\tilde{U})}{N}\right]^{s+3}} \tag{3.2.104}$$

where,

$$A = Cm^{*3} \frac{{N_0}^{2-p} - f^{2-p}}{2 - p} \tag{3.2.105}$$

ρ is density, $m^*(= 2\pi/2000$ [1/m]) is the characteristic vertical wavenumber, N is the Brunt-Väisälä frequency, f is the Coriolis parameter, p(=1.5) and s(=1) are constants, $\tilde{c} = c - {U_0}^{\phi}$ and $\tilde{U} = U^{\phi} - {U_0}^{\phi}$, with c as the ground based phase speed and U^{ϕ} as the velocity in the direction of the azimuth ϕ , with the subscript 0 referring to the launch level.

The launch level is 450 hPa and the total launch momentum flux F_0 (corresponding to the integrated C in the phase speed direction) is formulated as a function of the latitude y [°]:

$$F_0 = \rho_0 F_{\text{launch,bs}}\ \eta \tag{3.2.106}$$

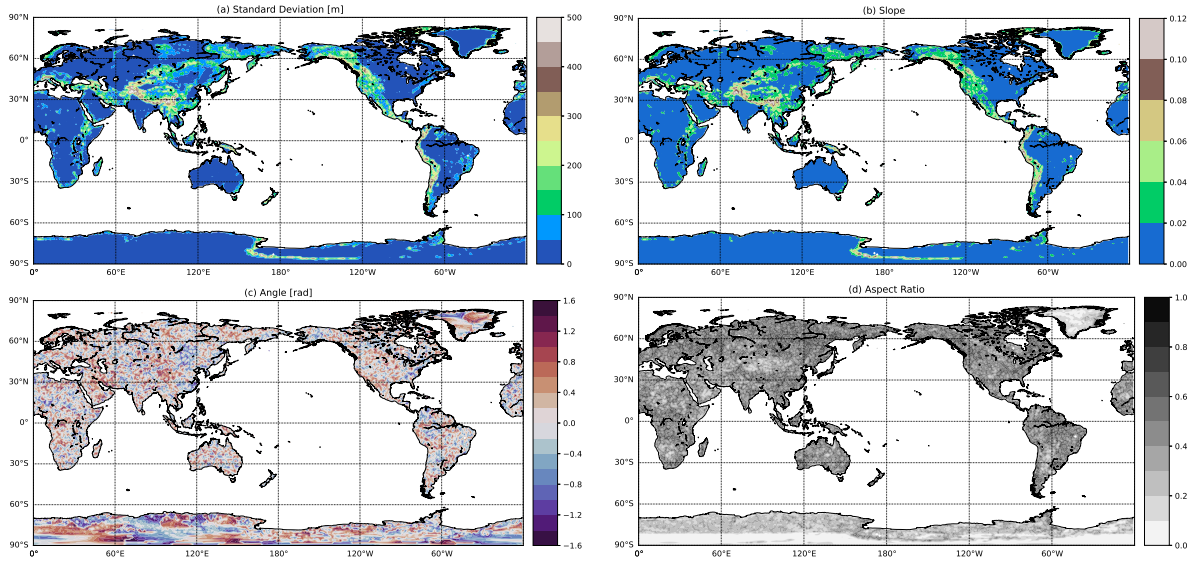


Figure 3.2.3: The four parameters used in blocked flow drag and gravity wave drag parametrizations at TQ959, (a) standard deviation [m], (b) orographic slope, (c) angle [rad] and (d) aspect ratio. Note that only land grid points are plotted.

$$\eta = 1 - \left(1 - R_{\rm eq}\right)\exp\left(-\left(\frac{y}{2w_{\rm eq}}\right)^2\right) - \left(1 - R_{\rm pl}\right)\exp\left(-\left(\frac{y - 90{\rm sgn}(y)}{2w_{\rm pl}}\right)^2\right) \tag{3.2.107}$$

$$R_{\rm eq} = \frac{\rho_0 F_{\rm launch,eq}}{\rho_0 F_{\rm launch,bs}},\ R_{\rm pl} = \frac{\rho_0 F_{\rm launch,pl}}{\rho_0 F_{\rm launch,bs}} \tag{3.2.108}$$

where $\rho_0 F_{\rm launch,bs} = 3.2 \times 10^{-3}$ Pa, $\rho_0 F_{\rm launch,eq} = 1.8 \times 10^{-3}$ Pa, $\rho_0 F_{\rm launch,pl} = 0.6 \times 10^{-3}$ Pa, $w_{\rm eq} = 10^\circ$ and $w_{\rm pl} = 7.5^\circ$.

The momentum fluxes are discretized in four equally spaced azimuths (north, south, west and east) and 50 phase speed bins in the range of 0.25 to 2000 m/s.

As momentum deposition processes, critical level filtering and nonlinear saturation are considered. In critical level filtering calculation, when $\tilde{c} - \tilde{U} < 0$ (assuming that waves encounter the critical level), the momentum flux is deposited to the mean flow in this layer. In nonlinear saturation calculation, when the upward propagating momentum flux $\rho\tilde{F}$ exceeds the saturated momentum flux $\rho\tilde{F}^{sat}$, the excess momentum flux ($\rho\tilde{F} - \rho\tilde{F}^{sat}$) is deposited to the mean flow. The saturated momentum flux is expressed as follows:

$$\rho\tilde{F}^{sat} = \rho C^* A \frac{\tilde{c} - \tilde{U}}{N}\left(\frac{\tilde{c} - \tilde{U}}{\tilde{c}}\right)^{2-p} \tag{3.2.109}$$

where C^*(=10) is the tuning parameter introduced by McLandress and Scinocca (2005). These vertical momentum flux depositions are calculated at each level for each azimuthal direction and phase speed. To reduce the computational cost, parameterization is performed only once an hour.

3.2.10 Land Surface Processes

The land surface model in the GSM employs a two-layer energy balance scheme based on the Simple Biosphere scheme (SiB; Sellers *et al.* 1986; Sato *et al.* 1989). It has evolved to a complex representation for snow and soil, with reference to Oleson *et al.* (2010)

The model is composed of vegetation canopy, snow and soil components, each of which has its own variables of temperature, water and ice content. Canopy air space is used to express paths of heat and water transfer. Figure 3.2.4 shows heat and water flows in an analogy of an electric circuit.

3.2.10.1 Fluxes

The main role of the land surface model is to provide lower boundary conditions of fluxes to the atmospheric model. The zonal and meridional momentum fluxes $(\tau_x, \tau_y) \equiv \overline{(w'\boldsymbol{v}')}_s$, sensible heat flux $H \equiv \overline{(w'\theta')}_s$, and water vapor flux $E \equiv \overline{(w'q')}_s$ are based on bulk formulae, as detailed in Subsection 3.2.6. Among these, τ_x and τ_y can be computed simply using the atmospheric wind velocity $\boldsymbol{v}_1$, whereas determination of H and E requires the surface temperature and specific humidity values provided by the land surface model. This model in the GSM defines the values as "canopy air space temperature T_a (= θ_s) and specific humidity q_a", where sensible heat and water vapor fluxes from vegetation canopy and ground surface are connected.
Sensible heat H is balanced by the sum of sensible heats from the canopy H_c and the ground surface H_g, since canopy air space is assumed to have negligible heat and water vapor capacities,

$$H = H_c + H_g \tag{3.2.110}$$

Similarly, water vapor E is balanced by the sum of evaporations from the canopy E_c^e, the ground surface E_g^e and bare soil E_{bs}, as well as transpiration from the canopy E_c^t and the ground surface E_g^t

$$E = E_c^e + E_c^t + E_g^e + E_g^t + E_{bs} \tag{3.2.111}$$

These fluxes are determined with canopy temperature T_c and ground surface temperature T_g. Here, T_g represents grass or bare soil temperature, but is regarded as snow surface temperature when the ground surface is covered with snow.

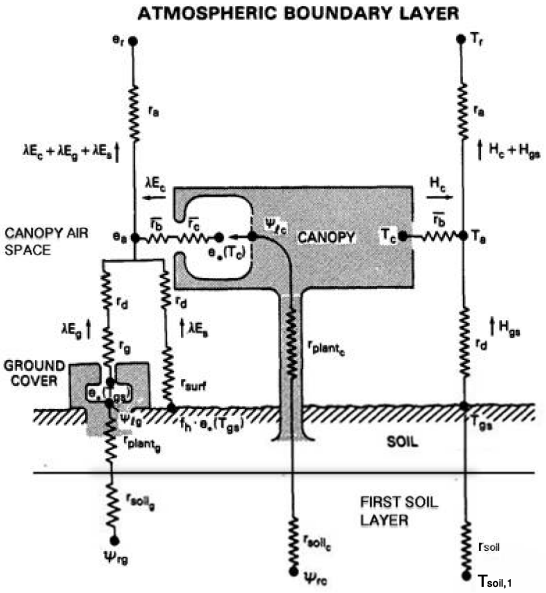


Figure 3.2.4: Schematic illustration of the land surface model (modified from Sellers *et al.* (1986)). Details are provided in the original paper.

3.2.10.2 Radiation and Albedo

The net radiation fluxes at the canopy R_c^n and the ground surface R_g^n are estimated based on the radiation balance equations

$$R_c^n = (1 - \alpha_c)\, S_{atm}^{\downarrow} + L_c^{\downarrow} \tag{3.2.112}$$

$$R_g^n = (1 - \alpha_g)\, S_{atm}^{\downarrow} + L_g^{\downarrow} \tag{3.2.113}$$

where α is the albedo, and $S_{atm}^{\downarrow}$ and $L^{\downarrow}$ are downward shortwave and longwave radiation from the atmosphere, respectively. In (3.2.112) and (3.2.113), R_c^n and R_g^n can be estimated using the albedos.

The surface albedo α_s can be determined as an average of the canopy albedo α_c and the ground albedo α_g, weighted by the fraction of canopy cover f_c

$$\alpha_s = f_c\, \alpha_c + (1 - f_c)\, \alpha_g \tag{3.2.114}$$

Similarly, the ground albedo α_g is an average of the grass albedo α_{grs} and the bare soil albedo α_{bs}, weighted by the fraction of grass cover f_{grs} in snow-free areas. When the ground is covered with snow, f_{grs} is set to zero, and α_{bs} is replaced by the value for snow α_{sn}

$$\alpha_g = \begin{cases} f_{grs}\, \alpha_{grs} + (1 - f_{grs})\, \alpha_{bs} & \text{(snow-free)} \\ \alpha_{sn} & \text{(snow-covered)} \end{cases} \tag{3.2.115}$$

α_c and α_{grs} are calculated with radiative transfer equations (Sellers *et al.* 1986) for leaf and stem area. α_{bs} is provided from climatological data of the MODIS albedo product (Schaaf *et al.* 2002) and modified using the solar zenith angle and soil moisture near the soil surface. α_{sn} evolves with time, and is corrected using the solar zenith angle.

3.2.10.3 Energy and Water Balances

The prognostic equations for T_c and T_g are given as

$$C_c \frac{\partial T_c}{\partial t} = R_c^n - H_c - L_{vap} E_c \tag{3.2.116}$$

$$C_g \frac{\partial T_g}{\partial t} = R_g^n - H_g - L_{vap} E_g - G_g \tag{3.2.117}$$

where the subscripts c and g denote canopy and ground surface, respectively, C is heat capacity, R^n net radiation, L_{vap} latent heat of vaporization, G_g ground surface heat flux, and $E_c \equiv E_c^e + E_c^t$, $E_g \equiv E_g^e + E_g^t + E_{bs}$.

Water storage on canopy leaves M_c and grass leaves M_g, which are sources of evaporation, are predicted by

$$\frac{\partial M_c}{\partial t} = P_{i,c} - P_{d,c} - E_c^e \tag{3.2.118}$$

$$\frac{\partial M_g}{\partial t} = P_{i,g} - P_{d,g} - E_g^e \tag{3.2.119}$$

where P_i is precipitation intercepted by leaves, and P_d water drip falling from leaves. The difference between these two values, $I_{cept} = P_{i,c} + P_{i,g} - (P_{d,c} + P_{d,g})$, represents interception by canopy and grass. When T_c (T_g) is below the freezing point of water, M_c (M_g) represents ice on canopy or grass leaves.

3.2.10.4 Snow

Snow temperature T_{sn} is predicted based on the principle of energy conservation and Fourier's law of heat conduction

$$C_{sn} \frac{\partial T_{sn}}{\partial t} = \frac{\partial G_{sn}}{\partial z} \tag{3.2.120}$$

$$G_{sn} = -\lambda_{sn} \frac{\partial T_{sn}}{\partial z} \tag{3.2.121}$$

where the subscript sn denotes snow, G is heat flux (positive downward), z snow depth from the snow surface, and λ thermal conductivity. Using a multi-layer model, a snow column is discretized into up to four layers (Figure 3.2.5). The top boundary condition the ground surface heat flux, while the bottom boundary condition is conductive heat flux with the first soil layer

$$G_{sn,0} = G_g \tag{3.2.122}$$

$$G_{sn,k_{max}} = \Lambda_{sn,k_{max}} (T_{sn,k_{max}} - T_{sl,1}) \tag{3.2.123}$$

where the subscript sl denotes soil, and Λ is a thermal conduction coefficient. For integration in time t, implicit methods are adopted and tri-diagonal matrices are solved.

Snow mass M_{sn} is predicted based on the snow mass balance equation

$$\frac{\partial M_{sn}}{\partial t} = S_{fall} + \left(S_{frst} - S_{sub}\right) + \left(S_{frz} - S_{melt}\right) \tag{3.2.124}$$

where S_{fall} is snowfall reaching the snow surface, including ice drip falling from leaves, S_{frst} frost, S_{frz} freezing, S_{sub} sublimation, and S_{melt} snowmelt.

Liquid water content in snow W_{sn} is predicted using

$$\frac{\partial W_{sn}}{\partial t} = \left(Q_{sn,infl} - Q_{sn,drng}\right) + \left(S_{dew} - S_{evap}\right) - \left(S_{frz} - S_{melt}\right) \tag{3.2.125}$$

where $Q_{sn,infl}$ is infiltration into snow including rainfall and water drip falling from leaves, $Q_{sn,drng}$ gravitational drainage from the bottom, and S_{dew} liquid dew, S_{evap} evaporation. Snow mass and snow water content are predicted using the same k_{max} layers as those for the snow temperature. When snow water content in a layer exceeds the layer's holding capacity, excess water is moved to the underlying layer. Snow depth data from Snow Depth Analysis (see Section 2.8) are used to set the initial value of snow water equivalent.

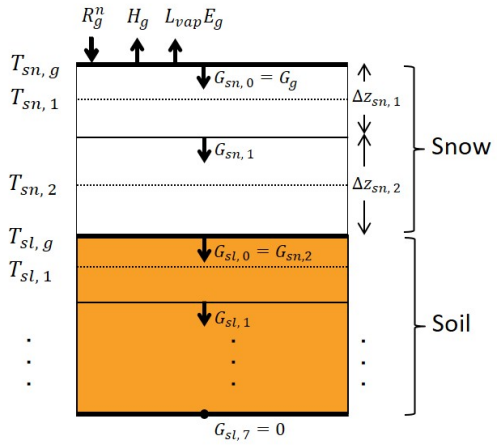


Figure 3.2.5: Schematic diagram of numerical discretization used to solve for snow temperatures $T_{sn,k}$ with two snow layers on top of soil. The subscripts *sn*, *sl* and k denote snow, soil and the k-th layer, respectively.

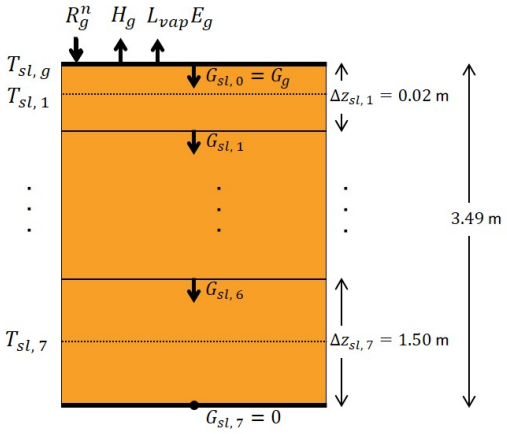


Figure 3.2.6: Schematic diagram of numerical discretization used to solve for soil temperatures $T_{sl,k}$ in snow-free areas.

3.2.10.5 Soil

Soil temperature T_{sl} is predicted in the same way as snow temperature, and soil is discretized into seven layers (Figure 3.2.6) of soil thicknesses $\Delta z_{sl,1-7}$ (m) = (0.02, 0.05, 0.12, 0.3, 0.5, 1.0, 1.5). The boundary conditions are

$$G_{sl,0} = \begin{cases} G_g & \text{(snow-free)} \\ \Lambda_{sl,k_{max}} \left(T_{sn,k_{max}} - T_{sl,1}\right) & \text{(snow-covered)} \end{cases} \tag{3.2.126}$$

$$G_{sl,7} = 0 \tag{3.2.127}$$

Soil moisture is predicted using Richard's equation, and root extraction for transpiration is also considered. The prognostic equation for the degree of saturation W is as follows:

$$\frac{\partial W}{\partial t} = \frac{1}{\rho_{wtr}\,\theta_{sat}} \left(-\frac{\partial Q}{\partial z} - S^t\right) \tag{3.2.128}$$

where ρ_{wtr} is water density, θ_{sat} soil porosity, Q water flux caused by differences in matric potential and gravitational potential, and S^t root extraction for transpiration. The top water flux is $Q_{infl} - E_{bs}$, where Q_{infl} is infiltration into soil including precipitation, drip falling from leaves and snowmelt, and the bottom is gravitational drainage Q_{drng}.

Surface runoff R_{off} and gravitational drainage Q_{drng} are counted as total runoff R_{total}

$$R_{total} = R_{off} + Q_{drng} \tag{3.2.129}$$

$$= T_{fall} - \left(Q_{infl} - Q_{drng}\right) - \left(Q_{sn,infl} - Q_{sn,drng}\right) \tag{3.2.130}$$

where

$$R_{off} = T_{fall} - Q_{infl} - \left(Q_{sn,infl} - Q_{sn,drng}\right) \tag{3.2.131}$$

$$T_{fall} = P_g - I_{cept} \tag{3.2.132}$$

P_g is gross rainfall, and T_{fall} throughfall reaching the ground surface. Q_{infl} is limited due to the maximum surface infiltration capacity.

3.2.10.6 Datasets

The climatological data of base soil albedo is derived from the MODIS albedo product[2] of NASA (Schaaf *et al.* 2002).

Each grid point on land is classified by a specific vegetation type provided from GLC2000[3] (Global Land Cover 2000; Bartholomé and Belward 2005) of the European Commission's Joint Research Center (JRC). The LAI (Leaf Area Index) is based on the MODIS LAI product[4] of NASA (Myneni *et al.* 2002). Fractions of canopy and grass covers are calculated using the 1 km MODIS-based Maximum Green Vegetation Fraction of USGS (Broxton *et al.* 2014a), the 1 km Tree Cover Continuous Fields product of GLCF (DeFries *et al.* 2000), and the Cropland and Pasture Area fraction[5] of EarthStat (Ramankutty *et al.* 2008).

The initial condition for soil moisture below the fourth soil layer is given by climatological data calculated using an offline model with the atmospheric forcing dataset of GSWP3 (Global Soil Wetness Project Phase 3; Kim 2017)[6]. The soil property is from HWSD (Harmonized World Soil Database; FAO *et al.* 2012).

3.2.11 Parallelization

In the GSM, Open Multiprocessing (OpenMP) is employed for shared memory parallelization, and the Message Passing Interface (MPI) is used for distributed memory parallelization. A two-dimensional decomposition method is adopted for parallelization among processes.

Figure 3.2.7 shows the schematic design of parallelization. There are five computational stages, and appropriate decompositions are selected in each stage. The base is the Grid stage. The Fourier and Legendre stages are used for calculating spherical harmonic transformation, and the Wavenumber stage is used for calculating Helmholtz equations in the semi-implicit scheme. The Horizontal Advection stage is used in the implementation of the semi-Lagrangian advection scheme.

At the Grid stage, since all vertical levels exist in a same rank for the computation of physical processes and non-linear terms of dynamical processes, variable arrays are decomposed into east-west and north-south directions. North-south decomposition follows a cyclic order, and is applied in such a way that the order of ranking is reversed alternately. This helps to mitigate load imbalances associated with physical parameterization and the number of grid points, since their computational loads depend mainly on latitudinal zones. At the Fourier stage, since all east-west grid points exist in a same rank for the performance of Fourier transformation, variable arrays are decomposed into north-south and vertical directions. At the Legendre stage, since all north-south grid points exist in a same rank for the performance of Legendre transformation, variable arrays are decomposed into vertical and longitudinal wavenumber directions. At the Wavenumber stage, since all vertical levels exist in a same rank for solving Helmholtz equations in the semi-implicit scheme, variable arrays are decomposed into longitudinal and total wavenumber directions. Communication among these four stages can be performed independently within each subset based on the provision of two restrictions for the number of decompositions: 1) the number of decompositions for the east-west direction, the vertical direction and the total wavenumber direction must be the same, and 2) the number of decompositions for the north-south direction and the longitudinal wavenumber direction must be the same.

At the Horizontal Advection stage, variable arrays are decomposed into vertical and north-south directions. To reduce the amount of communication relating to halo regions, the number of decompositions for the north-south direction is made as small as possible. Unlike communication in the stages described above, global communication is required for interaction between the Grid stage and the Horizontal Advection stage.

[2] https://search.earthdata.nasa.gov/search
[3] http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php
[4] https://search.earthdata.nasa.gov/search
[5] http://www.earthstat.org/cropland-pasture-area-2000/
[6] http://hydro.iis.u-tokyo.ac.jp/GSWP3/index.html, https://www.isimip.org/gettingstarted/details/4/

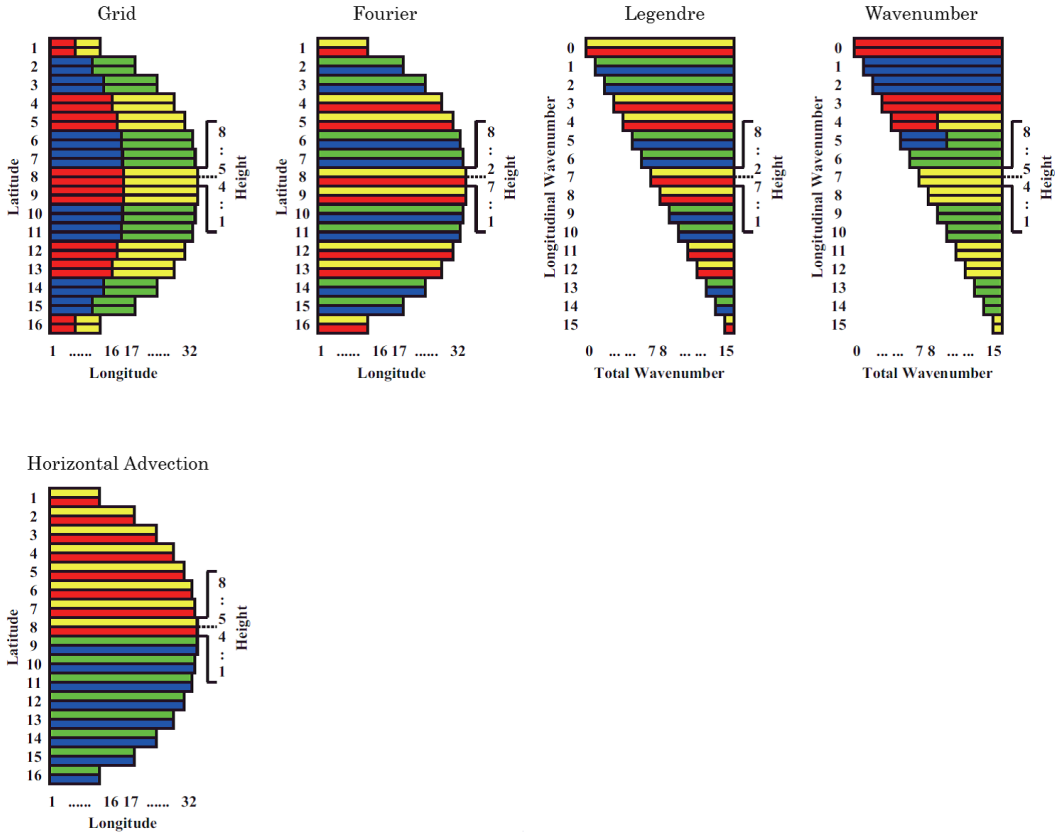


Figure 3.2.7: Schematic design of the parallelization. The number of processes used is assumed to be 4 in this example. Colors in the figure represent the rank for the computation in that area; red is rank 0, yellow is rank 1, blue is rank 2 and green is rank 3.

3.2.12 Surface Specifications

3.2.12.1 Orography

Model orography in the GSM is based on combined Global 30 Arcsecond lat-lon grid elevation orographic data generated using an average operation from the MERIT DEM (Yamazaki and D. Ikeshima 2017) dataset covering the 90° N – 60° S region and the RAMP2 (Liu *et al.* 2015) dataset covering 60° S – 90° S (Kanehama *et al.* 2023c). These combined elevation data are interpolated onto the reduced Gaussian grid after smoothing by multiplying the spectral coefficients by

$$f(n) = \left\{ \left(1 + df \left(\frac{n(n+1)}{N(N+1)} \right)^2 \right) \left(1 + 4.0 \left(\frac{n(n+1)}{N(N+1)} \right)^8 \right) \right\}^{-1}, \tag{3.2.133}$$

where n is the total wavenumber, N is the truncation total wavenumber, and $df = 0.866033186407302$ is a tunable smoothing parameter.

3.2.12.2 Grid Type

Land-ocean distribution is determined in reference to the Global Land Cover Characteristics (GLCC) database (Loveland *et al.* 2000) compiled by USGS and others. Model grid sections in which the land area ratio is more than 49% in GLCC are regarded as land grids in the GSM to keep the same global ocean area ratio as

GLCC. Grids not defined as land are sea (ocean) types and can have two tiles fractions (open water and ice). Inland water grids are treated as sea. Meanwhile, each land grid has a particular vegetation type based on the Global Land Cover 2000 database provided by the European Commission's Joint Research Centre in 2003. See Subsection 3.2.10 for the surface properties of land grids.

3.2.12.3 Sea and Lake Surfaces

On sea grids in the GSM, sea surface temperature (SST) and sea ice concentration (SIC) are given as boundary conditions minimally affected by the atmosphere. The amount of change in these variables during the time-integration of the model is equivalent to the time interpolated variation in daily climatological data.

On lake grids, monthly climatology based on the MODIS/Aqua product (Wan 2014) is used for water surface temperature. Lake ice concentration (LIC) is diagnosed from the surface temperature (T) using

$$\mathrm{LIC} = \begin{cases} 0 & T \geq 273.15 \\ 1 - \exp(0.1(T - 273.15)) & T < 273.15 \end{cases} . \tag{3.2.134}$$

The direct beam albedo α_B of water surfaces is basded on Briegleb *et al.* (1986):

$$\alpha_B = \frac{0.026}{(\mu^{1.7} + 0.065)} + 0.15(\mu - 0.1)(\mu - 0.5)(\mu - 1.0) , \tag{3.2.135}$$

where μ is the cosine of the solar zenith angle. The diffused albedo α_D is constant (= 0.06).

3.2.12.4 Sea and Lake Ice

The ice fraction is modelled as a slab with open water underneath and a skin temperature for thermal contact with the lowest part of the atmosphere. Sea ice parameterization involves (i) a fixed slab depth (with a constant volume regardless of melting), (ii) homogeneous and constant slab material properties, and (iii) climatological snow accumulation on the ice. Ice heat transfer is assumed to obey the Fourier law of diffusion:

$$(\rho C)\frac{\partial T_{ice}}{\partial t} = \frac{\partial}{\partial z}\left(\lambda \frac{\partial T_{ice}}{\partial z}\right) , \tag{3.2.136}$$

where $\rho C = 1.93 \times 10^6$ [$\mathrm{Jm^{-3}K^{-1}}$] is volumetric ice heat capacity, T_{ice} is ice temperature, and $\lambda = 2.03$ [$\mathrm{Wm^{-1}K^{-1}}$] is ice thermal conductivity. As a sea ice boundary condition, the temperature at the bottom of the slab is given as $T_{ice} = 271.51$K. The temperature at the top of the slab is diagnosed from the net heat flux at the top skin. The lake ice boundary condition at the bottom is the temperature of frozen water ($T_{ice} = 273.15$K), and the top boundary condition is the same as for sea ice. In the GSM, the ice slab is vertically discretized into four layers, with the temperature at each level solved via implicit time-integration.

The effect of snow over ice on the surface albedo is climatologically parameterized based on Hunke and Lipscomb (2006). The albedos of ice (α_{ice}) and snow (α_{snw}) depend on the surface temperature T_{skin}:

$$\alpha_{ice} = \begin{cases} \alpha_{ice0} & T_{skin} < 272.15 \\ \alpha_{ice0} - 0.075(T - 272.15) & T_{skin} \geq 272.15 \end{cases} , \\ \alpha_{snw} = \begin{cases} \alpha_{snw0} & T_{skin} < 272.15 \\ \alpha_{snw0} - k(T_{skin} - 272.15) & T_{skin} \geq 272.15 \end{cases} , \tag{3.2.137}$$

where T_{skin} is the surface temperature, α_{ice0} for visible is 0.78 and for near infrared is 0.36, α_{snw0} and k are 0.98 and 0.1 for visible, and 0.7 and 0.15 for near infrared. The total albedo is the area-weighted average of ice and snow albedos.

3.2.13 Initial Conditions

Initial conditions of subsystems such as the atmosphere and land are required for GSM time integration. The specifications of these are detailed in Table 3.2.3, where the term "forecast guess" represents the use of forecast variables with a lead time of six hours. However, land and snow variables are adjusted via snow analysis for consistency.

Table 3.2.3: Initial Conditions of GSM

Subsystem	Variable	Origin
Atmosphere	Zonal wind	4D-Var global objective analysis (see Section 2.5)
	Meridional wind	
	Temperature	
	Specific humidity	
	Surface pressure	
	Cloud water content	Forecast guess
	Diagnosed cloud water content in convection updraft	
	Cloud cover of stratiformis and convective cloud	
	Convective mass flux at cloud base	
Land surface	Grass (or bare soil) temperature	Forecast guess
	Canopy temperature	
	Liquid and ice water content on vegetation	
Land soil	Temperature	
	Liquid and ice water content in voids	Soil moisture analysis (see Section 2.9) and Climatological values in the deep layers
Snow	Snow water equivalent	Snow analysis (see Subsection 2.8.1)
	Temperature	Forecast guess
	Density	
	Liquid water content in voids	
	Albedo	
	Age	
Sea ice	Ice concentration	Sea ice analysis (see Section 5.6)
	Temperature	Forecast guess
Sea surface	Sea surface temperature	SST analysis (see Section 5.2)
	Friction velocity	Forecast guess
Lake ice	Temperature	Forecast guess

3.2.14 Forecast Performance

Figure 3.2.8 shows the root mean square error (RMSE) for 24-, 72- and 120-hour forecasts of 500 hPa geopotential height against analysis in the Northern Hemisphere extra-tropics. Dashed lines indicate monthly means, and solid lines represent 12-month running means. There are decreasing trends in RMSEs corresponding to GSM changes (see Subsection 3.2.1), although the impact of the changes in recent years has been relatively small.

Tropical cyclone (TC) track predictions are verified against the best track as analyzed by JMA's RSMC Tokyo - Typhoon Center. The mean position error of GSM TC track predictions in the western North Pacific (Figure 3.2.9) exhibits a gradual reduction from 1996 to 2022 due to GSM improvements, but considerable inter-annual variations are seen in 72-hour forecast errors and elsewhere.

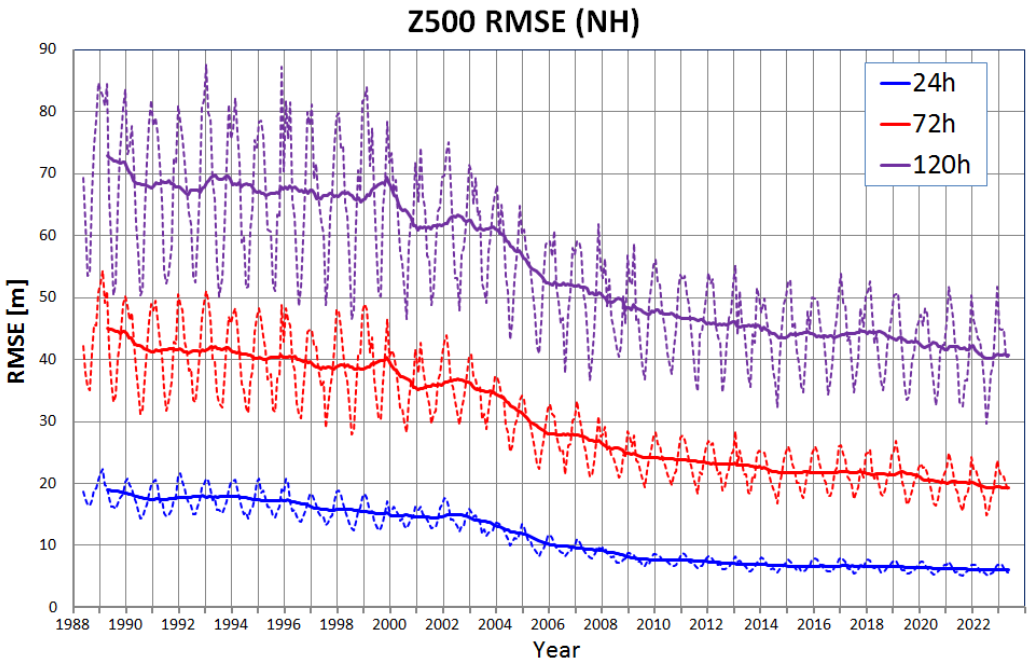


Figure 3.2.8: Root mean square error of GSM 500 hPa geopotential height (Z500) predictions against analysis in the Northern Hemisphere extra-tropics (20°N – 90°N). Dashed lines indicate monthly means, and solid lines represent running means calculated for the previous 12 months.

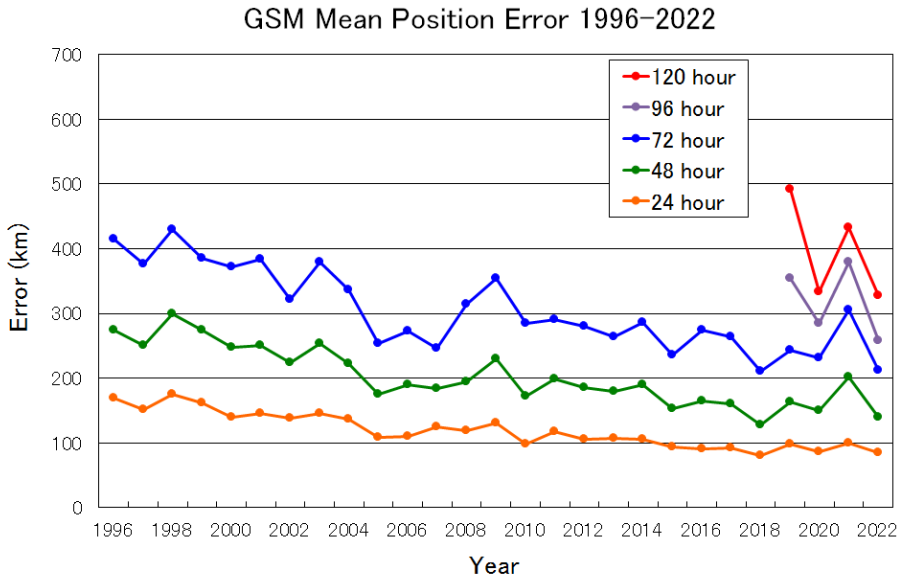


Figure 3.2.9: Mean position error of GSM TC track predictions in the western North Pacific from 1996 to 2022. The lines represent 24- (orange), 48- (green), 72- (blue), 96- (purple), and 120-hour forecasts (red).

3.3 Global Ensemble Prediction System

3.3.1 Introduction

The Global Ensemble Prediction System (Global EPS), which has been operational since January 2017, produces forecasts with lead times of up to 34 days to support the issuance of Five-day Tropical Cyclone (TC) Forecasts, One-week Forecasts, Two-week Temperature Forecasts, Early Warning Information on Extreme Weather, and One-month Forecasts. The system took over the roles of JMA's previous Typhoon EPS, One-week EPS and One-month EPS (JMA 2013, 2017; Yamaguchi *et al.* 2014; Hirai *et al.* 2014), thereby helping to reduce computational cost and focus resource development on a single EPS. The Typhoon EPS and One-week EPS were replaced by the Global EPS in January 2017, and the system inherited the roles of the One-month EPS in March 2017. In March 2019, initial perturbations of the Global EPS were upgraded to incorporate the usage of perturbations from the six-hour forecast ensemble based on the Local Ensemble Transform Kalman Filter (LETKF) of the previous analysis instead of analysis perturbations (Ota *et al.* 2019). In March 2020, a major upgrade in the Global EPS was implemented, which incorporated improvements in the forecast model, a two-tiered sea surface temperature (SST) approach and direct application of initial perturbations from JMA's hybrid data assimilation system (Yamaguchi *et al.* 2020). In March 2021, another major upgrade was implemented to incorporate more levels in the forecast model, increased ensemble size and improved initial perturbations (Yamaguchi *et al.* 2021). In March 2022, the effective horizontal resolution of the forecast model was enhanced to approximately 27 km for forecasts of up to 18 days, and to 40 km thereafter, with the introduction of a quadratic grid. SST boundary conditions were also upgraded (Yamaguchi *et al.* 2022; Sekiguchi *et al.* 2022). The forecast model was further upgraded in March 2023, with a two-tiered SST approach expanded to the whole globe (Ota *et al.* 2023; Yamaguchi *et al.* 2023).

3.3.2 System

3.3.2.1 Configuration, and Initial and Boundary Conditions

The specifications of the Global EPS are shown in Table 3.3.1.

A low-resolution version of JMA's Global Spectral Model (GSM; Section 3.2) is used in the EPS. Accordingly, the dynamical framework and physical processes involved are essentially identical to those of the GSM except for a lower horizontal resolution, which is further reduced after a lead time of 18 days. Unperturbed initial conditions are based on interpolation of the global analysis field (Section 2.5). A two-tiered SST approach (Takakura and Komori 2020) combining anomaly-fixed SSTs (Subsection 3.2.12) with SST prediction from the Seasonal EPS (Section 3.4) is used for the lower-boundary condition. Sea ice concentration (SIC) for the lower-boundary condition is the climatological value with the daily analysis anomaly. For the first 14 forecast days, SIC anomaly and sea ice extent anomaly are largely maintained. Thereafter, the SIC anomaly is adjusted to maintain the sea ice extent anomaly.

When a forecast is used for issuing Two-week Temperature Forecasts, Early Warning Information on Extreme Weather and One-month Forecasts, systematic biases estimated from hindcast experiments (Subsection 3.3.6.3) are removed.

3.3.2.2 Ensemble Size, Forecast Range and Frequency

The Global EPS, which consists of 50 perturbed members and a control, is run daily with initial times of 00, 06, 12 and 18 UTC. 00 UTC has a forecast range of 11 days, with data used for One-week Forecasts. The 12 UTC run on Tuesdays and Wednesdays is extended to 34 days and on other days to 18 days. After the 18-day lead time, the ensemble size is reduced to 25 per initial time. A 50-member ensemble is composed using lagged averaged forecasting (LAF; Hoffman and Kalnay 1983) from 25-member ensemble forecasts with two consecutive initial times. The 12 UTC runs are also used for daily Two-week Temperature Forecasts, and for Early Warning Information on Extreme Weather issued on Mondays and Thursdays when high probability in data for seven-day averages of very high/very low temperatures or heavy snow is expected for the week starting five to eight days ahead of issuance. A 50-member lagged ensemble with a one-month forecast range is used

Table 3.3.1: JMA Global EPS specifications

Start of operation (Latest major implementation)	January 2017 (March 2023)
Initial time	00, 06, 12 and 18 UTC
Forecast range	Initial time 00 UTC: 11 days Initial time 12 UTC: 34 days on Tuesdays and Wednesdays, 18 days otherwise Initial time 06 and 18 UTC: 132 hours
Ensemble size	Forecasts up to 18 days: 51 members Longer forecasts: 25 members (50-member lagged ensemble with 2 initial times)
Model type	GSM (an atmospheric general circulation model)
Horizontal resolution	Forecasts up to 18 days: Spectral triangular 479 (TQ479), reduced Gaussian grid system, roughly equivalent to $0.25^\circ \times 0.25^\circ$ (27 km) in latitude and longitude Longer forecasts: Spectral triangular 319 (TQ319), reduced Gaussian grid system, roughly equivalent to $0.375^\circ \times 0.375^\circ$ (40 km) in latitude and longitude
Vertical resolution (model top)	128 stretched sigma pressure hybrid levels (0.01 hPa)
Integration time step	Forecasts up to 18 days: 600 seconds Longer forecasts: 720 seconds
Initial perturbation generator	SV method, LETKF and LAF method
Initially perturbed area	Global
Model ensemble method	Stochastic physics scheme
Surface boundary conditions	SST: Two-tiered SST
Surface boundary perturbations	SST perturbations

for One-month Forecasts issued on Thursdays. Runs initialized at 06 and 18 UTC with a forecast range of 132 hours are used for Five-day Tropical Cyclone Forecasts in addition to runs with initial times at 00 and 12 UTC.

3.3.3 Initial Ensemble Perturbations

The initial perturbation of the Global EPS is generated by adding perturbations calculated using the singular vector (SV) method (Buizza and Palmer 1995) and the LETKF (Hunt *et al.* 2007) approach. The subsections below describe these methods and outline how atmospheric ensemble initial perturbations are generated.

3.3.3.1 Local Ensemble Transform Kalman Filter

Initial perturbations based on the LETKF are generated using 50 of the 100-member ensemble of hybrid 4D-Var global analysis, each of which is a six-hour forecast valid at the initial time from previous runs of the LETKF analysis ensemble (Section 2.5). These 50 are recentered to a zero ensemble mean and inflated using the same coefficient as hybrid 4D-Var global analysis. The initial perturbations from the LETKF are then multiplied by a factor of 0.85 to avoid overestimation upon addition to initial perturbations from the SV method.

3.3.3.2 Singular Vector Method

The SV method involves extraction of atmospheric growing modes as the SVs of the linearized low-resolution model. Table 3.3.2 summarizes the specifications of SV calculation for the Global EPS. The tangent-linear and adjoint models of JMA's 4D-Var system (Section 2.5) as of March 2021 generate SVs, although the resolution, fourth-order horizontal diffusion and cumulus convection parameterization of the models are modified from those of the originals. The models involve full dynamical core and physical processes including surface turbulent fluxes, vertical turbulent transports, gravity wave drag, long-wave radiation, cloud and large-scale precipitation, and cumulus convection. SVs based on tangent-linear and adjoint models incorporating full physical processes are called moist SVs, while those based on models incorporating simplified physical processes involving surface fluxes and vertical diffusion are called dry SVs.

Table 3.3.2: SV calculation specifications

Resolution	Spectral triangular truncation 63 (TL63), 40 levels		
Norm	Moist total energy		
Target area	Northern hemisphere (30 – 90°N)	Southern hemisphere (90 – 30°S)	Tropics (30°S – 30°N)
Physical process	Simplified physics		Full physics
Optimization time	48 hours		24 hours
Number of perturbations	25		

1. SV Definition

   SV calculations are conducted for the Northern Hemisphere (30 – 90°N), the tropics (30°S – 30°N) and the Southern Hemisphere (90 – 30°S). Dry SVs with 48-hour optimization are computed for the Northern Hemisphere and the Southern Hemisphere, while moist SVs with a 24-hour optimization time are computed for the tropics.

2. Norm of SV Calculation

   The norm for evaluating the growth rate of dry and moist SVs is based on a total energy norm that includes a specific humidity term (Ehrendorfer *et al.* 1999):

$$
\begin{aligned}
(x, Ey) \quad = \quad & \frac{1}{2}\int_0^1 \int_S \left[U_x U_y + V_x V_y + \frac{c_p}{T_r} T_x T_y \right. \\
& \left. + w_q \frac{L_c^2}{c_p T_r} q_x q_y \right] dS \left(\frac{\partial p}{\partial \eta} \right) d\eta + \frac{1}{2}\int_S \left[\frac{R_d T_r}{P_r} P_x P_y \right] dS . \qquad (3.3.1)
\end{aligned}
$$

   Here, U_x, V_x, T_x, q_x and P_x are the zonal wind, meridional wind, temperature, specific humidity and surface pressure components of the state vector x, respectively. The left-hand side of Eq. (3.3.1) is an inner product of state vectors x and y with a weighting of norm operator E. c_p is the specific heat of dry air at a constant pressure, L_c is the latent heat of condensation, and R_d is the gas constant for dry air. T_r = 300 K is a reference temperature, P_r = 800 hPa is a reference pressure, and w_q is a constant (here 0.04). $\int dS$ is the horizontal integration for the whole globe, and $\int \left(\frac{\partial p}{\partial \eta}\right) d\eta$ gives the vertical integration from the surface to the model top. Horizontal integration is performed over each target area instead of the whole globe at the end of the optimization time. The norm at the initial time is also vertically integrated with a weight dependent on the model level; all terms are neglected above the 31st model level, and the specific humidity term is neglected above the 9th model level. When the surface pressure is 1,000 hPa, the 31st and 9th model levels correspond to around 50 and 750 hPa, respectively. This suppresses initial perturbation with a large peak of energy in the upper stratosphere and confines initial specific humidity perturbation in the lower troposphere.

3. Generation of SV-based Perturbations

   SV-based perturbations are linear combinations of SVs. A total of 50 SVs are generated for the Northern Hemisphere and Southern Hemisphere, and 35 for the tropics. SVs with extremely high growth rates and those with large humidity perturbations over desert areas in low latitudes are discarded, as their modes do not grow appropriately in the Global EPS. The SVs are transformed in a variance minimum rotation (Yamaguchi *et al.* 2009) to generate 25 SV-based perturbations for each targeted area. The perturbations for both hemispheres are scaled so that related amplitudes of temperature at the 15th model level (or the 6th model level for the tropics) in the targeted area are 0.21 K (or 0.20 K for the tropics). When the surface pressure is 1,000 hPa, the 15th and 6th model levels correspond to around 500 and 850 hPa, respectively. The perturbations for the three targeted areas are linearly combined to create global perturbations.

   Initial conditions for 50 perturbed members are generated by adding or subtracting the 25 SV-based perturbations from the analysis.

3.3.4 Perturbed Physics

The stochastic physics scheme (Buizza *et al.* 1999b) is used for perturbed members to represent model uncertainties in physical parameterizations. This scheme represents random errors associated with physical processes as follows:

$$\frac{\partial \mathbf{x}}{\partial t} = F(\mathbf{x}) + \alpha(\lambda, \phi, t)P(\mathbf{x}). \tag{3.3.2}$$

Here, t, $\mathbf{x}$, $F(\mathbf{x})$ and $P(\mathbf{x})$ represent time, the set of forecast variables, the overall tendency of the forecast model and the tendency of parameterized physical processes at latitude λ and longitude ϕ, respectively. The random variable $\alpha(\lambda, \phi, t)$ is expressed via a triangularly truncated spherical harmonics expansion (Berner *et al.* 2009) with a truncation wave number of 20. It has a time correlation of six hours based on a first-order autoregressive process. The average of α is set as zero. Its value is limited to a specific range (-0.7 to 0.7) to avoid excessive perturbation, and its value in the stratosphere is also set as zero.

3.3.5 Sea Surface Temperature and Related Perturbations

The SST boundary condition of the Global EPS combines anomaly-fixed SSTs (Subsection 3.2.12) from JMA analysis (Subsection 5.2.1) and SST prediction from the Seasonal EPS (Section 3.4) using a two-tiered approach (Takakura and Komori 2020) in which an anomaly-fixed value covers the entire globe for up to 6 days and is linearly relaxed to the bias-corrected ensemble mean of the Seasonal EPS from 6 to 11 days.

SST perturbations (Hotta and Ota 2019) are then generated and added to the SST of perturbed members. The perturbation of the anomaly-fixed SST for the member i (ΔSST_i^f) in the forecast from the initial time T_0 is constructed as

$$\Delta SST_i^f \left(T_0; t_f\right) = \alpha \left[SSTA\left(T_i + \Delta T + t_f\right) - SSTA\left(T_i\right)\right] \tag{3.3.3}$$

where *SSTA* is the SST anomaly from the daily climatology, t_f is the forecast lead time and T_i is a past date randomly sampled from the range within ± 29 days from the calendar date of the initial time. ΔT and α are tunable parameters, and are set as 1 day and 1.0, respectively. Perturbation of SST prediction from the Seasonal EPS is the same as that of the anomaly-fixed SST, except α is set as 0.85 because the RMSE of SST prediction from the Seasonal EPS is smaller than that of the anomaly-fixed SST in the tropics and subtropics. SST perturbation is set as 0 where the maximum sea ice concentration on T_0, T_i and $T_i + \Delta T + t_f$ is above a certain threshold (0.001) to prevent the use of excessively large perturbations.

3.3.6 Performance

The performances of each EPS product are described below. For the sake of completeness, the period before Global EPS operation is also incorporated.

3.3.6.1 Typhoon Forecasting

Typhoon forecasting is supported by the Global EPS, as it was previously by the Typhoon EPS. The results of related verification are provided in the Annual Report on Activities of the RSMC Tokyo – Typhoon Center [7].

Ensemble TC tracks derived from the EPS enable JMA forecasters to integrate TC track forecast uncertainty into their operational processes. Strike probability data, which indicate the chances of a TC center passing within 120 km of a grid point, are routinely produced as a form of probabilistic guidance. Figure 3.3.1 shows the reliability of this probability data for the coming five days [8].

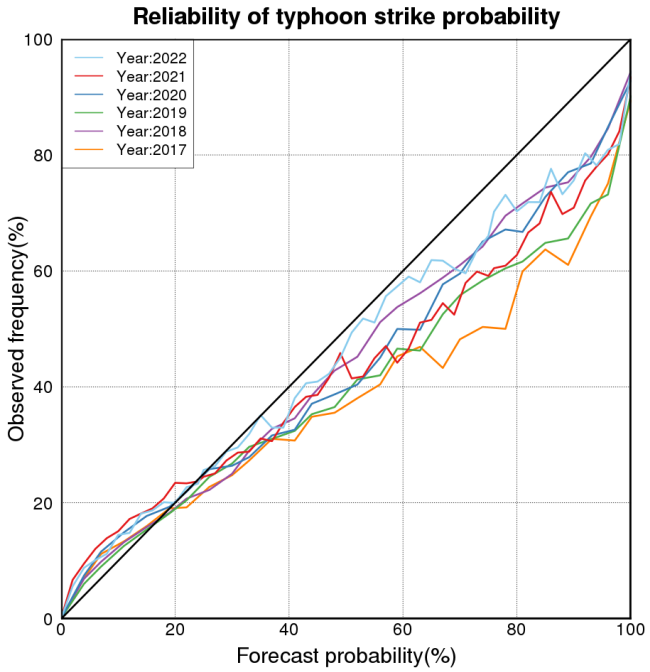


Figure 3.3.1: Reliability diagram for probabilistic typhoon-position forecasts as derived from the EPS for 2017 (yellow), 2018 (purple), 2019 (green), 2020 (blue), 2021 (red) and 2022 (light blue). RSMC Tropical Cyclone Best Track information was referenced as observation data.

3.3.6.2 One-week Forecasting

One-week forecasting is now supported by the Global EPS rather than the previous One-week EPS. The results of related verification are provided on the website of the WMO Lead Centre for EPS Verification[9].

Figure 3.3.2 shows a time-series representation of monthly-averaged RMSEs for the 500-hPa geopotential height ensemble mean forecast against analysis for the Northern Hemisphere (NH; 20 – 90°N). Figure 3.3.3 compares RMSEs of ensemble means, unperturbed members and the spread of the ensemble averaged for the periods of DJF (December/January/February) 2021/2022 and JJA (June/July/August) 2022. A higher level of skill is observed for ensemble means than for unperturbed members, especially for longer lead times. For shorter forecast lead times, the spread is almost the same size as that for the ensemble mean RMSE; however, as the forecast lead time increases, values tend to fall slightly. Figure 3.3.4 shows the Brier skill score (BSS) for 500-hPa geopotential height probabilistic forecasts in the NH. The reference forecast for the skill score is the climatological probability given by the frequency derived from analysis fields for each month. Since the start of operation, performance has improved annually in ensemble mean forecasts and probabilistic forecasts.

[7] https://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/annualreport.html

[8] The 2021 result was calculated using forecast data only for the period from April to December to maintain verification data consistency,

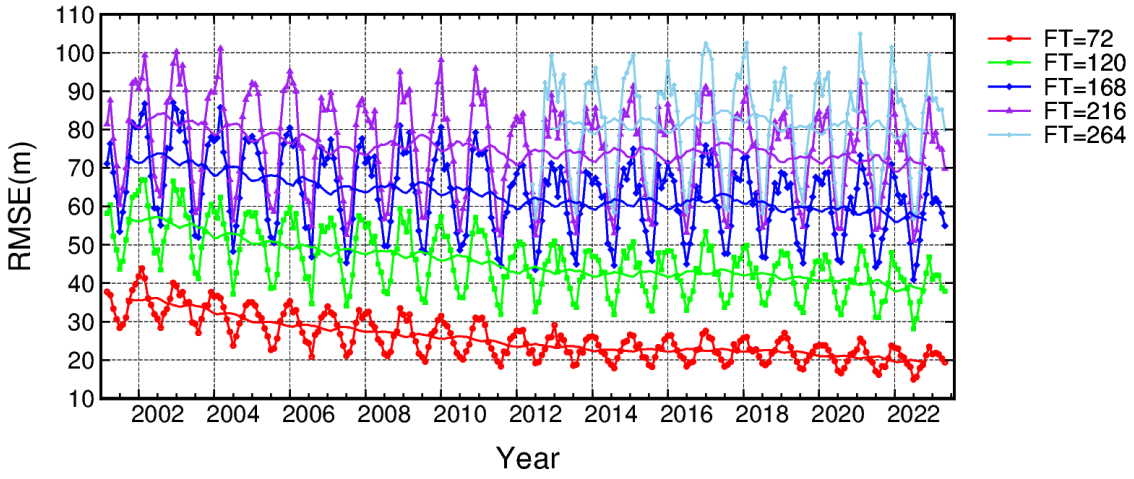


Figure 3.3.2: Time-series representation of ensemble mean scores for the EPS (where the score is the monthly average RMSE of the ensemble mean) for Northern Hemisphere (20 – 90°N) 500-hPa geopotential height forecasts with lead times of 72 (red), 120 (green), 168 (dark blue), 216 (violet) and 264 (light blue) hours from March 2001 to May 2023. 13-month running means are also shown.

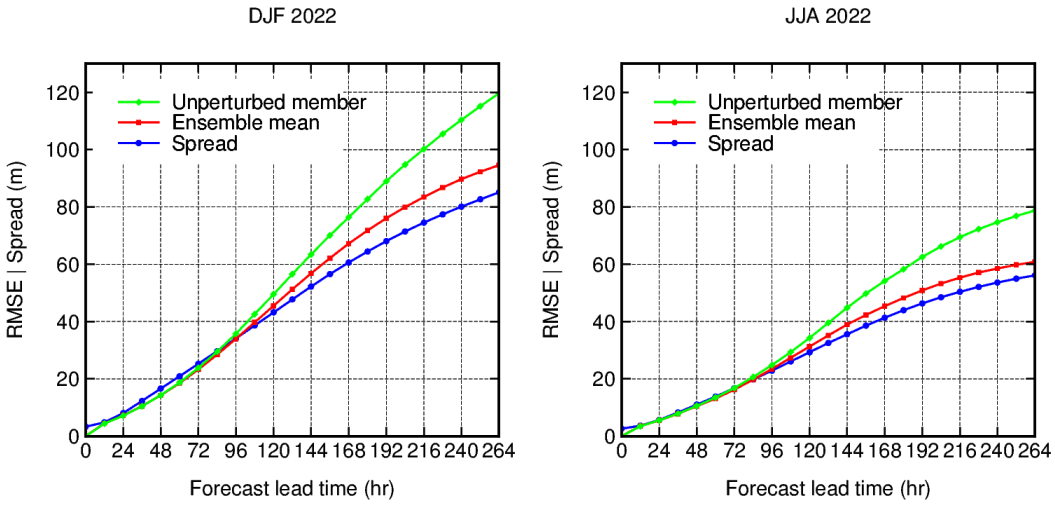


Figure 3.3.3: RMSEs for Northern Hemisphere (20 – 90°N) 500-hPa geopotential height forecasts for the ensemble mean (red) and the unperturbed member (green) for DJF 2021/2022 and JJA 2022 from the EPS. The spread of the ensemble (blue) is also shown.

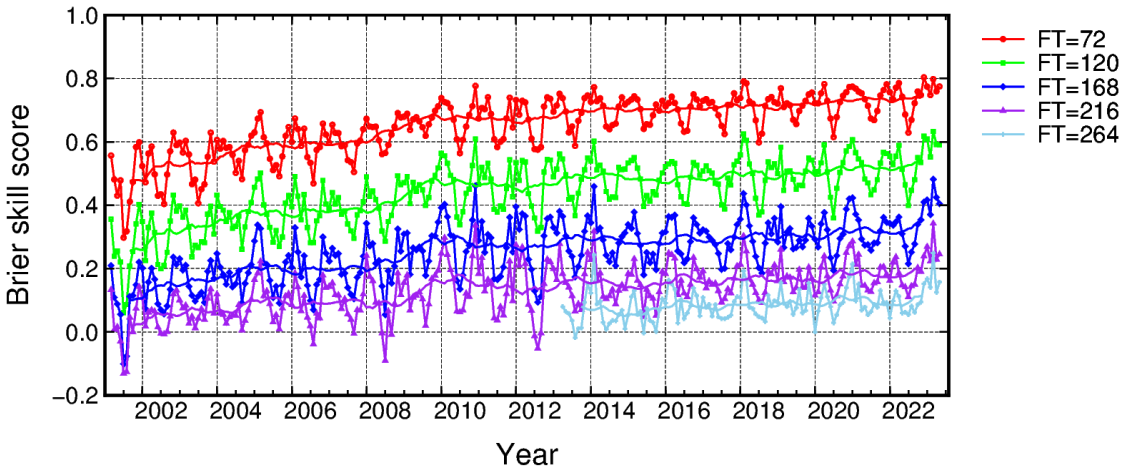


Figure 3.3.4: As per Figure 3.3.2, but for Brier skill score with probabilistic forecasts of 500-hPa geopotential height negative anomalies with magnitudes less than one climatological standard deviation from JMA's EPS.

3.3.6.3 One-month Forecasting

One-month forecasting is now supported by the Global EPS rather than the previous One-month EPS. The results of prediction skill evaluation based on hindcast experiments and real-time forecasts are provided on the Tokyo Climate Center website[10]. Hindcast experiments covering the period from 1991 to 2020 with 13 ensemble members were conducted using atmospheric initial conditions produced from JRA-3Q (Section 2.11). Initial perturbations were created from a combination of initial and evolved SVs. For computational efficiency, perturbations from LETKF application in real-time operation were not adopted.

The skill of ensemble mean forecasts was evaluated using the Anomaly Correlation Coefficient (ACC) and the RMSE for selected areas with respect to several physical variables. Probabilistic forecast skill was also evaluated based on BSS, Reliability Skill Score, Resolution Skill Score and Relative Operating Characteristics (ROCs).

Figure 3.3.5 shows a time-series representation of NH 500-hPa geopotential height ACC for ensemble mean forecasts averaged for days 2 – 29 (the running mean of 52 forecasts) based on operational forecasting conducted from 1997 to 2022. Skill exhibits a rising trend with fluctuations corresponding to El Niño Southern Oscillation (ENSO). Table 3.3.3 shows ROC areas of 2-m temperature (T2m) and 500-hPa geopotential height (Z500) anomalies based on hindcast experiments verified against JRA-3Q analysis, and indicates that skill for the tropics is higher than for the extratropics in all forecasts from each initial season.

as the ensemble size of the Global EPS was increased in March 2021.

[9] http://epsv.kishou.go.jp/EPSv/

[10] https://www.data.jma.go.jp/tcc/tcc/index.html

Table 3.3.3: ROC areas of 28-day (from day 3 to day 30) mean 2-m temperature (T2m) and 500-hPa geopotential height (Z500) anomaly prediction for positive anomaly events in the Northern Hemisphere (NH; 20 – 90°N), the tropics (20°S – 20°N), and the Southern Hemisphere (SH; 90 – 20°S) based on hindcast experiments covering the period from 1991 to 2020. The figures in the table are multiplied by 100.

T2m

Init	NH	Tropics	SH
DJF	81.9	86.2	80.6
MAM	79.3	84.3	78.7
JJA	79.6	83.5	78.1
SON	79.5	84.2	79.3

Z500

Init	NH	Tropics	SH
DJF	80.3	92.2	76.5
MAM	76.2	91.3	74.7
JJA	75.9	88.1	75.6
SON	74.9	89.0	77.8

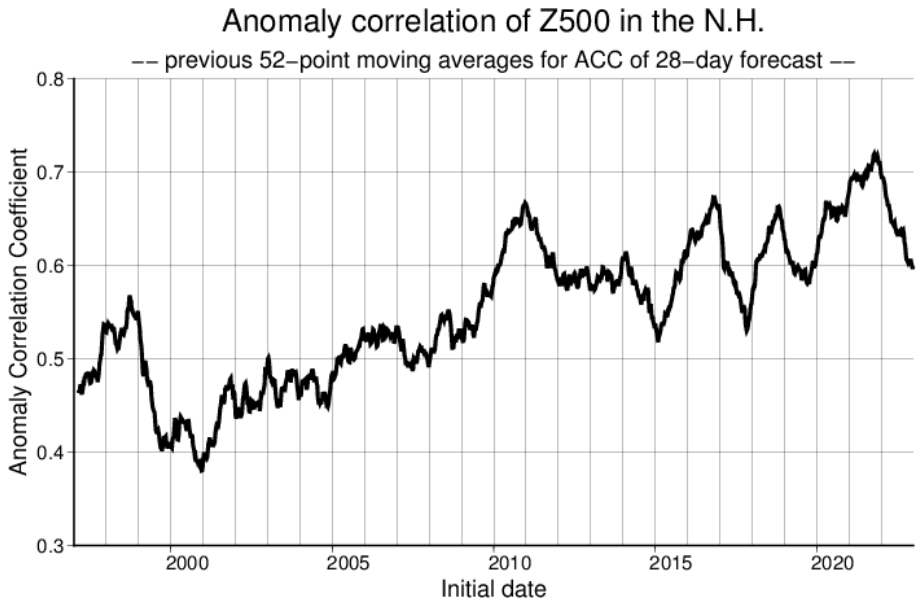


Figure 3.3.5: Time-series representation for the Northern Hemisphere (NH; 20 – 90°N) 500-hPa geopotential height anomaly correlation coefficient in ensemble mean forecasts averaged over 28 days from day 2 to day 29 (the running mean of 52 forecasts) based on operational forecasting conducted from 1997 to 2022.

3.4 Seasonal Ensemble Prediction System

3.4.1 Introduction

JMA operates the Seasonal EPS to support a wide range of seasonal forecast products, including three-month forecasts, warm- and cold-season forecasts, and El Niño outlooks. JMA began operating an ensemble prediction system with an atmosphere-ocean coupled model in July 1998 with and objective limited to El Niño outlooks. Since February 2010, seasonal forecasts have also been covered by upgraded systems (Takaya *et al.* 2017). The latest major update was introduced in February 2022. This section describes the details of the latest JMA/MRI-CPS3 (referred to simply as CPS3) (Hirahara *et al.* 2023).

3.4.2 System Configuration

3.4.2.1 Forecast Model

The CPS3 forecast model is an atmosphere/land/ocean/sea-ice coupled type (Table 3.4.1).

The atmosphere/land model is based on a low-resolution version of JMA's Global Spectral Model as of March 2020 (GSM2003; Yonehara *et al.* 2020; JMA 2019), with a horizontal resolution of TL319 (triangular truncation at total wavenumber 319 with a linear grid), which corresponds to approximately 55-km grid spacing, and 100 vertical levels with the model top at 0.01 hPa. In CPS3, several schemes were upgraded from GSM2003 to improve representation of atmosphere/ocean climatology and sub-seasonal to inter-annual variability. Refinements for the cumulus convection scheme included treating the dissipation time scale of kinetic energy separately for shallow and deep cumulus clouds. The cumulus entrainment rate is modified to take into account empirical dependence on altitude and humidity, rather than being constant in the vertical direction (Bechtold *et al.* 2008; Komori *et al.* 2020), and the lower limit of it is proportional to the depth of the planetary boundary layer (Tokioka *et al.* 1988). The stratocumulus scheme is triggered when the conditions proposed by Kawai *et al.* (2017) are met (Chiba and Kawai 2021). Gustiness near the sea surface is parameterized with downdrafts generated by deep convection (Redelsperger *et al.* 2000) and boundary layer free convection (Godfrey and Beljaars 1991). The skin sea surface temperature (SST) scheme, which solves the heat budget in the upper few meters of the ocean in consideration of diurnal changes in the vertical water temperature profile and the surface skin layer in radiative equilibrium, is introduced by expanding Zeng and Beljaars (2005). A fractional land ratio is introduced to consider sub-grid-scale water surfaces. The thermodynamic lake scheme is also adopted to predict lake ice formation and lake temperature variations through water phase changes and heat transfer among water, ice and snow. Calculation of solar zenith angles (Hogan and Hirahara 2016) and treatment of surface albedo (Hogan and Bozzo 2015) for shortwave radiation are also adopted. A monthly climatology of ozone concentration calculated from 1981 – 2010 MRI-CCM2 reanalysis is used (Subsection 3.9.1, Deushi and Shibata 2011). Greenhouse gas concentrations are based on observations conducted up to 2016 and on the Coupled Model Intercomparison Project phase 6 (CMIP6) emission scenario SSP2-RCP4.5 (Van Vuuren *et al.* 2011) thereafter. A three-dimensional monthly aerosol concentration climatology (Yabu *et al.* 2017) is used both for re-forecasts and operational forecasts, and CPS3 provides a mode allowing evaluation of the direct radiative effect from volcanic aerosols in the event of eruptions.

The ocean/sea-ice component is the Meteorological Research Institute Community Ocean Model version 4.6 (MRI.COM; Tsujino *et al.* 2017). For CPS3, the model is configured to cover the whole globe with the tripolar grid of Murray (1996) at a horizontal resolution of 0.25 x 0.25° . This resolution is sufficient to

Table 3.4.1: JMA Seasonal EPS specifications

Last update	February 2022
Initial Time	00 UTC, every day
Forecast Range	up to seven months
Ensemble Size	5 members per start date
Forecast Model	GSM coupled with the Meteorological Research Institute Community Ocean Model (MRI.COM)
Horizontal Resolution	GSM: Spectral triangular 319 (TL319) reduced Gaussian grid system, equivalent to approximately 55km MRI.COM: 0.25 x 0.25° in latitude and longitude
Vertical Resolution (model top)	GSM: 100 stretched sigma pressure hybrid levels (0.01 hPa) MRI.COM: 60 levels
Initial perturbation method	Atmosphere: BGM method and LAF method Ocean: Perturbations calculated using 4DVAR minimization history and LAF method
Model perturbation method	Stochastic physics scheme for the atmosphere

resolve Tropical Instability Waves, which promote heat exchange between the equator and its north and south, providing negative feedback to the Pacific equatorial SST during El Niño Southern Oscillation (ENSO) periods (Vialard *et al.* 2001; An 2008). A z* vertical coordinate (Adcroft and Campin 2004) is used to accurately capture flow along steep seafloor topography. The sea ice model deals with sea ice advection, formation, growth, and melting using five ice-thickness categories (Tsujino *et al.* 2017).

The atmosphere and ocean models are integrated at different time steps, with a simple coupler (Yoshimura and Yukimoto 2008) absorbing the difference and exchangeing sea surface temperature, sea ice cover and sea surface fluxes (radiation, latent and sensible heat, momentum and fresh water) between the models every hour.

Atmospheric conditions are initialized with the Japanese Reanalysis for Three Quarters of a Century (JRA-3Q-provisional, Kosaka *et al.* 2024) for re-forecast, and the JMA Global Analysis (GA) for operational forecast, in which atmospheric conditions are updated with a shorter time delay. CPS3 has a different land-sea mask from this analysis and a unique lake scheme that requires initialization. To fill the gaps between these differences while avoiding initial shock, offline surface simulation is separately run with JRA-3Q forcing and GA (for operational forecasts only). Oceanic and sea ice initial conditions are from the ocean data assimilation system (MOVE-G3; Fujii *et al.* 2023, Section 5.3).

3.4.2.2 Ensemble Size, Forecast Range and Frequency

To represent uncertainties in the initial condition, CPS3 adopts a combination of the Lagged Average Forecast (LAF; Hoffman and Kalnay 1983) method and the initial perturbation method described below. CPS3 runs a control member and four perturbed members from an initial time of 00 UTC every day up to seven months. Operational products for seasonal forecasting and El Niño outlooks at JMA are composed of 51 members (involving 3 members and 17 start dates). Atmospheric initial perturbations are generated via the Breeding of Growing Modes (BGM) method (Chikamoto *et al.* 2007). Oceanic initial perturbations are generated using 4DVAR minimization history, with which daily analysis error covariance can be approximated (Niwa and Fujii 2020). To represent physical processes in the model, a stochastic physics scheme (Yonehara and Ujiie 2011; Buizza *et al.* 1999b) is applied to perturbed ensemble members (Subsection 3.3.4).

3.4.2.3 BGM Method

Atmospheric perturbation breeding is performed separately for the Northern Hemisphere (NH; 20 – 90°N), the tropics (20°S – 20°N) and the Southern Hemisphere (SH; 20 – 90°S). First, perturbed and unperturbed initial conditions are integrated up to 24 hours. Then, the difference between perturbed and unperturbed fields is normalized so that the area-averaged root mean square difference for 500-hPa geopotential height over the NH and SH (200-hPa velocity potential for the tropics) is equal to 14.5 (20.0) % of the climatological variance. The normalized perturbations are subsequently orthogonalized to each other and added to the analysis to create the next set of initial perturbations. In CPS3, the NH, SH and tropical initial perturbations are combined and added to/subtracted from JRA-3Q (re-forecast) and GA (operational forecast).

3.4.3 Performance

The forecast skill of CPS3 is evaluated in line with the World Meteorological Organization (WMO) Global Data Processing and Forecasting System manual (WMO 2019). To verify performance, a 30 year hindcast for 1991 – 2020 was performed with forecast model identical to the operational application, except for an ensemble size of 10 rather than 51.

Figure 3.4.1 shows the Anomaly Correlation Coefficients (ACCs) for use between ensemble mean forecasts and Merged satellite and in-situ data Global Daily SSTs (MGDSSTs, Kurihara *et al.* 2006) in NINO.3 (5°S – 5°N, 150 – 90°W), NINO.WEST (0° – 15°N, 130 – 150°E) and IOBW (20°S – 20°N, 40 – 100°E). CPS3 effectively reproduces inter-annual modes of variability in tropical oceans, which are the primary known basis for seasonal predictability. The Relative Operating Characteristic (ROC) areas of 2-m temperature (T2m) anomalies and 500-hPa geopotential height (Z500) anomalies are shown in Table 3.4.2. Skill for the tropics is superior to that for the NH and SH.

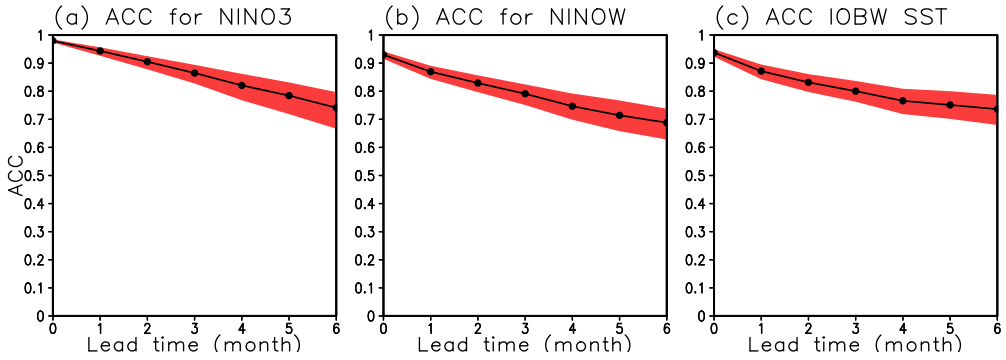


Figure 3.4.1: ACC for MGDSSTs in (a) NINO.3 (5°S – 5°N, 150 – 90°W), (b) NINO.WEST (0° – 15°N, 130 – 150°E) and (c) IOBW (20°S – 20°N, 40 – 100°E). Shading indicates a 95% confidence interval estimated using the bootstrap method (1,000 samples).

Table 3.4.2: ROC areas of three-month mean (JJA and DJF) T2m and Z500 for positive anomaly events (upper tercile) in the NH, the tropics and the SH. Statistics are based on hindcast experiments for 1991 – 2020. The numbers in the table are multiplied by 100. The initial months are May (10-member LAFs from April 11 and 26) for JJA and November (10-member LAFs from October 13 and 28) for DJF.

T2m	NH	Tropics	SH
JJA(Initial:May)	66.8	77.2	66.3
DJF(Initial:Nov)	67.3	80.9	67.4

Z500	NH	Tropics	SH
JJA(Initial:May)	68.1	90.2	64.9
DJF(Initial:Nov)	66.8	90.0	67.8

Other verification scores can be found on the Tokyo Climate Center website (https://www.data.jma.go.jp/tcc/tcc/products/model/hindcast/CPS3/index.html).

3.5 Meso-Scale Model (JMA-MSM2206)

3.5.1 Introduction

The meso-scale numerical prediction system has been operated since March 2001 to provide information for disaster prevention and aviation safety. The Meso-Scale Model (MSM) was initially a hydrostatic spectral model, producing 18-hour forecasts every 6 hours at 00, 06, 12 and 18 UTC. The domain covered Japan and its surrounding areas (3,600 km × 2,880 km) at a horizontal resolution of 10 km with 40 vertical layers.

In September 2004, the MSM was replaced with a non-hydrostatic grid model (JMA-NHM; Saito *et al.* 2006, 2007) with retention of similar general configurations in areas such as resolution, forecast time and forecast frequency. In March 2006, resolution and operation frequency were enhanced to produce 15-hour forecasts at 00, 03, 06, 09, 12, 15, 18 and 21 UTC with 5-km horizontal-grid spacing and 48 vertical layers. After subsequent model updates, the forecast period of the MSM was extended to 39 hours for all eight daily operations with an enlarged domain (4,080 km × 3,300 km) in 2013. The forecast period was extended to 51 hours at 00 and 12 UTC in March 2019, and subsequently to 78 hours at 00 and 12 UTC in June 2022. The extension of the forecast period and the model domain supported improved provision of useful information for disaster prevention and aviation operations with a two-day lead time.

In February 2017, the new-generation nonhydrostatic model ASUCA (Ishida *et al.* 2009, 2010, 2022) was incorporated into the operational MSM following its application for LFM usage in January 2015 (Aranami *et al.* 2015). The development of ASUCA was begun in 2007 after the development and widespread adoption of new nonhydrostatic equations allowing conservation of mass as well as sophisticated numerical methods

in computational fluid dynamics. Efficient operation of numerical models on scalar multi-core architecture was also required against a background of rapid expansion in the market for massive scalar computers in the supercomputer field (Hara *et al.* 2012).

ASUCA has great potential to meet these demands. In the model, flux-form fully compressible governing equations are adopted and discretized using the finite volume method to guarantee mass conservation. The three-stage Runge-Kutta scheme (Wicker and Skamarock 2002) is employed for time integration, leading to better computational stability, even with a longer time-step interval, than the JMA-NHM. Improvement of parallelization and coding methods yields more effective computation on massive scalar multi-core architecture.

Physical processes equivalent to or better than those of the JMA-NHM are implemented via the use of the Physics Library [11] , in which various subroutines related to physical processes are collected as vertical one-dimensional models with unified coding and interface rules (Hara *et al.* 2012; Hara 2015). This simple one-dimensional realization helps to improve computational efficiency, especially on scalar computers, and facilitates efficient development of physical processes such as evaluation of the straightforward responses of specific processes of interest via idealized single-column model experiments.

As described above, the MSM was significantly upgraded with the introduction of ASUCA in February 2017, updating of cloud microphysics, cloud fractions and land process in March 2020, and a layer-number was increase from 76 to 96 in March 2022. A one-dimensional ocean mixed-layer model predicting sea surface temperature (SST) was also introduced to represent SST cooling related to wind stress over ocean. Here Subsection 3.5.2 describes general configurations, Subsection 3.5.3 outlines the design of the dynamical core, and other subsections detail physical processes such as cloud physics, convective parameterization and radiation. Improvement of parallelization in ASUCA is described in Subsection 3.5.11, and forecast performance evaluation is detailed in Subsection 3.5.12.

3.5.2 General Configuration

The current ASUCA-based MSM is operated eight times a day, providing 78-hour forecasts at 00 and 12 UTC and 39-hour forecasts at 03, 06, 09, 15, 18 and 21 UTC. Its forecast domain is a rectangular flat area of 4,080 km × 3,300 km covering Japan and its surroundings, with a grid spacing of 5 km. The domain configuration is identical to that of 4D-Var Meso-scale Analysis (MA; see Section 2.6) as depicted in Figure 2.6.2. The rectangular plane is determined via a Lambert conformal conic map projection of the Earth's sphere with a map scale factor applied to correct plane expansion or shrinkage associated with projection from the sphere. Hybrid terrain following the relevant coordinates is adopted for the vertical coordinate to reduce the influences of topography as height increases (Subsection 3.5.3). The lowest atmospheric layer is 10 m above the surface, and the model top is at 37,500 m with 96 layers at intervals increasing from 20 m at the bottom to approximately 915 m at the top.

The prognostic variables are horizontal and vertical momentum, mass-virtual potential temperature, total mass density, density of water vapor and hydrometeors (cloud water, cloud ice, rain, snow and graupel), ground temperature, soil water and four of the second-order moments of turbulent fluctuations (including turbulent kinetic energy). The model is operated with a 100/3-second time step.

Initial conditions for the model are generated via MA. Lateral boundary conditions come from the latest available GSM (Section 3.2) forecast with a 3- or 6-hour time lag. Thus, for example, the MSM at 03 and 06 UTC has lateral boundaries from the GSM initiated at 00 UTC.

The model terrain setting relies on the GTOPO30 data set, which is a global digital elevation model with a horizontal grid spacing of 30 arc seconds developed by the U.S. Geological Survey's EROS Data Center (EDC). To avoid computational instability related to steep slopes on terrain, smoothing is performed so that the valid resolution of the terrain adopted in the model is 1.5 times as coarse as that of the model itself.

The Global Land Cover Characteristics (GLCC) data set, also provided by EDC, is used to determine the land-sea attributes of all grids in the model. To alleviate discontinuities in surface wind and temperature fields around coastlines, a tiling approach is introduced in which land/sea sub-grid effects can be considered in surface flux evaluation. Surface-related parameters such as heat capacity, thermal conductivity and roughness are also

[11] The term ASUCA in this section refers to an NWP model incorporating physical processes from the Physics Library. The term sometimes refers only to the related dynamical core in a more narrow sense.

based on land use as described by the GLCC data set. The National Land Numerical Information data set provided by Japan's Ministry of Land, Infrastructure, Transport and Tourism is also referenced for parameters over Japan. Albedo and initial values of soil moisture are prescribed from monthly mean climatology.

Grids on land are further classified in terms of snow presence, and sea grids may be covered with ice. This gives a total of four surface categories: land, snow-covered land, sea and ice-covered sea. Snow-covered areas are analyzed using the high-resolution snow depth analysis system (Subsection 2.8.2), and ice-covered areas are identified from sea ice analysis conducted by the Office of Marine Prediction under JMA's Atmosphere and Ocean Department. As described previously, surface-related parameters are essentially based on land use without assumption of snow- or ice-covered areas. Accordingly, the parameters for these covered grid areas need to be modified with corresponding values.

3.5.3 Dynamics

3.5.3.1 Basic Equations

The governing equations used in the MSM consist of non-hydrostatic, fully compressible equations on spherical curvilinear orthogonal and hybrid terrain-following coordinates with the shallow assumption. The equations are described in flux form.

1. Momentum equations

   The equations of motion are described as

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{1}{J}\rho u\right)+\frac{\partial}{\partial \xi}\left(\frac{1}{J}\rho u U\right)+\frac{\partial}{\partial \eta}\left(\frac{1}{J}\rho u V\right)+\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho u W\right)\\
&\qquad+\gamma R_d \pi\left\{\frac{1}{J}\xi_x\frac{\partial}{\partial \xi}(\rho\theta_m)'+\frac{1}{J}\eta_x\frac{\partial}{\partial \eta}(\rho\theta_m)'+\frac{1}{J}\zeta_x\frac{\partial}{\partial \zeta}(\rho\theta_m)'\right\}\\
&=-\sum_\alpha\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho u q_\alpha W_{t_\alpha}\right)-\frac{1}{J}\rho v\Gamma-\frac{1}{J}\rho v f+\frac{1}{J}F_{\rho u},
\end{aligned}\tag{3.5.1}$$

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{1}{J}\rho v\right)+\frac{\partial}{\partial \xi}\left(\frac{1}{J}\rho v U\right)+\frac{\partial}{\partial \eta}\left(\frac{1}{J}\rho v V\right)+\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho v W\right)\\
&\qquad+\gamma R_d \pi\left\{\frac{1}{J}\xi_y\frac{\partial}{\partial \xi}(\rho\theta_m)'+\frac{1}{J}\eta_y\frac{\partial}{\partial \eta}(\rho\theta_m)'+\frac{1}{J}\zeta_y\frac{\partial}{\partial \zeta}(\rho\theta_m)'\right\}\\
&=-\sum_\alpha\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho v q_\alpha W_{t_\alpha}\right)+\frac{1}{J}\rho u\Gamma+\frac{1}{J}\rho u f+\frac{1}{J}F_{\rho v},
\end{aligned}\tag{3.5.2}$$

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{1}{J}\rho w\right)+\frac{\partial}{\partial \xi}\left(\frac{1}{J}\rho w U\right)+\frac{\partial}{\partial \eta}\left(\frac{1}{J}\rho w V\right)+\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho w W\right)\\
&\qquad+\gamma R_d \pi\left\{\frac{1}{J}\zeta_z\frac{\partial}{\partial \zeta}(\rho\theta_m)'\right\}+\left(\frac{\rho'}{J}-\frac{\pi'}{\overline{\pi}}\frac{\overline{\rho}}{J}\right)\mathrm{g}\\
&=-\sum_\alpha\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho w q_\alpha W_{t_\alpha}\right)+\frac{1}{J}F_{\rho w},
\end{aligned}\tag{3.5.3}$$

   where

$$\Gamma=u\frac{m_2}{m_1}\frac{\partial m_1}{\partial \eta}-v\frac{m_1}{m_2}\frac{\partial m_2}{\partial \xi}.\tag{3.5.4}$$

Here, J is the Jacobian of coordinate transformation from Cartesian coordinates (x, y, z) to generalized coordinates (ξ, η, ζ), defined as

$$J \equiv \begin{vmatrix} \xi_x & \xi_y & \xi_z \\ \eta_x & \eta_y & \eta_z \\ \zeta_x & \zeta_y & \zeta_z \end{vmatrix}, \tag{3.5.5}$$

where, $(\partial\xi/\partial x)_{y,z}$ - a metric of coordinate transformation - is described as ξ_x, and the same description applies to other metrics. A limitation for vertical coordinate to satisfy $\xi_z = \zeta_z = 0$ is introduced, enabling utilization of the Split-Explicit time integration scheme (see Subsection 3.5.3.3). (u, v, w) and (U, V, W) represent velocity components in Cartesian coordinates and generalized coordinates, respectively. $\gamma = C_p/C_v$, where C_p and C_v are the specific heat of dry air at constant pressure and constant volume, respectively. R_d is the gas constant for dry air, and ρ is the total mass density defined as

$$\rho = \rho_d + \rho_v + \rho_c + \rho_r + \rho_i + \rho_s + \rho_g, \tag{3.5.6}$$

where the subscripts d, v, c, r, i, s and g represent dry air, water vapor, cloud water, rain, cloud ice, snow and graupel, respectively.

π is the Exner function defined by

$$\pi = \left(\frac{p}{p_0}\right)^{\frac{R_d}{C_p}}. \tag{3.5.7}$$

The overlined variables $\overline{\rho}$, $\overline{\rho\theta_m}$ and $\overline{\pi}$ represent the hydrostatic state as

$$\gamma R_d \overline{\pi} \frac{1}{J} \zeta_z \frac{\partial}{\partial \zeta}\left(\overline{\rho\theta_m}\right) + \frac{\overline{\rho}\mathfrak{g}}{J} = 0, \tag{3.5.8}$$

and the variables with prime ρ', $(\rho\theta_m)'$ and π' represent perturbation from the hydrostatic state. $\mathfrak{g}$ is gravity acceleration, and f is the Coriolis parameter. q_α is the ratio of the density of water substances α to the total mass density ($\alpha = v, c, r, i, s, g$). $W_{t\alpha}$ is the terminal fall velocity of water substance α. θ_m is defined as

$$\theta_m \equiv \theta\left(1 + \left(\frac{1-\epsilon}{\epsilon}\right) q_v - q_c - q_r - q_i - q_s - q_g\right), \tag{3.5.9}$$

where ϵ is the ratio of R_d to the gas constant for water vapor. $F_{\rho u}$, $F_{\rho v}$ and $F_{\rho w}$ are terms of the surface friction.

Lambert conformal projection is employed, and the map factors m_1 and m_2 (for the x and y directions) are given by

$$m_1 = m_2 = m = \left(\frac{\cos\varphi}{\cos\varphi_1}\right)^{a-1} \left(\frac{1+\sin\varphi_1}{1+\sin\varphi}\right)^{a}, \tag{3.5.10}$$

where φ is the latitude of the relevant point, $\varphi_1 = 30°$, $\varphi_2 = 60°$ and a is given by

$$a = \ln\left(\frac{\cos\varphi_1}{\cos\varphi_2}\right) \Big/ \ln\left\{\frac{\tan\left(45° - \frac{\varphi_1}{2}\right)}{\tan\left(45° - \frac{\varphi_2}{2}\right)}\right\}. \tag{3.5.11}$$

The hybrid terrain-following vertical coordinate which is based on the same approach as the η coordinate (Simmons and Burridge 1981) is adopted to reduce the influences of topography as height increases (Ishida 2007). The vertical coordinate ζ is transformed using the equation:

$$z = \zeta + z_s h(\zeta), \tag{3.5.12}$$

where z is the height and z_s is the surface height. The function $h(\zeta)$ is given by,

$$h(\zeta) = \frac{b\left\{1 - \left(\frac{\zeta}{z_T}\right)^n\right\}}{b + \left(\frac{\zeta}{z_T}\right)^n}, \quad b = \frac{\left(\frac{z_l + z_h}{2z_T}\right)^n}{1 - 2\left(\frac{z_l + z_h}{2z_T}\right)^n}, \tag{3.5.13}$$

where z_T is the model top, $z_l = 2000$m, $z_h = 12000$m and $n = 3$, respectively.

2. Continuity equations

   The continuity equation is described as follows:

$$\frac{\partial}{\partial t}\left(\frac{1}{J}\rho'\right) + \frac{\partial}{\partial \xi}\left(\frac{1}{J}\rho U\right) + \frac{\partial}{\partial \eta}\left(\frac{1}{J}\rho V\right) + \frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho W\right) = -\sum_{\alpha}\frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho q_\alpha W_{t_\alpha}\right) + \frac{1}{J}F_\rho, \tag{3.5.14}$$

   where F_ρ is the tendency by water vapor flux from the surface.

3. Prognostic equation of potential temperature

   The thermodynamic equation is described as

$$\frac{\partial}{\partial t}\left(\frac{1}{J}(\rho\theta_m)'\right) + \frac{\partial}{\partial \xi}\left(\frac{1}{J}\rho\theta_m U\right) + \frac{\partial}{\partial \eta}\left(\frac{1}{J}\rho\theta_m V\right) + \frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho\theta_m W\right) = \frac{1}{J}\left(\rho_d + \frac{\rho_v}{\epsilon}\right)Q_\theta, \tag{3.5.15}$$

   where Q_θ is the diabatic heating.

4. Prognostic equation of water substances

   The prognostic equations for the density of water substances are described as

$$\frac{\partial}{\partial t}\left(\frac{1}{J}\rho q_\alpha\right) + \frac{\partial}{\partial \xi}\left(\frac{1}{J}\rho q_\alpha U\right) + \frac{\partial}{\partial \eta}\left(\frac{1}{J}\rho q_\alpha V\right) + \frac{\partial}{\partial \zeta}\left(\frac{1}{J}\rho q_\alpha (W + W_{t_\alpha})\right) = \frac{1}{J}F_{\rho\alpha}, \tag{3.5.16}$$

   where $F_{\rho\alpha}$ is source or sink term and tendency by flux from the surface for $\alpha = v$.

5. State equation

   The state equation is

$$p = R_d \pi \rho \theta_m. \tag{3.5.17}$$

3.5.3.2 Spatial discretization

The grid structures of the model are the Arakawa C type in the horizontal direction and the Lorenz type in the vertical direction. The equations are spatially discretized using the finite volume method (FVM) to conserve total mass throughout the whole domain in consideration of lateral boundary inflow and outflow. The third-order upwind scheme with the flux limiter function proposed by Koren (1993) is employed to calculate horizontal and vertical advection terms for monotonicity in order to prevent numerical oscillation, and enhance computational efficiency.

3.5.3.3 Time integration

The Runge-Kutta (RK3) scheme (Wicker and Skamarock 2002) is adopted for system time integration. The terms responsible for sound waves and gravity waves are treated using a split-explicit time integration scheme with a short time step. Other time-splitting methods are also used to treat vertical advection associated with strong wind and vertical advection of water substances with high terminal velocity such as rain or graupel.

1. Split-Explicit (HE-VI) Scheme

   The horizontally explicit and vertically implicit (HE-VI) scheme (Klemp *et al.* 2007) is employed. RK3 scheme is also used for the short time step of HE-VI. Forward time integrations with the short time step $\Delta\tau$ are used for the horizontal momentum equations:

$$\left(\frac{1}{J}\rho u\right)^{\tau+\Delta\tau} = \left(\frac{1}{J}\rho u\right)^{\tau} - \gamma R_d \pi^t \left\{\frac{1}{J}\xi_x \frac{\partial}{\partial\xi}(\rho\theta_m)'^{\tau} + \frac{1}{J}\eta_x \frac{\partial}{\partial\eta}(\rho\theta_m)'^{\tau} + \frac{1}{J}\zeta_x \frac{\partial}{\partial\zeta}(\rho\theta_m)'^{\tau}\right\}\Delta\tau + R_u^t \Delta\tau, \quad (3.5.18)$$

$$\left(\frac{1}{J}\rho v\right)^{\tau+\Delta\tau} = \left(\frac{1}{J}\rho v\right)^{\tau} - \gamma R_d \pi^t \left\{\frac{1}{J}\xi_y \frac{\partial}{\partial\xi}(\rho\theta_m)'^{\tau} + \frac{1}{J}\eta_y \frac{\partial}{\partial\eta}(\rho\theta_m)'^{\tau} + \frac{1}{J}\zeta_y \frac{\partial}{\partial\zeta}(\rho\theta_m)'^{\tau}\right\}\Delta\tau + R_v^t \Delta\tau, \quad (3.5.19)$$

   where

$$R_u = -\frac{\partial}{\partial\xi}\left(\frac{1}{J}\rho u U\right) - \frac{\partial}{\partial\eta}\left(\frac{1}{J}\rho u V\right) - \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\rho u W\right) + \frac{1}{J}F'_u, \quad (3.5.20)$$

$$R_v = -\frac{\partial}{\partial\xi}\left(\frac{1}{J}\rho v U\right) - \frac{\partial}{\partial\eta}\left(\frac{1}{J}\rho v V\right) - \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\rho v W\right) + \frac{1}{J}F'_v, \quad (3.5.21)$$

   and $\frac{1}{J}F'_u$ and $\frac{1}{J}F'_v$ are the right hand side of Eq. (3.5.1) and Eq. (3.5.2), respectively. Backward time integrations are used for equations of vertical momentum, potential temperature and density:

$$\left(\frac{1}{J}\rho w\right)^{\tau+\Delta\tau} = \left(\frac{1}{J}\rho w\right)^{\tau} - \left\{\gamma R_d \pi^t \frac{1}{J}\zeta_z \frac{\partial}{\partial\zeta}(\rho\theta_m)'^{\tau+\Delta\tau} + \frac{\rho'^{\tau+\Delta\tau}}{J}\mathfrak{g} - \frac{\pi'^t}{\overline{\pi}}\frac{\overline{\rho}}{J}\mathfrak{g}\right\}\Delta\tau + R_w^t \Delta\tau, \quad (3.5.22)$$

$$\begin{aligned}\left(\frac{1}{J}(\rho\theta_m)'\right)^{\tau+\Delta\tau} &= \left(\frac{1}{J}(\rho\theta_m)'\right)^{\tau} - \left\{\frac{\partial}{\partial\zeta}\left(\frac{1}{J}\zeta_z \theta_m^{\tau}(\rho w)^{\tau+\Delta\tau}\right)\right\}\Delta\tau \\ &\quad - \left\{\frac{\partial}{\partial\xi}\left(\frac{1}{J}\theta_m^{\tau}(\widetilde{\rho U})\right) + \frac{\partial}{\partial\eta}\left(\frac{1}{J}\theta_m^{\tau}(\widetilde{\rho V})\right) + \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\theta_m^{\tau}(\widetilde{\rho W})\right)\right\}\Delta\tau + \frac{1}{J}F^t_{\rho\theta_m}\Delta\tau,\end{aligned} \quad (3.5.23)$$

$$\begin{aligned}\left(\frac{1}{J}\rho'\right)^{\tau+\Delta\tau} =& \left(\frac{1}{J}\rho'\right)^{\tau} - \left\{\frac{\partial}{\partial\zeta}\left(\frac{1}{J}\zeta_z(\rho w)^{\tau+\Delta\tau}\right)\right\}\Delta\tau \\ &- \left\{\frac{\partial}{\partial\xi}\left(\frac{1}{J}\widetilde{(\rho U)}\right) + \frac{\partial}{\partial\eta}\left(\frac{1}{J}\widetilde{(\rho V)}\right) + \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\widetilde{(\rho W)}\right)\right\}\Delta\tau + \frac{1}{J}F'_\rho\Delta\tau,\end{aligned} \tag{3.5.24}$$

where

$$R_w = -\frac{\partial}{\partial\xi}\left(\frac{1}{J}\rho w U\right) - \frac{\partial}{\partial\eta}\left(\frac{1}{J}\rho w V\right) - \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\rho w W\right) + \frac{1}{J}F'_w, \tag{3.5.25}$$

and $\frac{1}{J}F'_w$, $\frac{1}{J}F'_\rho$ and $\frac{1}{J}F_{\rho\theta_m}$ are the right hand side of Eq. (3.5.3), Eq. (3.5.14) and Eq. (3.5.15), respectively, and

$$\widetilde{(\rho U)} = \xi_x(\rho u)^{\tau+\Delta\tau} + \xi_y(\rho v)^{\tau+\Delta\tau}, \tag{3.5.26}$$

$$\widetilde{(\rho V)} = \eta_x(\rho u)^{\tau+\Delta\tau} + \eta_y(\rho v)^{\tau+\Delta\tau}, \tag{3.5.27}$$

$$\widetilde{(\rho W)} = \zeta_x(\rho u)^{\tau+\Delta\tau} + \zeta_y(\rho v)^{\tau+\Delta\tau}. \tag{3.5.28}$$

Here, we can exclude $(\rho w)^{\tau+\Delta\tau}$ from Eq. (3.5.26) - Eq. (3.5.28) due to the limitation for vertical coordinate to satisfy $\xi_z = \zeta_z = 0$ as mentioned in Subsection 3.5.3.1, which enables the vertical implicit treatment of Eq. (3.5.22) - Eq. (3.5.24). Eliminating $\left(\frac{1}{J}(\rho\theta_m)'\right)^{\tau+\Delta\tau}$ and $\left(\frac{1}{J}\rho'\right)^{\tau+\Delta\tau}$ from Eq. (3.5.22) using Eq. (3.5.23) and Eq. (3.5.24), we obtain the one dimensional Helmholtz type equation of $\omega \equiv \left(\frac{1}{J}\rho w\right)^{\tau+\Delta\tau}$ as

$$-\Delta\tau^2\gamma R_d\pi^t\frac{1}{J}\zeta_z\frac{\partial}{\partial\zeta}\left(J\frac{\partial}{\partial\zeta}\left(\zeta_z\theta_m^\tau\omega\right)\right) - \Delta\tau^2\mathfrak{g}\frac{\partial}{\partial\zeta}\left(\zeta_z\omega\right) + \omega = R, \tag{3.5.29}$$

where

$$R = \left(\frac{1}{J}\rho w\right)^{\tau} - \gamma R_d\pi^t\Delta\tau\frac{1}{J}\zeta_z\frac{\partial}{\partial\zeta}\left\{(\rho\theta_m)'^{\tau} + JR'_{\theta_m}\Delta\tau\right\} - \Delta\tau\mathfrak{g}\left(\frac{1}{J}\rho'^{\tau} + R'_\rho\Delta\tau\right) + R'_w\Delta\tau, \tag{3.5.30}$$

and

$$R'_w = \frac{\pi'^t}{\overline{\pi}}\frac{\overline{\rho}}{J}\mathfrak{g} + R_w = -\left(1 - \frac{\pi^t}{\overline{\pi}}\right)\frac{\overline{\rho}}{J}\mathfrak{g} + R_w, \tag{3.5.31}$$

$$R'_{\theta_m} = -\left\{\frac{\partial}{\partial\xi}\left(\frac{1}{J}\theta_m^\tau\widetilde{(\rho U)}\right) + \frac{\partial}{\partial\eta}\left(\frac{1}{J}\theta_m^\tau\widetilde{(\rho V)}\right) + \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\theta_m^\tau\widetilde{(\rho W)}\right)\right\} + \frac{1}{J}F^t_{\rho\theta_m}, \tag{3.5.32}$$

$$R'_\rho = -\left\{\frac{\partial}{\partial\xi}\left(\frac{1}{J}\widetilde{(\rho U)}\right) + \frac{\partial}{\partial\eta}\left(\frac{1}{J}\widetilde{(\rho V)}\right) + \frac{\partial}{\partial\zeta}\left(\frac{1}{J}\widetilde{(\rho W)}\right)\right\} + \frac{1}{J}F^t_\rho. \tag{3.5.33}$$

Considering $W = 0$ at the upper and lower boundary and $u = 0$, $v = 0$ at the lower boundary, upper and lower boundary conditions are given by $\omega = 0$.

2. Time splitting of vertical advection

Using RK3 as a time integration scheme and a flux limiter function as an advection scheme, the CFL condition of 3-dimensional advection is given by

$$C_\xi + C_\eta + C_\zeta < 1.25, \tag{3.5.34}$$

where C_ξ, C_η and C_ζ are the Courant number in the ξ, η and ζ direction, respectively. As this condition can be hard to fulfill with typhoons characterized by stormy horizontal winds and strong updrafts, time splitting of vertical advection is adopted in consideration of computational efficiency and the model's memory alignment with vertical indices placed innermost.

In the time splitting method, each RK3 stage is divided into substeps depending on the relevant Courant numbers. As each RK3 stage can be regarded as a forward time integration with the time steps of $\Delta t/3$, $\Delta t/2$ and Δt, respectively (as shown in Figure 3.5.1), these time steps are used to evaluate the Courant numbers for each stage. For each column, the number of substeps N is set to satisfy

$$C_\xi + C_\eta + \frac{C_\zeta}{N} < 1.25, \tag{3.5.35}$$

at each RK3 stage. When time splitting is invoked, the forward form integration at each RK3 stage is replaced with RK3 (i.e., RK3 is nested in the original RK3 time integration) as shown in Figure 3.5.2. This involves greater computational cost, but produces the desired higher stability.

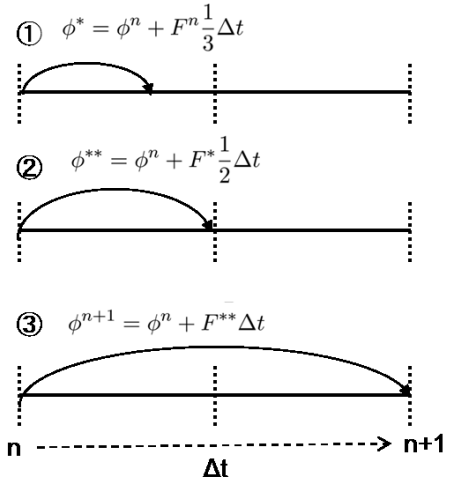


Figure 3.5.1: Schematic of RK3 time integration scheme.

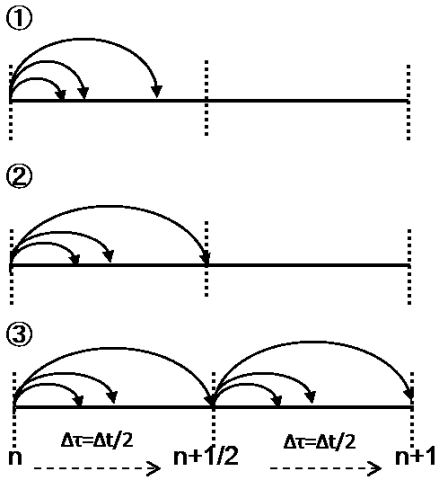


Figure 3.5.2: Schematic of time-splitting of vertical advection. The case needs to split ③ stage into 2 substeps.

When time-splitting is invoked, fields are updated using the horizontal flux F_ξ and F_η first, and the vertical flux F_ζ is then evaluated with the integrated field as follows.

$$\phi^{H*} = \phi^n - \left(\frac{\partial}{\partial \xi} {F_\xi}^n + \frac{\partial}{\partial \eta} {F_\eta}^n \right) \Delta\tau, \tag{3.5.36}$$

$$\phi^{n+1} = \phi^{H*} - \left(\frac{\partial}{\partial \zeta} {F_\zeta}^{H*} \right) \Delta\tau. \tag{3.5.37}$$

3. Time splitting of vertical advection of water substances

   To stabilize integration for the vertical advection of water substances with high terminal velocity, a time splitting method is adopted. The short time step $\Delta\tau_1$ for sedimentation is determined from the Courant number $C_{t\zeta}$ as follows.

$$\Delta\tau_1 = \begin{cases} \Delta t & (\max(C_{t\zeta}) \leq 1) \\ \beta \dfrac{\Delta t}{\max(C_{t\zeta})} & (\max(C_{t\zeta}) > 1), \end{cases} \tag{3.5.38}$$

   where

$$C_{t\zeta} = (W^n + W^n_{t\alpha})\Delta t/\Delta\zeta. \tag{3.5.39}$$

   Here, W^n is W at the time level n, β is a parameter for determining the short time step, and $c = 0.9$ is used.

   After time integration with $\Delta\tau_1$, the residual time step is $\Delta t' = \Delta t - \Delta\tau_1$. The next short time step $\Delta\tau_2$ is decided from the Courant number $C'_{tz} = (W^n + W^{n+\tau_1}_{t\alpha})\Delta t'/\Delta\zeta$ and the time integration with $\Delta\tau_2$ is calculated. This procedure is repeated until no residual time step is left.

3.5.3.4 Boundary Conditions

Rayleigh damping,

$$\mathrm{D}_R = -n(x,y,z)\{\phi - \phi_{EXT}\}, \quad n(x,y,z) = max\left(\frac{\mathrm{D}}{m_L}, \frac{\mathrm{D}}{m_U}\right), \tag{3.5.40}$$

is added near the lateral and upper boundaries to the time tendencies of horizontal and vertical momentum, potential temperature and the mixing ratio of water vapor, where ϕ is the prognostic variable and ϕ_{EXT} is the value of the external model. m_L and m_U represent coefficients that determine the 1/e-folding time for the lateral and upper boundaries, respectively, while $m_L = 250$ seconds and $m_U = 125$ seconds. The location-based function D is unity at the boundary and decreases with subsequent grid point distance.

3.5.4 Cloud Microphysics

An explicit three-ice bulk microphysics scheme (Ikawa and Saito 1991) based on Lin *et al.* (1983) is incorporated. The scheme predicts the mixing ratios of water vapor and five hydrometeors designated by q_x where x denotes categories defined as v for water vapor, w for cloud water, r for rain, i for cloud ice, s for snow, and g for graupel. The cloud microphysical processes simulated in this scheme are illustrated in Figure 3.5.3 (see Table 3.5.1 for a list of symbols used in the figure). In this scheme, some basic cloud microphysical processes (e.g., nucleation of cloud particles, conversion from cloud particles to precipitation particles) are parameterized because the related processes occur within a shorter time than the integration time step. However, most of the cloud microphysical processes can be applied directly to calculation related to the size distribution assumed in each hydrometeor category.

The number-weighted mean of the temporal tendency of one cloud microphysical variable ϕ relating to one cloud microphysical process in each particle gives the grid-mean temporal tendency of ϕ as

$$\frac{d\phi}{dt} = \int_0^\infty \frac{d\phi_0(D)}{dt} n(D)\, dD, \tag{3.5.41}$$

where $\frac{d\phi_0(D)}{dt}$ is the temporal tendency of ϕ relating to one cloud microphysical process in a particle with diameter D, and $n(D)\, dD$ is the number of particles per unit volume of air with diameters from D to $D + dD$. Hydrometeor size distribution therefore significantly affects time tendency of cloud microphysical variables relating to cloud microphysical processes.

3.5.4.1 Mass-size Relationships

The mass-size relationships represent particle mass m_x as a function of particle diameter D_x for determination of mixing ratios or mass weighted-mean variables. For example, the mixing ratio q_x is generally formulated as

$$q_x = \frac{1}{\rho_a} \int_0^\infty m_x(D_x) n_x(D_x) dD_x, \tag{3.5.42}$$

where ρ_a is the density of air.

The mass-size relationship ($m_x(D_x)$) with the power-law is formulated as

$$m_x(D_x) = a_x D_x^{b_x}. \tag{3.5.43}$$

For the category of cloud ice and snow, a_x and b_x (x for i and s) are set to 0.0185kg m$^{-1.9}$ and 1.9 respectively (Brown and Francis 1995). Particle sphericity ($b_x = 3$) is assumed in the hydrometer categories of graupel, rain and cloud water, and density is constant in each category (ρ_x). Accordingly, the mass-size relationship is formulated as

$$m_x(D_x) = \frac{\pi}{6} \rho_x {D_x}^3. \tag{3.5.44}$$

3.5.4.2 Size Distribution Functions

1. **Rain, cloud ice and graupel**
   The size distributions of rain, cloud ice and graupel are assumed to follow an exponential function:

   $$n_x(D_x) = N_{0x} \exp(-\lambda_x D_x), \tag{3.5.45}$$

   where N_{0x} is the intercept and λ_x is the slope parameter of the size distribution. Accordingly, the moment formula for rain, cloud ice and graupel is calculated as

   $$M_x(p) = \int_0^\infty {D_x}^p n_x(D_x)\, dD_x = N_{0x} \frac{\Gamma(1+p)}{{\lambda_x}^{1+p}}, \tag{3.5.46}$$

   where Γ is the gamma function and $M_x(p)$ is the p-th moment of $n_x(D_x)$. The number concentration N_x is the 0-th moment of $n_x(D_x)$, and is therefore calculated as

   $$N_x = \int_0^\infty n_x(D_x)\, dD_x = M_x(0) = \frac{N_{0x}}{\lambda_x}. \tag{3.5.47}$$

   The mixing ratio q_x is calculated using $m_x(D_x)$ as

   $$q_x = \frac{1}{\rho_a} \int_0^\infty m_x(D_x) n_x(D_x)\, dD_x = \frac{a N_{0x}}{\rho_a} \frac{\Gamma(1+b_x)}{{\lambda_x}^{1+b_x}}. \tag{3.5.48}$$

   The intercepts N_{0x} are assumed to be constant or dependent only on temperature for cloud ice. The slope parameter is therefore expressed:

   $$\lambda_x = \left(\frac{a_x N_{0x} \Gamma(1+b_x)}{\rho_a q_x} \right)^{\frac{1}{1+b_x}}. \tag{3.5.49}$$

2. **Snow**
   The size distribution function itself is not directly used for snow, but moments are parameterized based on Field *et al.* (2007) as follows:

   $$M_s(p) = \int_0^\infty D_s^p n_s(D_s)\, dD_s = A(p) \exp\left[B(p) T_c\right] M_s^{c(p)}(2), \tag{3.5.50}$$

   $$A(p) = \exp\left[13.6 - 7.76p + 0.479p^2\right], \tag{3.5.51}$$

$$B(p) = -0.0361 + 0.0151p + 0.00149p^2, \tag{3.5.52}$$

$$C(p) = 0.807 + 0.00581p + 0.0457p^2, \tag{3.5.53}$$

where T_c is the temperature in celsius (°C).
The number concentration N_s, the 0-th moment of $n_s(D_s)$, is calculated as

$$N_s = M_s(0) = A(0)\exp\left[B(0)\,T_c\right] M_s^{c(0)}(2)\,. \tag{3.5.54}$$

Based on (3.5.43), the mixing ratio q_s is

$$q_s = \frac{1}{\rho_a}\int_0^\infty m_s(D_s) n_s(D_s)\, dD_s = \frac{1}{\rho_a} a_s M_s(b_s) = \frac{1}{\rho_a} a_s A(b_s)\exp\left[B(b_s)(T_c)\right] M_s^{c(b_s)}(2)\,. \tag{3.5.55}$$

Accordingly, the second moment $M_s(2)$ is calculated as

$$M_s(2) = \left[\frac{\rho_a q_s}{a_s}\frac{1}{A(b_s)\exp\left[B(b_s)\,T_c\right]}\right]^{\frac{1}{c(b_s)}}. \tag{3.5.56}$$

3. **Cloud water**
As cloud water is assumed to be monodisperse, its size distribution follows the δ-function:

$$n_w(D_w) = N_w \delta(D_w - \overline{D_w}), \tag{3.5.57}$$

where $\delta(x)$ satisfies the equation $\int_{-\infty}^{\infty} \delta(x-a) f(x)\, dx = f(a)$ and $\overline{D_w}$ represents the diameter of the monodisperse particle. The moment formula for cloud water is given by

$$M_w(p) = N_w \overline{D_w}^p. \tag{3.5.58}$$

The number concentration of cloud water N_w is always assumed to be constant and is set to 1.0×10^8 m^{-4} in this scheme.
The mixing ratio is calculated as

$$q_w = \frac{\rho_w}{\rho_a}\frac{\pi}{6} M_w(3) = \frac{\rho_w}{\rho_a}\frac{\pi}{6} N_w \overline{D_w}^3. \tag{3.5.59}$$

The diameter is therefore determined as

$$\overline{D_w} = \left(\frac{6\rho_a q_w}{\pi \rho_w N_w}\right)^{\frac{1}{3}}. \tag{3.5.60}$$

3.5.4.3 Fall Velocity and Sedimentation

The simple power law is adopted for the fall velocity-size relationship ($U_x(D_x)$) given by

$$U_x(D_x) = \alpha_{ux} D_x^{\beta_{ux}} \left(\frac{\rho_0}{\rho_a}\right)^{\gamma_{ux}}, \tag{3.5.61}$$

where ρ_0 is the density of the reference air, and α_{ux}, β_{ux} and γ_{ux} are constants in each hydrometeor category. The forecast model for the MSM calculates sedimentation processes for hydrometers in its dynamical core. The cloud microphysics scheme diagnoses mass-weighed mean fall velocities as terminal velocities W_{tx}:

$$W_{tx} = \frac{\int_0^\infty U_x(D_x) m_x(D_x) n_x(D_x)\, dD_x}{\int_0^\infty m_x(D_x) n_x(D_x)\, dD_x}. \tag{3.5.62}$$

The cloud microphysics scheme passes W_{tx} to the dynamical core.

Table 3.5.2 shows the characteristics of each hydrometeor class. More information on the treatment of each cloud microphysical process in this scheme can be found in the references.

Table 3.5.1: List of symbols in Figure 3.5.3

Notation	Description
	Production terms
p_a__ppp_b	Production of category "a" converted from category "b" via the process "ppp"
p_a__ppp_a_b	Growth of category "a" based on capture of category "b" via the process "ppp"
p_a__ppp_b_c	Generation of category "a" based on category "b" capturing category "c" via the process "ppp"
	Categories by hydrometeor
v	Water vapor
w	Cloud water
r	Rain
i	Cloud ice
s	Snow
g	Graupel
	Cloud microphysical processes
evp	Evaporation
cnd	Condensation
aut, cn	Conversion
ac	Accretion
mlt	Melting
nud	Nucleation
dep	Deposition
sub	Sublimation
frz	Freezing

Table 3.5.2: Assumed hydrometeor parameters and characteristics

Variable	Rain q_r[kg kg^{-1}]	Snow q_s[kg kg^{-1}]	Graupel q_g[kg kg^{-1}]	Cloud ice q_i[kg kg^{-1}]	Cloud water q_w[kg kg^{-1}]
Size distribution [m^{-4}]	$n_r(D_r) = N_{0r}\exp(-\lambda_r D_r)$ $N_{0r} = 8.0\times10^6$	not directly used	$n_g(D_g) = N_{0g}\exp(-\lambda_g D_g)$ $N_{0g} = 1.1\times10^6$	$n_i(D_i) = N_{0i}\exp(-\lambda_i D_i)$ $N_{0i} = N_{00i}\exp(-0.12T_c)$ $N_{00i} = 4.0\times10^7$ $T_c = \max(T_c, -40°\text{C})$	monodisperse, $\overline{D_w} = \left[\frac{6q_w\rho_a}{\pi N_w\rho_w}\right]^{\frac{1}{3}}$ $N_w = 1.0\times10^8$
Mass[kg]-size[m] relationship	$m_r = \frac{\pi}{6}\rho_r D_r^3$	$m_s = 0.0185D_s^{1.9}$	$m_g = \frac{\pi}{6}\rho_g D_g^3$	$m_i = 0.0185D_i^{1.9}$	$m_w = \frac{\pi}{6}\rho_w D_w^3$
Density [kg m^{-3}]	$\rho_r = 1.0\times10^3$	not used	$\rho_g = 3.0\times10^2$	not used	$\rho_w = 1.0\times10^3$
Fall velocity [m/s]	$U_x(D_x) = \alpha_{ux}D_x^{\beta_{ux}}\left(\frac{\rho_0}{\rho_a}\right)^{\gamma_{ux}}$				not considered
	$\alpha_{ur} = 842$ $\beta_{ur} = 0.8$ $\gamma_{ur} = 0.5$	$\alpha_{us} = 17$ $\beta_{us} = 0.5$ $\gamma_{us} = 0.5$	$\alpha_{ug} = 124$ $\beta_{ug} = 0.64$ $\gamma_{ug} = 0.5$	$\alpha_{ui} = 71.34$ $\beta_{ui} = 0.6635$ $\gamma_{ui} = 0.5$	

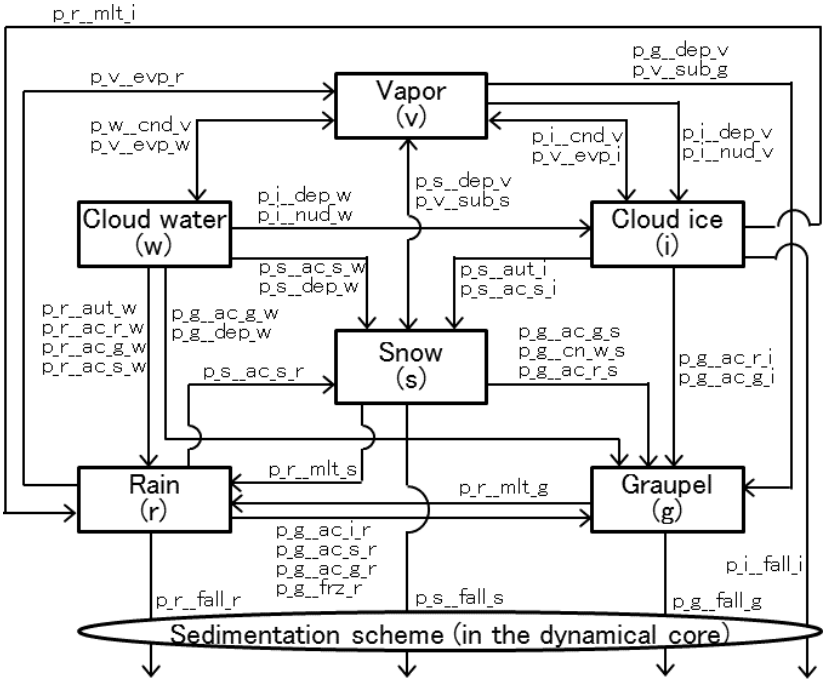


Figure 3.5.3: Cloud microphysical processes in the MSM. For a list of symbols, see Table 3.5.1.

3.5.5 Convective Parameterization

A temporal tendency of a grid mean value ϕ associated with subgrid convection under the isobaric coordinate is generally described as

$$\rho\left(\frac{\partial\overline{\phi}}{\partial t}\right)_{\text{convection}} = \rho g\frac{\partial\rho\overline{\phi' w'}}{\partial p} + S_{\phi}, \tag{3.5.63}$$

where $\overline{\phi' w'}$ and S_{ϕ} represent a subgrid transport flux and a source term respectively. To parameterize the subgrid flux and source term for heat and moisture, a mass flux convective parameterization based on the Kain-Fritsch (KF) (Kain and Fritsch 1990; Kain 2004) type scheme is employed in the MSM.

3.5.5.1 Cloud Model

The scheme parameterizes convection using a cloud model based on a one-dimensional entraining/detraining plume model incorporating detailed treatment for interactions between convective updraft and the surrounding air. The cloud model consists of an updraft mass flux representing the convective activity within a column. Downdraft is not treated in the scheme. The cloud model involves the assumption that the convective updraft area is in a steady state. Thus, mass conservation is represented as

$$0 = \rho g\frac{\partial M^u}{\partial p} + E^u - D^u, \tag{3.5.64}$$

where M, E and D represent convective mass flux, and entrainment / detrainment from / to environmental air respectively. The superscript u shows updraft. For ϕ, the steady state equations can be represented as

$$0 = \rho g\frac{\partial M^u\phi^u}{\partial p} + E^u\overline{\phi} - D^u\phi^u + S^u_{\phi}. \tag{3.5.65}$$

The scheme also employs the assumption that the area of convection is small enough relative to that of a grid-box. Based on this assumption, the subgrid flux can be represented as

$$\rho\overline{\phi' w'} = M^u\left(\phi^u - \overline{\phi}\right). \tag{3.5.66}$$

Substituting Eq. (3.5.64), Eq. (3.5.65) and Eq. (3.5.66) into Eq. (3.5.63), the temporal tendency of ϕ due to subgrid convection can be re-written as a summation of the detrainment and compensating subsidence terms:

$$\rho\left(\frac{\partial\overline{\phi}}{\partial t}\right)_{\text{convection}} = D^u\left(\phi^u - \overline{\phi}\right) - \rho g M^u \frac{\partial\overline{\phi}}{\partial p}. \tag{3.5.67}$$

3.5.5.2 Determination of Variables in Updraft

Vertical integration with Eq. (3.5.64) and Eq. (3.5.65) from the lifting condensation level (LCL), is applied to determine ϕ^u and vertical profiles of M^u. During this integration, E^u and D^u, representing entrainment and detrainment, are calculated with consideration of the mixing process between the updraft and environmental air.

Following the original KF scheme (Kain and Fritsch 1990), the interaction between updraft and the environment associated with the turbulent mixing is estimated at each vertical model level to determine E^u and D^u. It is assumed that the turbulent mixing occurs very near the periphery of the updraft, and that large number of subparcel-like mixtures of the updraft and the environment form at various ratios that can be described by a Gaussian probability distribution function the mean of with a mean of 0.5, representing a scenario in which environmental mass and updraft mass are likely to be equally mixed in subparcels.

The relationship linking δM_e, δM_u and δM_t, defined as the entrained mass from the environment, the updraft mass mixed with the entrained mass and the total mass respectively, can be expressed as

$$\delta M_u + \delta M_e = \delta M_t = \delta M_t \int_0^1 f(x)dx, \tag{3.5.68}$$

$$\delta M_e = \delta M_t \int_0^1 x f(x)dx, \tag{3.5.69}$$

$$\delta M_u = \delta M_t \int_0^1 (1-x) f(x)dx, \tag{3.5.70}$$

where x and $f(x)$ are the fraction of environmental mass in mixed subparcels and the probability distribution function as a function of x respectively.

δM_e, which determines the amounts of entrainment and detrainment, is inversely proportional to the updraft radius, R:

$$\delta M_e = M_{u0}(a\delta P/R), \tag{3.5.71}$$

where δP is the vertical grid thickness in the pressure coordinate, and the factor a is set to 0.03 m Pa^{-1} as a constant. The radius of the updraft is used only for entrainment rate estimation. The radius R is set to a constant of 750 m.

Consequently, mixtures with positive buoyancy against the environment entrain into updraft, while those with negative buoyancy detrain from it. When x_n is the fraction with which mixed air is neutral against environmental air, E^u and D^u can be formulated as follows:

$$E^u \delta p = \rho g \delta M_t \int_0^{x_n} x f(x)dx, \tag{3.5.72}$$

$$D^u \delta p = \rho g \delta M_t \int_{x_n}^1 (1-x) f(x)dx. \tag{3.5.73}$$

The vertical velocity of the updraft depends on buoyancy and hydrometeor weights. The updraft terminates when the mass flux becomes emaciated through detrainment or when its vertical velocity vanishes.

In the scheme, parameterized convection is divided into deep and shallow convection types. An updraft that does not reach the minimum cloud depth for deep convection is regarded as shallow convection. The minimum cloud depth is a function of temperature at the cloud base (LCL).

3.5.5.3 Treatment of convective precipitation

For cloud water content q_c, Eq. (3.5.65) applies as follows:

$$0 = \rho g \frac{\partial M^u q_c^u}{\partial p} + E^u \overline{q_c} - D^u q_l^u + c_c^u - P_c, \tag{3.5.74}$$

where c_c^u[kg m^{-3}s^{-1}] and P_c[kg m^{-3}s^{-1}] are condensation/glaciation and precipitation generation terms respectively. In updraft evaluation, water condensate /glaciate (depending on updraft temperature at each level) are considered.

The precipitation generation rate is formulated as a Kessler type autoconversion scheme:

$$P_c = A \max\left(q_c^u - q_{c0}, 0\right). \tag{3.5.75}$$

The threshold q_{c0} is set to a constant value of 2.0×10^{-3}kg kg^{-1}. The conversion rate A is set to $\rho g M^u / \Delta p$ so that excess hydrometeors over the threshold are immediately taken out of the updraft as precipitation. The generated precipitation is added to the tendencies shown as per Eq. (3.5.67) for rain, snow and graupel rather than being represented as falling to the ground. The precipitation fall process is calculated in a sedimentation scheme outside the convective parameterization scheme.

3.5.5.4 Closure

The closure process finally determines the magnitude of convective mass flux. Closure is the only difference between deep and shallow convection.

For deep convection, the magnitude of mass flux is determined so that stabilized vertical profiles after convection satisfy the condition that CAPE in the final state should be less than 15% of the initial value. Seeking the stabilized state usually requires iterative adjustment of mass fluxes. The updraft mass flux at the LCL, M_{u0}[kgm^{-2}s^{-1}], is initially given as follows with the assumption that vertical velocity w_{one} is 1 m/s and the initial area occupied by convection is 1 % of a grid:

$$M_{u0} = 0.01 \rho_{u\mathrm{LCL}} w_{\mathrm{one}}, \tag{3.5.76}$$

where $\rho_{u\mathrm{LCL}}$ is the density of the updraft mass flux at the LCL. Using M_{u0}, the vertical integration of Eqs. (3.5.64) and (3.5.65) from the LCL are calculated. The value of $\overline{\phi}_{\mathrm{stabilized}}$, the grid mean $\overline{\phi}$ after stabilization, is then calculated based on time integration of Eq. (3.5.67) with the period of the lifetime of convection τ_{lifetime}. If the post-stabilization CAPE is still 15 % more than the initial value, the mass flux at the LCL is increased. Iterating this process, the final value of M_{u0} is determined.

For shallow convection, the mass flux at the LCL is determined using the maximum turbulent kinetic energy (TKE),

$$M_{\mathrm{u0,shallow}} = \frac{\mathrm{TKE}_{\mathrm{max}}}{k_0} \frac{\Delta p_{\mathrm{parcel}}}{g \tau_{\mathrm{lifetime}}}, \tag{3.5.77}$$

where $\mathrm{TKE}_{\mathrm{max}}$ is the maximum value of TKE and is set to 1 m^2 s^{-2} for the MSM. k_0 is set to 20 m^2 s^{-2}. $\Delta p_{\mathrm{parcel}}$ [Pa] is the pressure depth from the LCL to the highest model level at which the depth is no larger than 50 hPa.

The temporal tendency of the adjusted physical quantity ϕ can be determined as

$$\left.\frac{d\overline{\phi}}{dt}\right|_{\mathrm{convection}} = \frac{\overline{\phi}_{\mathrm{stabilized}} - \overline{\phi}_{\mathrm{initial}}}{\tau_{\mathrm{lifetime}}}, \tag{3.5.78}$$

where τ_{lifetime} is set to a constant value of 600 s.

3.5.5.5 Diagnosis of the Convection as a Triggering Process

Diagnosis is performed to determine whether each column is convectively unstable enough to activate the scheme. This is done for every timestep to identify grids where parameterized convection should occur, and involves two steps.

Firstly, the temperature of a lifted parcel is compared with the environmental temperature $\overline{T}$. The parcel temperature T_{diag} is defined as the lifted air mass at the LCL T_{LCL} with a perturbation:

$$T_{\mathrm{diag}} = T_{\mathrm{LCL}} + \Delta T'. \tag{3.5.79}$$

The second term on the right represents perturbation corresponding to subgrid-scale buoyant flux associated with the planetary boundary layer process.

$$\Delta T' = \max\left[A_{\mathrm{plume}}, \min\left[B_{\mathrm{plume}}\sigma_{T_v}, G_{\max} z_h\right]\right]\Pi, \tag{3.5.80}$$

$$\sigma_{T_v} = 1.93\overline{w'\theta'_{vs}}/w_m, \tag{3.5.81}$$

$$w_m^3 = u_*^3 + 0.25 z_h \overline{w'b'_s}, \tag{3.5.82}$$

where Π, z_h, u^* and $\overline{w'b'_s}$ are the Exner function, the height of the planetary boundary layer, friction velocity and turbulent buoyant flux at the surface respectively. The parameters are set to A_{plume} =0.2 K, B_{plume} =3.26 and $G_{\max}$ =10^{-3} Km^{-1}. The diagnosis process lifts the potential updraft source layer (USL) of the lowest 50 hPa depth adiabatically to its LCL for determination of T_{LCL}. If $T_{\mathrm{diag}} > \overline{T}$ is not satisfied, the base of the potential USL is moved up to the next model level and comparison of T_{diag} with $\overline{T}$ is repeated as long as the base of the potential USL is below the lowest 300 hPa of the atmosphere.

As the second step of diagnosis, if $T_{\mathrm{diag}} > \overline{T}$ is satisfied, the scheme calculates CAPE. To determine this value, the updraft variables are provisionally calculated by vertically integrating Eq. (3.5.64) and Eq. (3.5.65). If the updraft parcel has a positive CAPE value, parameterized convection is activated in the column.

3.5.6 Radiation

The radiation process employed in the MSM is almost identical to that in the GSM, as the codes of the GSM radiation process were ported into the MSM. The details are described in Subsection 3.2.3. Some differences are outlined below.

3.5.6.1 Radiatively Active Constituents

Radiatively active gases accounted for in the MSM are identical to those in the GSM, although certain representations of optical properties differ. Some gas concentrations differ (O_3:375 ppmv, CH_4:1.75 ppmv, N_2O:0.28 ppmv) and aerosol optical depth climatology is based on total-column values from MODerate resolution Imaging Spectroradiometer (MODIS) and Total Ozone Mapping Spectrometer (TOMS) observations with seasonal variations. Other optical properties of aerosols are specified as continental and maritime background values without seasonal variation.

3.5.6.2 Cloud Properties

The effective radius of cloud ice particles $r_{\mathrm{e}}[\mu\mathrm{m}]$ is diagnosed using a linearized equation based on Ou and Liou (1995):

$$r_{\mathrm{e}} = \begin{cases} 20 & T < -60 \\ 20 + \dfrac{130(T+60)}{40} & -60 \le T \le -20 \\ 150 & -20 < T \end{cases}, \tag{3.5.83}$$

where T [°C] is the air temperature. The effective radius of cloud water droplets is fixed at 15 μm.

Cloud optical properties for short-wave radiation are parameterized following Slingo (1989) for liquid droplets and Ebert and Curry (1992) for ice particles. Cloud optical thickness for long-wave radiation is parameterized as per Hu and Stamnes (1993) for liquid droplets and as per Ebert and Curry (1992) for ice particles.

3.5.6.3 Cloud Fraction

The cloud fraction for the radiation scheme is diagnosed using a partial condensation method based on Sommeria and Deardorff (1977) and Mellor (1977), which is also employed to evaluate subgrid scale buoyancy flux in the boundary layer scheme (Subsection 3.5.7). This method involves calculation to determine the variance of the gridbox saturation deficit, associated with fluctuations of liquid water potential temperature (θ_l') and total water specific humidity (q_w'). Assuming unimodal Gaussian distribution for the deficit, the liquid cloud fraction (R_l) and the gridbox mean liquid water content ($\overline{q_l}$) are given by

$$R_l = \frac{1}{\sqrt{2\pi}(2\sigma_s)} \int_0^\infty \exp\left[-\frac{(q_l - a_l\Delta\overline{q})^2}{2(2\sigma_s)^2}\right] dq_l = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{Q_1}{\sqrt{2}}\right)\right], \tag{3.5.84}$$

$$\overline{q_l} = \frac{1}{\sqrt{2\pi}(2\sigma_s)} \int_0^\infty q_l \exp\left[-\frac{(q_l - a_l\Delta\overline{q})^2}{2(2\sigma_s)^2}\right] dq_l = 2\sigma_s\left[R_l Q_1 + \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{Q_1^2}{2}\right)\right], \tag{3.5.85}$$

where $\Delta\overline{q} = \overline{q_w} - q_{sat}(\overline{T_l})$, T_l denotes the liquid water temperature, and $q_{sat}(T)$ is saturated specific humidity at the temperature (T). a_l and Q_1 are given by

$$a_l = \left[1 + \frac{L}{C_p}\left(\frac{\partial q_{sat}}{\partial T}\right)_{T=T_l}\right]^{-1}, \tag{3.5.86}$$

$$Q_1 = \frac{a_l\Delta\overline{q}}{2\sigma_s}, \tag{3.5.87}$$

where L denotes the latent heat of condensation and C_p is specific heat at the constant pressure of dry air, and σ_s is a half of the standard deviation of the saturation deficit. σ_s can be represented using the turbulent prognostic variables ($\overline{\theta_l'^2}$, $\overline{q_w'^2}$, and $\overline{\theta_l' q_w'}$) in the boundary layer scheme (Subsection 3.5.7) as follows:

$$\sigma_s^2 = \frac{1}{4}\left(a_l^2\overline{q_w'^2} - 2a_l b_l\overline{\theta_l' q_w'} + b_l^2\overline{\theta_l'^2}\right), \tag{3.5.88}$$

$$b_l = a_l\Pi\left(\frac{\partial q_{sat}}{\partial T}\right)_{T=T_l}, \tag{3.5.89}$$

where Π is the Exner function.

For parametrizing the ice cloud fraction, a pragmatic approach based on Wilson and Ballard (1999) is employed. The ice cloud fraction R_i is given by

$$R_i = \begin{cases} 0 & \dfrac{\overline{q_f}}{b_{si}} = 0 \\ \dfrac{1}{2}\left(6\dfrac{\overline{q_f}}{b_{si}}\right)^{2/3} & 0 < \dfrac{\overline{q_f}}{b_{si}} \le \dfrac{1}{6} \\ 1 - 4\cos^2\phi & \dfrac{1}{6} < \dfrac{\overline{q_f}}{b_{si}} \le 1 \\ 1 & 1 < \dfrac{\overline{q_f}}{b_{si}} \end{cases}, \tag{3.5.90}$$

where $\overline{q_f}$ is the gridbox mean ice water content including snow aggregates. b_{si} and ϕ are given by

$$b_{si} = 2\sqrt{6}\sigma_s, \tag{3.5.91}$$

$$\phi = \frac{1}{3}\left\{\cos^{-1}\left[\frac{3}{2\sqrt{2}}\left(1 - \frac{\overline{q_f}}{b_{si}}\right)\right] + 4\pi\right\}. \tag{3.5.92}$$

A total cloud fraction R is obtained by combination of liquid and ice cloud fraction. Assuming that the liquid and ice clouds are minimally overlapped, R is given by

$$R = R_l + R_i - \min\langle R_l, R_i\rangle. \tag{3.5.93}$$

3.5.6.4 Radiative Timesteps

Long-wave and short-wave radiation schemes are fully calculated every 15 minutes, while heating rates associated with long-wave and short-wave radiation are corrected at every time step using the surface temperature and the solar zenith angle, respectively.

3.5.7 Boundary Layer

The boundary layer scheme represents vertical turbulent transport of momentum, heat and water. The fluxes exhibit the temporal tendency of the variable $\phi(= u, v, \theta_l, q_w)$ associated with turbulent transport as follows:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial}{\partial z}\overline{w'\phi'}. \tag{3.5.94}$$

The MSM employs the Mellor-Yamada-Nakanishi-Niino Level 3 model (MYNN3) (Nakanishi and Niino 2009) as a boundary layer scheme. This is a second order turbulent closure model in which it is assumed that the third-order moments of turbulent fluctuation can be depicted by lower-order moments.

3.5.7.1 Prognostic Equations and Fluxes

In the MYNN3 with boundary layer approximation, in which horizontal derivatives are ignored, just only four turbulent prognostic variables (including turbulent kinetic energy (TKE)), remain:

$$\frac{\partial q^2}{\partial t} = -2\left(\overline{u'w'}\frac{\partial \overline{u}}{\partial z} + \overline{v'w'}\frac{\partial \overline{v}}{\partial z}\right) + 2\frac{g}{\overline{\theta_v}}\overline{w'\theta_v'} - 2\varepsilon + \frac{\partial}{\partial z}\left(q\ell S_q \frac{\partial q^2}{\partial z}\right), \tag{3.5.95}$$

$$\frac{\partial \overline{\theta_l'^2}}{\partial t} = -2\overline{w'\theta_l'}\frac{\partial \overline{\theta_l}}{\partial z} - 2\varepsilon_\theta + \frac{\partial}{\partial z}\left(q\ell S_\theta \frac{\partial \overline{\theta_l'^2}}{\partial z}\right), \tag{3.5.96}$$

$$\frac{\partial \overline{q_w'^2}}{\partial t} = -2\overline{w'q_w'}\frac{\partial \overline{q_w}}{\partial z} - 2\varepsilon_q + \frac{\partial}{\partial z}\left(q\ell S_{q_w} \frac{\partial \overline{q_w'^2}}{\partial z}\right), \tag{3.5.97}$$

$$\frac{\partial \overline{\theta_l'q_w'}}{\partial t} = -\overline{w'\theta_l'}\frac{\partial \overline{q_w}}{\partial z} - \overline{w'q_w'}\frac{\partial \overline{\theta_l}}{\partial z} - 2\varepsilon_{\theta q} + \frac{\partial}{\partial z}\left(q\ell S_{\theta q} \frac{\partial \overline{\theta_l'q_w'}}{\partial z}\right), \tag{3.5.98}$$

where q^2 is a doubled TKE value, θ_l the liquid water potential temperature, q_w the total water content, and ℓ the mixing length (see Subsection 3.5.7.4). $\overline{\phi}$ represents an ensemble-averaged ϕ, and ϕ' is the turbulent fluctuation of ϕ. The buoyancy flux, $(g/\overline{\theta_v})\overline{w'\theta_v'}$, and the dissipation terms ε_X are described in Subsections 3.5.7.2 and 3.5.7.3, respectively. Here, q_l is the mixing ratio of liquid water (including the ice phase), and

$$q^2 = \left(\overline{u'^2} + \overline{v'^2} + \overline{w'^2}\right), \tag{3.5.99}$$

$$\theta_l = \theta - \frac{L}{C_p}\frac{\theta}{T}q_l, \tag{3.5.100}$$

$$q_w = q_v + q_l, \tag{3.5.101}$$

where L denotes the latent heat of condensation and C_p specific heat at the constant pressure of dry air. The turbulent fluxes are diagnosed as

$$\overline{u'w'} = -q\ell(S_{M2.5} + S'_M)\frac{\partial \overline{u}}{\partial z}, \tag{3.5.102}$$

$$\overline{v'w'} = -q\ell(S_{M2.5} + S'_M)\frac{\partial \overline{v}}{\partial z}, \tag{3.5.103}$$

$$\overline{w'\theta_l'} = -q\ell(S_{H2.5} + S'_H)\frac{\partial \overline{\theta_l}}{\partial z}, \tag{3.5.104}$$

$$\overline{w'q'_{\mathrm{w}}} = -q\ell(S_{\mathrm{H2.5}} + S'_{\mathrm{H}})\frac{\partial \overline{q_{\mathrm{w}}}}{\partial z}, \tag{3.5.105}$$

where S_X and S'_X are non-dimensional diffusion coefficients (see Subsection 3.5.7.5).

Once the prognostic equations (3.5.95), (3.5.96), (3.5.97), and (3.5.98) are integrated, the fluxes in Eqs. (3.5.102), (3.5.103), (3.5.104), and (3.5.105) and the tendencies of the turbulent prognostic variables can be calculated.

3.5.7.2 Buoyancy Flux

Buoyancy flux $(g/\overline{\theta_{\mathrm{v}}})\overline{w'\theta'_{\mathrm{v}}}$ is a major origin of TKE production. With consideration of partial condensation effects assuming that the fluctuations of θ_{l} and q_{w} from their mean values is expressed by the Gaussian probability density function (PDF) (Sommeria and Deardorff 1977), the width of which depends on $\overline{\theta'^2_{\mathrm{l}}}$, $\overline{q'^2_{\mathrm{w}}}$ and $\overline{\theta'_{\mathrm{l}}q'_{\mathrm{w}}}$, the buoyancy flux can be written as a function of the cloud fraction (R) and the gridbox mean liquid water content ($\overline{q_{\mathrm{l}}}$) determined as moments of the PDF (see Subsection 3.5.6). Following Sommeria and Deardorff (1977) and Mellor (1977), the value is given by

$$\frac{g}{\overline{\theta_{\mathrm{v}}}}\overline{w'\theta'_{\mathrm{v}}} = \frac{g}{\overline{\theta_{\mathrm{v}}}}\left(\beta_\theta \overline{w'\theta'_{\mathrm{l}}} + \beta_q \overline{w'q'_{\mathrm{w}}}\right), \tag{3.5.106}$$

$$\beta_\theta = 1 + 0.61\overline{q_{\mathrm{w}}} - 1.61\overline{q_{\mathrm{l}}} - \widetilde{R}a_{\mathrm{l}}b_{\mathrm{l}}c_{\mathrm{l}}, \tag{3.5.107}$$

$$\beta_q = 0.61\overline{\theta} + \widetilde{R}a_{\mathrm{l}}c_{\mathrm{l}}, \tag{3.5.108}$$

$$\widetilde{R} = R - \frac{\overline{q_{\mathrm{l}}}}{2\sigma_s}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{Q_{\mathrm{l}}^2}{2}\right), \tag{3.5.109}$$

$$c_{\mathrm{l}} = (1 + 0.61\overline{q_{\mathrm{w}}} - 1.61\overline{q_{\mathrm{l}}})\frac{\overline{\theta}}{T}\frac{L}{C_p} - 1.61\overline{\theta}. \tag{3.5.110}$$

Here, a_{l}, b_{l}, σ_s and Q_{l} are given by Eqs. (3.5.86), (3.5.89), (3.5.88), and (3.5.87).

3.5.7.3 Dissipation Terms

The dissipation terms ε_X appearing in the equations are parameterized on the basis of Kolmogorov's local isotropy assumption as

$$\varepsilon = \frac{q}{B_1\ell}q^2, \quad \varepsilon_\theta = \frac{q}{B_2\ell}\overline{\theta'^2_{\mathrm{l}}}, \quad \varepsilon_q = \frac{q}{B_2\ell}\overline{q'^2_{\mathrm{w}}}, \quad \varepsilon_{\theta q} = \frac{q}{B_2\ell}\overline{\theta'_{\mathrm{l}}q'_{\mathrm{w}}}, \tag{3.5.111}$$

with the closure constants B_1 and B_2 (Nakanishi and Niino 2009).

3.5.7.4 Mixing Lengths

The formulation of mixing length ℓ is based on the original work of Nakanishi and Niino (2009), but modified partially. ℓ is defined by blending three length scales as

$$\frac{1}{\ell^2} = \frac{1}{L_{\mathrm{S}}^2} + \frac{1}{L_{\mathrm{T}}^2} + \frac{1}{L_{\mathrm{B}}^2}, \tag{3.5.112}$$

where

$$L_{\mathrm{S}} = \begin{cases} kz/3.7 & (\zeta \geq 1) \\ kz(1 + 2.7\zeta)^{-1} & (0 \leq \zeta < 1) \\ kz(1 - 10\zeta)^{0.2} & (\zeta < 0) \end{cases}, \tag{3.5.113}$$

$$L_{\mathrm{B}} = \begin{cases} q/N_{\mathrm{l}} & (\partial\theta/\partial z > 0, \zeta \geq 0) \\ \left[1 + 5w_L(q_{\mathrm{c}}/L_T N_{\mathrm{l}})^{1/2}\right] q/N_{\mathrm{l}} & (\partial\theta/\partial z > 0, \zeta < 0) \\ \infty & (\partial\theta/\partial z \leq 0) \end{cases}, \tag{3.5.114}$$

with the Brunt-Väisälä frequency N_l, the von Kármán constant k, $q_\mathrm{c} = [(g/\overline{\theta_\mathrm{v}})\overline{w'\theta_\mathrm{v}'}L_\mathrm{T}]^{1/3}$, and $\zeta = z/L_\mathrm{MO}$ with the Monin-Obukhov length L_MO. w_L is a function given by

$$w_L(z) = \frac{1 - \tanh\left(\frac{z-(z_h+\Delta z)}{\Delta z/2}\right)}{2}, \tag{3.5.115}$$

with the boundary layer height z_h, and the transition layer depth into the free atmosphere Δz. z_h is calculated following the method of Olson *et al.* (2019) and Δz is assumed to be $0.3z_h$. w_L is equal to 1 in the boundary layer ($z < z_h$), gradually decreases in the transition layer ($z_h \le z < z_h + 2\Delta z$) and becomes 0 in the free atmosphere ($z \ge z_h + 2\Delta z$). It is introduced to switch mixing length formulation smoothly between the boundary layer and the free atmosphere.

L_T is defined by blending the length scales for the boundary layer ($L_\mathrm{T(BL)}$) and the free atmosphere ($L_\mathrm{T(FA)}$) as

$$L_\mathrm{T} = w_L L_\mathrm{T(BL)} + (1 - w_L)L_\mathrm{T(FA)}, \tag{3.5.116}$$

$$L_\mathrm{T(BL)} = 0.23\frac{\displaystyle\int_0^{z_h+\Delta z} qz\,dz}{\displaystyle\int_0^{z_h+\Delta z} q\,dz}. \tag{3.5.117}$$

$L_\mathrm{T(FA)}$ is defined in sections where q^2 exceeds a threshold $q_\mathrm{thr}^2 (= 10^{-4}\ \mathrm{m^2 s^{-2}})$ in the free atmosphere ($z > z_h$). If there are N sections, $L_\mathrm{T(FA)}$ is evaluated as

$$L_\mathrm{T(FA)}(z) = \begin{cases} L^1_\mathrm{T(FA)} & (z^1_\mathrm{min} \le z \le z^1_\mathrm{max}) \\ L^2_\mathrm{T(FA)} & (z^2_\mathrm{min} \le z \le z^2_\mathrm{max}) \\ \vdots & \vdots \\ L^N_\mathrm{T(FA)} & (z^N_\mathrm{min} \le z \le z^N_\mathrm{max}) \\ 0 & (\text{elsewhere}) \end{cases}, \tag{3.5.118}$$

$$L^n_\mathrm{T(FA)} = 0.23\frac{\displaystyle\int_{z^n_\mathrm{min}}^{z^n_\mathrm{max}} q(z - z^n_\mathrm{min})\,dz}{\displaystyle\int_{z^n_\mathrm{min}}^{z^n_\mathrm{max}} q\,dz} \quad (n = 1, 2, \cdots, N), \tag{3.5.119}$$

where z^n_min and z^n_max are the lower and upper boundary of the nth section.

3.5.7.5 Nondimensional Diffusion Coefficients

$S_\mathrm{M2.5}$ and $S_\mathrm{H2.5}$ are determined using the gradient Richardson number, TKE and the empirical constants appearing in closure assumptions. S'_M and S'_H are correction terms induced by enhancement from the level 2.5 model (in which only TKE is treated as a prognostic variable) to the level 3 model. The correction terms depend on the turbulent prognostic variables (q^2, $\overline{\theta_\mathrm{l}'^2}$, $\overline{q_\mathrm{w}'^2}$ and $\overline{\theta_\mathrm{l}'q_\mathrm{w}'}$). Following Nakanishi and Niino (2009), S_q, S_{θ_l}, $S_{\theta q}$ and S_{q_w} are assumed to be proportional to $S_\mathrm{M2.5}$ as $S_q = 3S_\mathrm{M2.5}$ and $S_{\theta_\mathrm{l}} = S_{\theta q} = S_{q_\mathrm{w}} = S_\mathrm{M2.5}$. For technical details, refer to Nakanishi (2001) and Nakanishi and Niino (2004, 2006, 2009).

3.5.8 Surface Fluxes

The main procedures relating to surface processes involve the evaluation of surface fluxes. The surface scheme in the MSM employs a tiled approach in which different subgrid surface types are represented for land and sea. Turbulent fluxes are calculated for all tiles based on properties such as albedo and surface temperature, and are averaged over tiles with land fractions.

3.5.8.1 Basic Equations

Within the surface layer, it is assumed that turbulent fluxes are constant with height and equivalent to surface values. These can be expressed in terms of differences between quantities in the lowest layer of the atmosphere (u_1,v_1,θ_{v1}, and q_{v1}) and the surface (θ_{vs} and q_{vs}) as

$$\overline{u'w'} = -C_m U_a u_1 \,, \tag{3.5.120}$$

$$\overline{v'w'} = -C_m U_a v_1 \,, \tag{3.5.121}$$

$$\overline{w'\theta'_v} = -C_h U_a (\theta_{v1} - \theta_{vs}) \,, \tag{3.5.122}$$

$$\overline{w'q'_v} = -C_q U_a (q_{v1} - q_{vs}) \,, \tag{3.5.123}$$

where u and v are horizontal wind velocity components, θ_v virtual potential temperature, q_v specific humidity, and U_a wind speed near the surface. Following Kitamura and Ito (2016), U_a is expressed using turbulent kinetic energy, E ($= q^2/2$), as

$$U_a = \sqrt{u_1^2 + v_1^2 + 2CE}, \tag{3.5.124}$$

where $C = 2/3$. q^2 is offered by the boundary layer scheme (see Subsection 3.5.7). q_{vs} is represented as a saturated specific humidity (q_{sat}) at the surface temperature (T_s):

$$q_{vs} = q_{sat}(T_s). \tag{3.5.125}$$

Over the sea surface, q_{vs} is reduced due to salinity by the factor of 0.98 (Zeng *et al.* 1998).

3.5.8.2 Transfer Coefficients for Momentum and Heat

The transfer coefficients are formulated as

$$\begin{aligned} C_m(z) &= \frac{k^2}{\left[\ln\dfrac{z}{z_{0m}} - \psi_m\left(\dfrac{z}{L_{MO}}\right) + \psi_m\left(\dfrac{z_{0m}}{L_{MO}}\right)\right]^2} \\ &\equiv \frac{k^2}{\Phi_m^2(z, L_{MO})} \,, \end{aligned} \tag{3.5.126}$$

$$\begin{aligned} C_h(z) &= \frac{k^2}{\left[\ln\dfrac{z}{z_{0m}} - \psi_m\left(\dfrac{z}{L_{MO}}\right) + \psi_m\left(\dfrac{z_{0m}}{L_{MO}}\right)\right]\left[\ln\dfrac{z}{z_{0h}} - \psi_h\left(\dfrac{z}{L_{MO}}\right) + \psi_h\left(\dfrac{z_{0h}}{L_{MO}}\right)\right]} \\ &\equiv \frac{k^2}{\Phi_m(z, L_{MO})\Phi_h(z, L_{MO})} \,, \end{aligned} \tag{3.5.127}$$

where z is the height of the lowest model layer, L_{MO} the Monin-Obukhov length, z_{0m} and z_{0h} the roughness length for momentum and heat, and $k = 0.4$ (von Kármán's constant). The integrated gradient functions for momentum, ψ_m, and heat, ψ_h, are given as functions of $\zeta = z/L_{MO}$ following Beljaars and Holtslag (1991) for unstable ($\zeta < 0$) and Gryanik *et al.* (2020) for stable ($\zeta \geq 0$):

$$\psi_m(\zeta) = \begin{cases} -3\dfrac{a_m}{b_m}\left\{(1 + b_m\zeta)^{1/3} - 1\right\} & (\zeta \geq 0) \\ \dfrac{\pi}{2} - 2\tan^{-1} x + \ln\dfrac{(1+x)^2(1+x^2)}{8} & (\zeta < 0) \end{cases} , \tag{3.5.128}$$

$$\psi_{\mathrm{h}}(\zeta) = \begin{cases} -\dfrac{a_h}{b_h}\ln(1 + b_h\zeta) & (\zeta \geq 0) \\ \\ 2\ln\dfrac{1+x^2}{2} & (\zeta < 0) \end{cases}, \tag{3.5.129}$$

with $a_m = 5$, $a_h = 5$, $b_m = 0.3$, $b_h = 0.4$ and $x = (1 - 16\zeta)^{1/4}$.

Here, u_1, v_1, θ_1 and q_{v1} are volume-averaged variables as ASUCA employs the finite volume method for the spatial descretization scheme. To make consistent with this, Φ_m and Φ_h should be averaged over the lowest model layer (Nishizawa and Kitamura 2018). The mean values of Φ_m and Φ_h are given by

$$\overline{\Phi_m} = \Phi_m(\zeta_1) - \frac{1}{\zeta_1}\int_{\zeta_{0m}}^{\zeta_1} \phi_m(\zeta')d\zeta', \tag{3.5.130}$$

$$\overline{\Phi_h} = \Phi_h(\zeta_1) - \frac{1}{\zeta_1}\int_{\zeta_{0h}}^{\zeta_1} \phi_h(\zeta')d\zeta', \tag{3.5.131}$$

where $\overline{\Phi_m}$ and $\overline{\Phi_h}$ are the mean values of Φ_m and Φ_h, $\zeta_1 = z_1/L_{\mathrm{MO}}$, $\zeta_{0m} = z_{0m}/L_{\mathrm{MO}}$, $\zeta_{0h} = z_{0h}/L_{\mathrm{MO}}$, $\phi_m = 1 - \zeta(\partial\psi_m/\partial\zeta)$, $\phi_h = 1 - \zeta(\partial\psi_h/\partial\zeta)$ and z_1 is the depth of lowest model layer.

The Monin-Obukhov length is determined from the following relation:

$$Ri_{\mathrm{B}} = \frac{z}{L_{\mathrm{MO}}}\frac{\overline{\Phi_h}(z, L_{\mathrm{MO}})}{\overline{\Phi_m}^2(z, L_{\mathrm{MO}})}, \tag{3.5.132}$$

which can be solved by using an iterative approach such as the Newton's method. Ri_{B} is the Bulk Richardson Number defined by

$$Ri_{\mathrm{B}} = \frac{gz}{\frac{1}{2}(\theta_{\mathrm{v1}} + \theta_{\mathrm{vs}})}\frac{(\theta_{\mathrm{v1}} - \theta_{\mathrm{vs}})}{U_{\mathrm{a}}^2}. \tag{3.5.133}$$

3.5.8.3 Transfer Coefficients for Moisture

C_{q} over sea areas is represented with the same formula as for C_{h}, except with the roughness length for moisture ($z_{0\mathrm{q}}$). For land, evaporation from soil, vegetation and water intercepted by the surface are modeled. C_{q} over land is represented as

$$C_{\mathrm{q}} = \left\{f_{\mathrm{intc}} + (1 - f_{\mathrm{intc}})\left[f_{\mathrm{veg}}(\psi_{\mathrm{stm}} + \psi_{\mathrm{can}}) + f_{\mathrm{soil}}\beta_{\mathrm{s}}\right]\right\} C_{\mathrm{h}} \tag{3.5.134}$$

where f_{intc} is a fraction of the interception reservoir, f_{veg} is a fraction of vegetation, f_{soil} is a fraction of bare soil, ψ_{stm} and ψ_{can} are aerodynamic resistance factors, and β_{s} is evaporation efficiency from bare soil. f_{veg} is calculated using the 1km MODIS-based Maximum Green Vegetation Fraction (Broxton *et al.* 2014b). f_{soil} is given by

$$f_{\mathrm{soil}} = 1 - f_{\mathrm{veg}} - f_{\mathrm{urban}}, \tag{3.5.135}$$

where f_{urban} is a fraction of the urbanized area, and f_{urban} is obtained by combining the fractions of buildings and roads derived from the National Land Numerical Information dataset provided by Japan's Ministry of Land, Infrastructure, Transport and Tourism. Parameterization for the interception reservoir is as described in Subsection 3.5.9.

ψ_{stm}, a factor for vegetation transpiration, is given by

$$\psi_{\mathrm{stm}} = \frac{1}{r_{\mathrm{s}}}\left(C_{\mathrm{h}}U_{\mathrm{a}} + \frac{1}{r_{\mathrm{s}}}\right)^{-1}, \tag{3.5.136}$$

where r_s is stomatal resistance. r_s is a function of the Leaf Area Index (LAI), downward shortwave radiation, root zone soil water, atmospheric water vapor deficit and air temperature. Following Noilhan and Planton (1989), r_s is given by

$$r_s = \frac{r_{s,min}}{LAI} F_1 F_2^{-1} F_3^{-1} F_4^{-1} \tag{3.5.137}$$

with $r_{s,min} = 250$ s m^{-1}. LAI is prescribed from monthly climatology based on the MODIS LAI product (Myneni 2015). F_1 parameterizes the influences of photosynthetically active radiation:

$$F_1 = (1+f)\left(f + \frac{r_{s,min}}{r_{s,max}}\right)^{-1}, \quad f = \frac{0.55 S_\downarrow}{S_0}\left(\frac{2}{LAI}\right), \tag{3.5.138}$$

where $S_\downarrow$ is downward short-wave radiation at the surface, $r_{s,max} = 1200$ s m^{-1} and $S_0 = 30$ W m^{-2}. F_2 describes the availability of soil water in the root zone for transpiration:

$$F_2 = \frac{w_2 - w_{wp}}{w_{fc} - w_{wp}} \quad (0 \le F_2 \le 1). \tag{3.5.139}$$

Here, w_2, w_{fc} and w_{wp} are volumetric soil water content in the root zone, at field capacity and at the wilting point, respectively. w_{fc} and w_{wp} are estimated using the fraction of clay (Subsection 3.5.9). F_3 (the effect of the vapor pressure deficit in air) and F_4 (air temperature dependence) are

$$\begin{array}{ll} F_3 = 1 - 0.025\{e_{sat}(T_1) - e_1\} & (0 \le F_3 \le 1) \\ F_4 = 1 - 0.0016(298 - T_1)^2 & (0 \le F_4 \le 1) \end{array}, \tag{3.5.140}$$

where T_1 and e_1 are temperature and vapor pressure at the lowest model levels and $e_{sat}(T_1)$ is saturated vapor pressure at T_1.

ψ_{can} is a factor for turbulent exchange between the vegetation canopy and the underlying soil, given by

$$\psi_{can} = f_{can}\left(\frac{\beta_s g_{can}}{C_h U_a + \beta_s g_{can}}\right), \tag{3.5.141}$$

where g_{can} is in-canopy conductance and f_{can} is a canopy gap fraction. Following Essery *et al.* (2003), g_{can} is given by

$$g_{can} = \frac{C_m^{1/2} U_a}{43}, \tag{3.5.142}$$

f_{can} is a function of LAI following Zeng and Wang (2007):

$$f_{can} = 1 - V, \tag{3.5.143}$$

$$V = \frac{1 - \exp\{-k_{can} \min\langle LAI, LAI_{cr}\rangle\}}{1 - \exp\{-k_{can} LAI_{cr}\}}, \tag{3.5.144}$$

with $k_{can} = 1.5$ and $LAI_{cr} = 2$.

β_s is formulated following Lee and Pielke (1992):

$$\beta_s = \begin{cases} \frac{1}{4}\left[1 - \cos\left(\pi \frac{w_g}{w_{fc}}\right)\right]^2 & w_g < w_{fc} \\ 1 & w_g \ge w_{fc} \end{cases}. \tag{3.5.145}$$

w_g is surface volumetric soil water content (Subsection 3.5.9). If the gridbox is covered with snow, β_s is set as 1.

3.5.8.4 Roughness Length

Momentum roughness length (z_{0m}) for land depends on land use at each grid point. Thermal roughness length (z_{0h}) for land is calculated using the relational expression $\ln(z_{0\mathrm{m}}/z_{0\mathrm{h}}) = kB^{-1}$ (e.g., Garratt and Francey 1978). For urban surfaces, kB^{-1} is set as 6 following Kanda *et al.* (2005), while $kB^{-1} = 2$ is applied for other land use types. Moisture roughness length (z_{0q}) for land is assumed to be the same as z_{0h}.

Roughness lengths for sea areas are calculated using friction velocity u_* following Beljaars (1995):

$$z_{0\mathrm{m}} = a_{\mathrm{m}} \frac{\nu}{u_*} + a_{\mathrm{Ch}} \frac{u_*^2}{g} , \tag{3.5.146}$$

$$z_{0\mathrm{h}} = a_{\mathrm{h}} \frac{\nu}{u_*} , \tag{3.5.147}$$

$$z_{0\mathrm{q}} = a_{\mathrm{q}} \frac{\nu}{u_*} , \tag{3.5.148}$$

where $a_{\mathrm{m}} = 0.11$, $a_{\mathrm{Ch}} = 0.018$, $a_{\mathrm{h}} = 0.40$ and $a_{\mathrm{q}} = 0.62$, and ν is kinematic viscosity ($= 1.5 \times 10^{-5}$ m^2 s^{-1}). u_* is defined by

$$u_* = \left(\overline{u'w'}^2 + \overline{v'w'}^2 \right)^{\frac{1}{4}} . \tag{3.5.149}$$

3.5.8.5 Screen Level Diagnostics

The screen level physical quantities such as temperature and dew point at 1.5 m height and wind at 10 m height are diagnosed by interpolation between the lowest model level and surface assuming the same gradient functions as in the scheme of surface process. Wind velocity at $z_{10\mathrm{m}}$ (10 m height), $u_{10\mathrm{m}}$, and virtual potential temperature at $z_{1.5\mathrm{m}}$ (1.5 m height), $\theta_{\mathrm{v}1.5\mathrm{m}}$ are diagnosed as

$$u_{10\mathrm{m}} = \sqrt{\frac{C_{\mathrm{m}}(z)}{C_{\mathrm{m}}(z_{10\mathrm{m}})}} u_1 , \tag{3.5.150}$$

$$\theta_{\mathrm{v}1.5\mathrm{m}} = \theta_{\mathrm{vs}} + \frac{C_{\mathrm{h}}(z)}{C_{\mathrm{h}}(z_{1.5\mathrm{m}})} \sqrt{\frac{C_{\mathrm{m}}(z_{1.5\mathrm{m}})}{C_{\mathrm{m}}(z)}} (\theta_{\mathrm{v}1} - \theta_{\mathrm{vs}}) . \tag{3.5.151}$$

3.5.9 Ground Temperature and Soil Moisture

3.5.9.1 Ground Surface Temperature

Ground surface temperature, which is used in evaluating surface fluxes, is predicted by solving a surface energy balance equation given by

$$c_{\mathrm{s}} \frac{\partial T_{\mathrm{s}}}{\partial t} = (1 - \alpha) S_{\mathrm{w}\downarrow} + L_{\mathrm{w}\downarrow} - \sigma T_{\mathrm{s}}^4 - H - LE - G_{\mathrm{s}} , \tag{3.5.152}$$

where $S_{\mathrm{w}\downarrow}$ and $L_{\mathrm{w}\downarrow}$ denote the fluxes of shortwave and longwave radiation towards the surface, σ is the Stefan-Boltzmann constant, α is the surface albedo, and G_{s} is heat flux towards the ground. H and LE represent fluxes of sensible heat and latent heat from the surface:

$$H = -C_{\mathrm{p}} \rho \, \overline{w'\theta'_{\mathrm{v}}} , \tag{3.5.153}$$

$$LE = -L\rho \, \overline{w'q'_{\mathrm{v}}} , \tag{3.5.154}$$

where C_{p} is the specific heat of dry air at a constant pressure, L is the latent heat of vaporization, and ρ is the density of air near the surface.

Ground temperature (T_{g}), which is used in evaluating G_{s}, is predicted using a multi-layer model. The basic formula adopted is the heat conduction equation:

$$c_{\mathrm{g}} \frac{\partial T_{\mathrm{g}}}{\partial t} = -\frac{\partial G}{\partial z}, \quad G = -\lambda \frac{\partial T_{\mathrm{g}}}{\partial z} , \tag{3.5.155}$$

where G is ground heat flux, c_g is heat capacity and λ is thermal conductivity. The soil column is discretized into eight layers to solve the above equations, numerically. The soil temperature for the lowest layer is fixed to a climatological value for forecasts. To obtain climatological data for ground temperature, monthly mean temperatures at standard pressure levels were first calculated from objective analysis conducted in 1985 and 1986. Next, these data were interpolated vertically to the model ground surface. Then, only the annual mean and the first harmonic component of annual change in surface temperature were extracted to obtain the climatological underground temperature at the k-th ground layer with the following equation:

$$T_g = \hat{T} + A \exp\left(-\frac{z}{d}\right) \cos\left\{\frac{2\pi}{365}(D - P) - \frac{z}{d}\right\} , \tag{3.5.156}$$

where $\hat{T}$ is the mean ground surface temperature, A and P are the amplitude and the phase of the annual component of surface temperature, respectively, z_k is the depth of the k-th ground layer, d(= 2.65 m) is the e-folding depth and D is the number of days since the beginning of the year.

Sea surface temperature is also spatially interpolated from the results of SST analysis (Section 5.2). The value is given as T_s, and is predicted using an ocean mixed-layer model (Subsection 3.5.10) for the MSM.

3.5.9.2 Interception Reservoir

Rainfall and dew are intercepted by vegetation, urban surface and bare ground, with intercepted water evaporating at the potential rate. The prognostic equation for intercepted water content W is

$$\frac{dW}{dt} = P_r - E_{intc} - R_r, \tag{3.5.157}$$

where P_r is the precipitation rate, and E_{intc} and R_r are evaporation and runoff rates from intercepted water. E_{intc} is given by

$$E_{intc} = f_{intc} C_h U_a (q_{vs} - q_{v1}). \tag{3.5.158}$$

The interception reservoir fraction f_{intc} is defined as

$$f_{intc} = \left(\frac{W}{W_m}\right)^{2/3} , \tag{3.5.159}$$

where W_m is the maximum value for intercepted water. W_m is defined from LAI, f_{urban} and f_{soil} as:

$$W_m = W_{m,veg} \mathrm{LAI} f_{veg} + W_{m,urban} f_{urban} + W_{m,soil} f_{soil}. \tag{3.5.160}$$

$W_{m,veg}$, $W_{m,urban}$ and $W_{m,soil}$ are set as 0.2 kg m^{-3}.

R_r is simply defined as the excess of intercepted water, given by

$$R_r = \frac{1}{\Delta t} \max \langle W + (P_r - E_{intc})\Delta t - W_m , 0 \rangle \tag{3.5.161}$$

where Δt is the model timestep.

3.5.9.3 Soil Moisture

Soil moisture is predicted using the force-restore method based on Noilhan and Planton (1989):

$$\frac{\partial w_g}{\partial t} = \frac{C_1}{\rho_w d_1}(R_r - E_{bs}) - \frac{C_2}{\tau}(w_g - w_{geq}) , \tag{3.5.162}$$

$$\frac{\partial w_2}{\partial t} = \frac{1}{\rho_w d_2}(R_r - E_{bs} - E_{tr}) , \tag{3.5.163}$$

where w_2 is the mean volumetric water content in the root zone, τ a time constant (86400 s), E_{bs} the evaporation rate from bare soil, E_{tr} the transpiration rate, ρ_w the density of liquid water, $d_1 = 0.1$ m, $d_2 = 0.5$ m, and w_{geq} the surface volumetric water content on which gravity balances the capillarity forces. Evaporations from soil (E_{bs} and E_{tr}) are given by

$$E_{bs} = \rho(1 - f_{intc})(f_{soil}\beta_s + f_{veg}\psi_{can})C_h U_a(q_{vs} - q_{v1}) \tag{3.5.164}$$

$$E_{tr} = \rho(1 - f_{intc})f_{veg}\psi_{stm}C_h U_a(q_{vs} - q_{v1}). \tag{3.5.165}$$

C_1, C_2 and w_{geq} are defined as

$$C_1 = \begin{cases} C_{1,sat}\left(\dfrac{w_g}{w_{sat}}\right)^{-b/2-1} & w_g \geq w_{wp} \\ C_{1,max}\exp\left[-\dfrac{(w_g - w_{max})^2}{2\sigma_g^2}\right] & w_g < w_{wp} \end{cases}, \tag{3.5.166}$$

$$C_2 = C_{2,ref}\left(\frac{w_2}{w_{sat} - w_2 + 0.01}\right), \tag{3.5.167}$$

$$w_{geq} = w_2 - a w_{sat}\left(\frac{w_2}{w_{sat}}\right)^p \left[1 - \left(\frac{w_2}{w_{sat}}\right)^{8p}\right]. \tag{3.5.168}$$

When soil is very dry, C_1 implicitly represents the vapor phase transfers within soil (Braud *et al.* 1993). Following Noilhan and Mahfouf (1996), $C_{1,max}$, w_{max}, and σ_g are expressed as

$$C_{1,max} = (1.19 w_{wp} - 5.09) \times 10^{-2} T_s + 1.464 w_{wp} + 17.86, \tag{3.5.169}$$

$$w_{max} = \eta w_{wp}, \tag{3.5.170}$$

$$\eta = (-1.815 \times 10^{-2} T_s + 6.41) w_{wp} + (6.5 \times 10^{-3} T_s - 1.4), \tag{3.5.171}$$

$$\sigma_g^2 = -\frac{w_{max}^2}{2 \ln \dfrac{0.01}{C_{1,max}}}. \tag{3.5.172}$$

Hydraulic parameters w_{sat}, w_{wp}, w_{fc}, b, $C_{1,sat}$, $C_{2,ref}$, a, p are related to soil texture. Using empirical equations based on Noilhan and Lacarrére (1995), these parameters are estimated as follows:

$$w_{sat} = (-1.08 f_{sand} + 494.305) \times 10^{-3}, \tag{3.5.173}$$

$$w_{wp} = 37.1342 \times 10^{-3} (f_{clay})^{1/2}, \tag{3.5.174}$$

$$w_{fc} = 89.0467 \times 10^{-3} (f_{clay})^{0.3496}, \tag{3.5.175}$$

$$b = 0.137 f_{clay} + 3.501, \tag{3.5.176}$$

$$C_{1,sat} = (5.58 f_{clay} + 84.88) \times 10^{-2}, \tag{3.5.177}$$

$$C_{2,ref} = 13.815 (f_{clay})^{-0.954}, \tag{3.5.178}$$

$$a = 732.42 \times 10^{-3} (f_{clay})^{-0.539}, \tag{3.5.179}$$

$$p = 0.134 f_{clay} + 3.4 \tag{3.5.180}$$

f_{sand} and f_{clay}, which are sand and clay fractions in the soil, are obtained from HWSD (Harmonized World Soil Database) dataset (FAO *et al.* 2012).

3.5.10 Ocean Mixed Layer

The ocean mixed layer (OML) scheme treats temporal variations in water temperature and salinity, especially during periods of typhoon passage, to provide temporally changing lower boundary conditions of flux over the

ocean. In such conditions, near-surface wind drives vertical mixing and upwelling of sea water, resulting in sea surface cooling. Formulation of one-dimensional prognostic equations for water temperature T_{ocn}, salinity S, and ocean current V_{ocn} in the ocean is based on Price *et al.* (1986):

$$\frac{\partial T_{\mathrm{ocn}}}{\partial t} = -\frac{1}{\rho_0 C}\frac{\partial F}{\partial z} + D, \tag{3.5.181}$$

$$\frac{\partial S}{\partial t} = -\frac{\partial E}{\partial z} + D, \tag{3.5.182}$$

$$\frac{\partial V_{\mathrm{ocn}}}{\partial t} = -f \times V_{\mathrm{ocn}} - \frac{1}{\rho_0}\frac{\partial G}{\partial z}, \tag{3.5.183}$$

where F, E, and G are air-sea heat fluxes, as given from the surface scheme (Subsection 3.5.8) and radiation scheme (Subsection 3.5.6) and freshwater fluxes (calculated from evaporation and precipitation rates), and G is wind stress. D is a Rayleigh damping term, ρ_0 is the reference density of water, C is the heat capacity of water, and f is the Coriolis parameter.

The OML scheme considers three vertical mixing processes based on:

$$\frac{\partial \rho}{\partial z} \leq 0 \tag{3.5.184}$$

for static stability,

$$R_{\mathrm{b}} = \frac{g\Delta\rho h}{\rho_0(\Delta V_{\mathrm{ocn}})^2} \leq 0.30 \tag{3.5.185}$$

for mixed layer shear flow stability, and

$$R_{\mathrm{g}} = \frac{g\partial\rho/\partial z}{\rho_0(\partial V_{\mathrm{ocn}}/\partial z)^2} \leq 0.25 \tag{3.5.186}$$

for stratified shear flow stability. g is the acceleration of gravity, h is the mixed layer depth, and $\Delta()$ takes the difference between the mixed layer and the level just beneath.

First, heat fluxes except short-wave radiation and freshwater fluxes are given to the surface grid, a density profile is calculated. Short-wave radiation is absorbed within the water column with a double exponential depth dependence. When static instability (3.5.184) matches, the mixing occurs from the surface downward and the mixed layer is homogenized until static stability is achieved. Second, wind stress drives mixed-layer motion, and the mixed-layer Richardson number R_{b} is calculated. The mixed layer entrains successively deeper levels as long as (3.5.185) is satisfied. Lastly, the gradient Richardson number R_{g} is calculated for each grid, and adjacent cells are mixed as long as (3.5.186) is satisfied.

The OML scheme is applied only at sea grids. To avoid drifts of water temperature and salinity, Rayleigh damping is applied as follows:

$$D = -L(b,d)(\phi - \phi_{\mathrm{ini}}), \tag{3.5.187}$$

$$L(b,d) = min\left[max\left\{\frac{L_{\mathrm{b}}(b)}{\tau}, \frac{L_{\mathrm{d}}(d)}{\tau}\right\}, \frac{1}{dt}\right], \tag{3.5.188}$$

$$L_{\mathrm{b}}(b) = \left\{\frac{b_{\mathrm{th}}}{min(b_{\mathrm{th}}, b)}\right\}^2, \tag{3.5.189}$$

$$L_{\mathrm{d}}(d) = \left\{\frac{d_{\mathrm{th}}}{min(d_{\mathrm{th}}, d)}\right\}^2, \tag{3.5.190}$$

where L is the location-based function, ϕ_{ini} is the value of the initial condition, $\tau = 86,400$ s, $b_{\mathrm{th}} = 50$ m, $d_{\mathrm{th}} = 400$ km, b is the ocean floor, d is the distance from land grids and dt is the integration time step. The e-folding time decreases when the ocean bottom is shallower than b_{th} and the distance from land grids is

less than d_{th} to vary water temperature and salinity smoothly between grids where the OML scheme turns on and off. The vertical grid spacing is 5 m, with the deepest bottom at 100 m to save computational time and memory. Sea surface temperature (SST) is given from SST analysis (Section 5.2). The initial conditions of ocean variables are spatially and temporally interpolated from the monthly climatology of the World Ocean Atlas 2018 (WOA18, Locarnini *et al.* 2018; Zweng *et al.* 2018). To fill the discontinuity between SST and water temperature, the SST anomaly from the surface water temperature of WOA18 is added to the water temperature from surface to bottom. Ocean currents are set to zero at initial condition.

3.5.11 Parallelization

The Open Multi Processing (OpenMP) interface is employed for shared memory parallelization in the model, and the Message Passing Interface (MPI) is used for distributed memory parallelization. The model domain is split into horizontally two-dimensional sub-domains, and each decomposed sub-domain is assigned to one of the MPI processes (Aranami and Ishida 2004).

The OpenMP interface is used for parallelization inside the sub-domains. OpenMP threads are applied to loops for the y direction, and some horizontal loops (i.e., for the x and y directions) are fused to increase the loop length such that the load imbalance between threads is minimized. The z direction is used as the innermost loop at which vectorization is applied. Thus, kij-ordering is adopted for nested loops.

The sub-domains have halo regions that are exchanged with immediately adjacent MPI processes. As MPI communication and file I/O are time-consuming operations with the current supercomputer architecture, two types of overlapping are used in the model to significantly improve computational efficiency. One is overlapping of halo exchanges with the computation (Cats *et al.* 2008) to minimize the overhead of communication between MPI processes. The OpenMP interface is also used for this operation; while one thread is communicating with another MPI process, the other threads continue independent computation. The other technique involves an I/O server approach (Selwood 2012) to overlap file I/O with computation. In this method, some MPI processes are dedicated to file I/O. While computation continues, dedicated I/O processes read data from files and send them to the relevant computational processes. When output is required, the processes save the data in a dedicated buffer to invoke send operation and immediately continue computation. I/O processes receive the data and output the data to the disk.

The domain decomposition and I/O server configuration involve 88 nodes and 1056 MPI processes (including 20 I/O processes), with 8 threads/MPI for 78-hour forecasts, and 51 nodes and 816 MPI processes (including 16 I/O processes), with 6 threads/MPI for 39-hour forecasts used on Cray XC50.

3.5.12 Forecast Performance

Forecast verification is an essential process for monitoring the quality of NWP products and improving the model itself. This subsection outlines the performance of MSM precipitation forecasts with evaluation based on comparison with actually observed values.

Figures 3.5.4 and 3.5.5 show time-series representations of threat and bias scores for three-hour cumulative precipitation forecasts produced by the MSM with a 10-mm threshold from January 2015 to December 2022. Verification is performed using Radar/Raingauge-Analyzed Precipitation data (referred to here as R/A; see Subsection 4.4.1) as reference observations. The verification grid size is 20 km, meaning that forecast and observed precipitation over land or sea within 40 km from the coast is averaged over 20 km meshes. Using all verification grids, contingency tables are created for each initial time by comparing forecasts and observations, and aggregated into monthly or annual tables.

These figures indicate that the threat score improved in 2017 in which the new-generation nonhydrostatic model ASUCA was introduced to the MSM (see Subsection 3.5.1). The bias score gradually approaches to unity over the previous eight years while the threat score remains at the same level since 2017. This steady progress is attributable to the ongoing development of the forecast model and its data assimilation system with more extensive use of observation data.

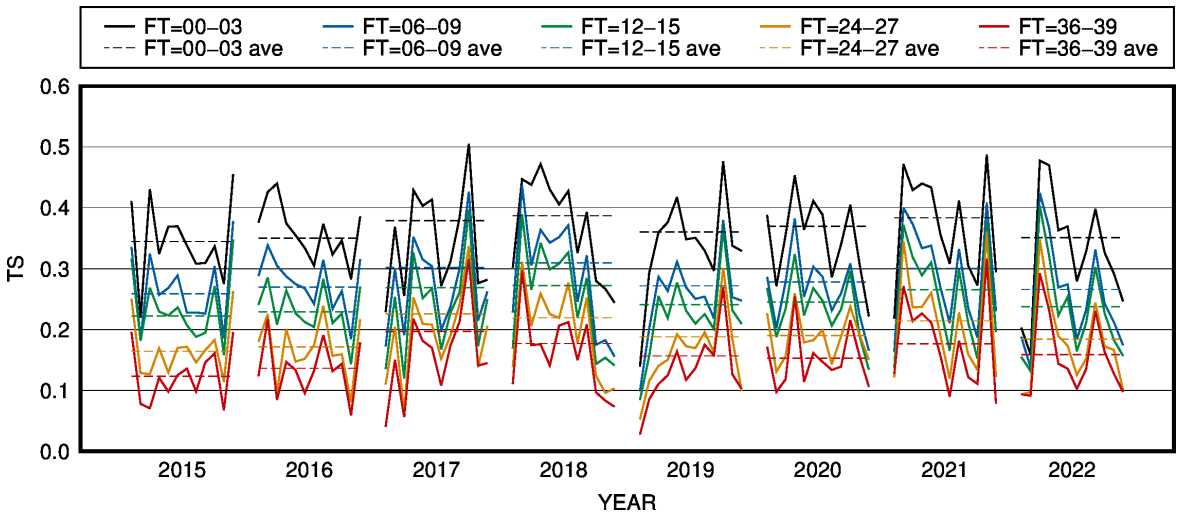


Figure 3.5.4: Monthly and annual threat scores of 3-hour cumulative precipitation at the 10 mm threshold, against the R/A within 20 km verification grids. The solid and dashed lines represent monthly and annual scores for each, FT represents the forecast range (hours). The verification period is from January 2015 to December 2022.

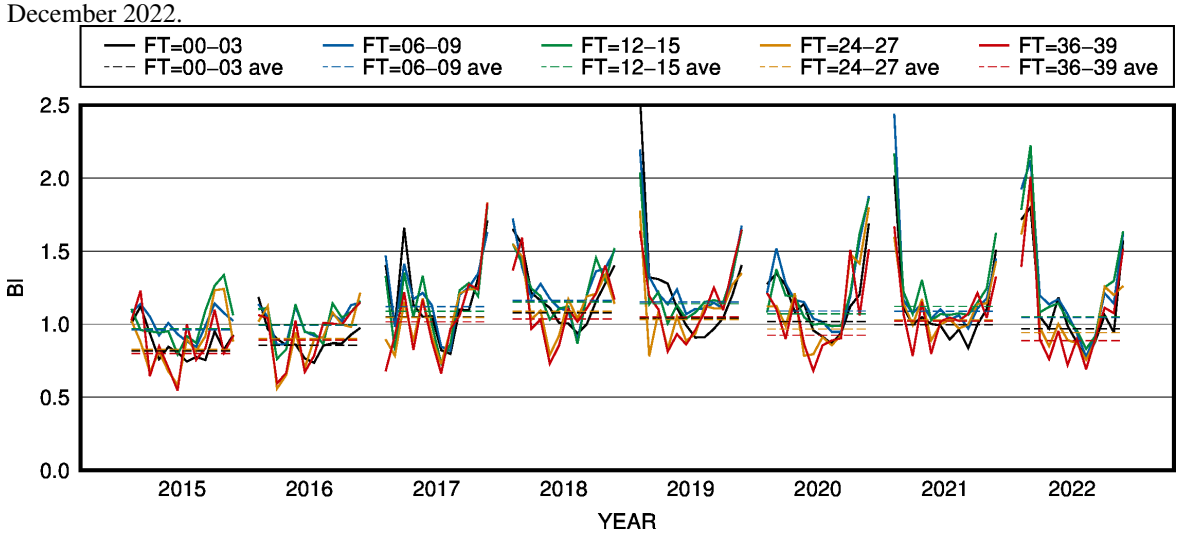


Figure 3.5.5: As per Figure 3.5.4, but for bias scores

3.6 Meso-scale Ensemble Prediction System

3.6.1 Introduction

This section describes the configuration of the regional model-based Mesoscale Ensemble Prediction System (MEPS), introduced in June 2019 based on Ono *et al.* (2021), primarily to provide uncertainty information for the MSM. Updates of initial and lateral boundary perturbations in September 2020 (Kakehata *et al.* 2021) and the introduction of a stochastic physics scheme in March 2023 (Kawada *et al.* 2023) are also outlined.

3.6.2 System Configuration

MEPS consists of 21 members, including 1 non-perturbed run. As the forecast model used in each ensemble member is identical to that of the MSM (5 km horizontal grid spacing, 96 vertical layers), non-perturbed control forecasting is identical to MSM forecasting with 39-hour forecasts at 0000, 0600, 1200 and 1800 UTC. Uncertainties in initial and lateral boundary conditions are considered, as described below, with initial perturbations (IPs) and lateral boundary perturbations (LBPs) derived from singular vectors (SVs). IPs comprise a linear combination of global SVs (GSVs) based on GSM and meso-scale SVs from the JMA-NHM(Saito *et al.* 2006), which has different spatial and temporal resolutions, and LBPs are supplied from linearly evolved GSVs. Model uncertainties are considered using the Stochastically Perturbed Parametrization Tendencies scheme (SPPT; Buizza *et al.* (1999a), Palmer *et al.* (2009)). Uncertainties in lower-boundary conditions are not considered in the current system (see Table 3.6.1 for MEPS specifications).

Table 3.6.1: MEPS specifications

System	
Title	Meso-scale Ensemble Prediction System
Implementation	27 June 2019 (Latest version: 28 March 2023)
Configuration	
Domain	Japan, Lambert projection, 817 × 661 grid points
Horizontal resolution (grid spacing)	5 km at 60 and 30°N (standard parallels)
Vertical resolution (model top)	96 stretched height hybrid levels (37.5 km)
Forecast length	39 hours (00, 06, 12, 18 UTC)
Members	1 unperturbed control forecast and 20 perturbed ensemble
Coupling to ocean / wave / sea ice models	–
Integration time step	100/3 seconds (3-stage Runge-Kutta method)
Initial conditions and perturbations	
Strategy	Singular vectors (SVs); linear combination of MSV40s, MSV80s and global SVs (GSVs)
Lateral boundary perturbations	
Strategy	Based on integration of GSV (a large-scale component of initial perturbation) using the tangent linear model
Model uncertainty perturbations	
Physics	Stochastic perturbation of physics tendency
Other model details	
	All ensemble members use exactly the same model as the MSM.

3.6.3 Initial Perturbations

3.6.3.1 Singular Vectors

The SV method computes perturbations with a large growth rate σ,

$$\sigma = \frac{\|\mathbf{M}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \tag{3.6.1}$$

where $\boldsymbol{x} \in \mathbf{R}_N$ is an N-dimensional perturbation vector and $\mathbf{M}$ is the tangent linear model (TLM) operator. Using the positive-definite operator $\mathbf{C}$ and a Euclidean inner product $(\cdot, \cdot)$, the perturbation norm is given by

$$\|\boldsymbol{x}\| = \sqrt{(\mathbf{T}\boldsymbol{x}, \mathbf{CT}\boldsymbol{x})}. \tag{3.6.2}$$

$\mathbf{T}$ is a local projection operator (LPO) that constrains the perturbation growth locations. At the initial time, $\mathbf{T}$ is usually an identity matrix.

To numerically solve (3.6.1) with a large growth rate σ, the eigenvalue problem of the matrix

$$\mathbf{A} \equiv \mathbf{C}^{-\frac{1}{2}*}\mathbf{M}^*\mathbf{T}^*\mathbf{CTMC}^{-\frac{1}{2}}, \tag{3.6.3}$$

is solved using the numerical algorithm called Lanczos algorithm. In (3.6.3), $\mathbf{M}^*$ denotes the adjoint model (ADM). The TLM/ADM integration time is called the optimization time interval (OTI).

Total energy (TE; Ehrendorfer *et al.* 1999) is used as the norm in MEPS, described as

$$\|\boldsymbol{x}\|^2 = \int_S \int_{Z_{\text{btm}}}^{Z_{\text{top}}} \frac{1}{2}\rho \left[u^2 + v^2 + w^2 + w_t \frac{C_p \theta^2}{T_r} + RT_r \left(\frac{p}{P_r} \right)^2 + w_q \frac{L^2}{C_p T_r} q^2 \right] dz dS \tag{3.6.4}$$

where ρ is density, u, v and w are zonal, meridional and vertical wind perturbations, respectively, θ, p and q are perturbations of potential temperature, pressure and the mixing ratio of water vapor, respectively, C_p is specific heat at a constant pressure, T_r = 300 K and P_r = 1,000 hPa are reference values of temperature and pressure, respectively, R is the gas constant, L is the latent-heat constant, and w_t and w_q are weight coefficients for θ and q, respectively. $\int_S$ and $\int_{Z_{\text{btm}}}^{Z_{\text{top}}}$ indicate integration extending over the horizontal target domain S and the vertical direction from Z_{btm} to Z_{top}, respectively.

3.6.3.2 Global SVs

GSM-based SVs are used as IPs and LBPs to represent large-scale uncertainties of synoptic flow in middle to upper troposphere, GSM-based SVs are used as IPs and LBPs. Global SVs (GSVs) are adopted in the global EPS (Subsection 3.3.3.2), while GSVs adopted in MEPS are calculated using a more suitable configuration for a regional EPS that targets the Japan area with a shorter OTI.

The horizontal resolution of each GSVs is TL63 (around 270 km in the mid-latitudes) with 40 vertical layers. The OTI is set as 45 hrs, which is longer than the forecast range of MEPS (39 hrs), because the operational limitation for the initial time of global analysis at the operational start time of MEPS is 6 hrs earlier than that of MEPS.

TE is also adopted as the norm for GSVs. However, in MEPS, the pressure term is not taken into account because its amplitude may be negligible in contrast to other TE terms. As moisture perturbations often cause excessive precipitation and unrealistic humidification results for land areas, the weight of the specific humidity term in TE is set as zero for GSV computations in MEPS ("Dry TE" in Table 3.6.2).

3.6.3.3 Mesoscale SVs

Mesoscale SVs (MSVs) are utilized for IPs in MEPS, with particular focus on the sensitivity of water vapor fields to severe weather. MSVs are based on a simplified version of JMA-NHM (Saito *et al.* 2006) and the TLM and ADM developed for JNoVA (Honda *et al.* 2005), along with the Lanczos algorithm.

Table 3.6.2: Configurations of MSV40, MSV80 and GSV

	MSV40	MSV80	GSV
Domain	Same as MSM (102 x 82)	Same as MSM (51 x 41)	Global
Resolution	40 km, L38	80 km, L38	TL63L40
OTI	6 hrs	15 hrs	45 hrs
norm	Moist TE	Moist TE	Dry TE
Horizontal target	125°–145° E, 25°–45° N	125°–145° E, 25°–45° N	120°–170° E, 25°–45° N
Vertical target	q: ~3 km, u, v, θ: ~5 km	q: ~3 km, u, v, θ: ~5 km	~9 km
No. of SVs	10	10	20
Use	IPs	IPs	IPs and LBPs

MSV40s computed in MEPS have a horizontal resolution of 40 km with 38 vertical layers and are designed to target meso-β- to meso-α-scale meteorological systems with a shorter OTI of 6 hrs. MSV80s with a horizontal resolution of 80 km, which can be expected to complement MSV40s in IP spatial distribution, are also employed. The OTI of each MSV80 is 15 hrs with targeting of meso-α-scale phenomena.

The TE norm is also adopted to evaluate MSV growth rates. In contrast to the norm used for GSVs, the moisture term here is explicitly considered for MSVs ("Moist TE" in Table 3.6.2). The weight of the potential temperature term w_t is 3.0, and that of the moisture term w_q is 0.6. Other terms are evaluated to an altitude of around 5 km, while vertical integration of the moisture term is limited to approximately 3 km because the desired MSVs for MEPS are related to water vapor fields in lower layers rather than to large-scale flows. Ten SVs are calculated for each MSV40 and MSV80 in the MSV target region around Japan (125–145° E and 25–45° N). Adaptive targeting is also used, with adaptive limitation of the MSV region depending on weather conditions. Here, grid points with 925 hPa vorticity values below a certain threshold are removed from the predetermined rectangular target region (Kakehata *et al.* 2021).

The specifications of these SVs are summarized in Table 3.6.2.

3.6.3.4 Linear Combination of SVs

IPs in MEPS comprise a linear combination of MSV40s, MSV80s and GSVs. Before blending, MSV80s and GSVs are interpolated to have the same resolution as MSV40s. The amplitude of each perturbation is then adjusted so that the maximum value of any one element is 5.13 m/s for wind components, 3.42 K for potential temperature and 5.13 g/kg for the mixing ratio of water vapor.

After these procedures, coefficients of the linear combination of all SVs are determined via variance minimum rotation (Yamaguchi *et al.* 2009). Ten sets of IPs are randomly chosen from 40 candidates calculated via variance minimum rotation because only this number is required for the 21 ensemble runs (including one non-perturbed run).

3.6.4 Lateral Boundary Perturbation

For a regional EPS, it is essential to consider uncertainties in lateral boundaries as well as in initial fields. It is preferable for LBPs to be consistent with the time evolution of IPs (Caron 2013). In consideration of the requirements for LBPs in a regional EPS, the linearly evolved GSVs used for IPs are adopted. Here, evolved GSVs are linearly combined using the variance minimum rotation coefficients used for IPs. This configuration enables provision of consistent and seamless perturbations during the forecast period even when their spatial scales are larger than those of the model domain.

The LBP amplitude is adjusted to reflect the uncertainty of lateral boundary conditions. In MEPS, amplitude is determined such that the average of temperature perturbations at around 500 hPa at the initial time of

the MEPS is equal to the climatological root mean square error (RMSE) of lateral boundary values against the MSM initial fields.

3.6.5 Perturbed Physics

The SPPT scheme (Buizza *et al.* (1999a), Palmer *et al.* (2009)) is used to represent model uncertainties in physical processes. The perturbed tendency of the physical process is

$$X_p = (1 + r_X)X_c, \tag{3.6.5}$$

where r_X is a Gaussian random number correlated in space-time, and X_c is the physical process tendency. The random number amplitude was set as 0.5, the correlation time as 6 hours, and the correlation length as 500 km. For computational stability, only convection and radiation scheme tendencies were perturbed.

3.7 Local Forecast Model (JMA-LFM2303)

3.7.1 Introduction

The Local Forecast Model (LFM) was launched in August 2012, along with the Local Analysis (LA) described in Section 2.7, leveraging a supercomputer upgrade implemented in June 2012. The LFM provides weather information for aviation and disaster prevention, running forecasts at an even higher resolution than the 5-km Meso-Scale Model (MSM, Section 3.5). It has 2-km horizontal grid spacing and 76 vertical layers up to a height of 21,801 m above sea level, and is designed to produce forecasts featuring greater detail with emphasis on predicting localized and short-lived extreme events.

The model focuses on providing very-short-range forecasts such as those covering the period 10 hours ahead, and allows quick and frequent updating of forecasts using initial conditions with the latest observations assimilated by LA. The operation of the LFM was started with 8 runs per day on a domain covering the eastern part of Japan (1,100 km × 1,600 km), and operation was extended in May 2013 to 24 runs per day on a domain covering Japan and its surrounding areas (3,160 km × 2,600 km).

A new-generation non-hydrostatic model known as ASUCA (Ishida *et al.* 2009, 2010, 2022; Hara *et al.* 2012) replaced the previous JMA-NHM (Saito *et al.* 2006, 2007) for LFM forecasting in January 2015 (Aranami *et al.* 2015), ahead of its implementation in the MSM in February 2017 (Subsection 3.5.1). Selected later upgrades of the ASUCA dynamical core and physics library applied to the MSM were also incorporated into the LFM in January 2017. The forecast range was extended from 9 to 10 hours in March 2019. The number of vertical layers was increased from 58 to 76 in March 2021, and the land surface model was updated to the one introduced for the MSM in March 2022. Physical process improvement included the implementation of cloud microphysics used in the MSM, and incorporation of turbulent orography form drag in March 2023, with updates of trace gas concentration and the topography data set.

3.7.2 General Configurations

The LFM provides 10-hour forecasts every hour on the hour. The forecast domain covers Japan and its surrounding areas, and has 2-km horizontal grid spacing.

The LFM employs a model identical to the MSM with similar configuration (see Section 3.5), incorporating the changes described below (see Table 3.1.3).

- Boundary conditions are obtained from MSM forecasts.
- The model is operated with a 12-second timestep.
- The main part of convection vertical transport is expected to be resolved with grid mean vertical velocity at a horizontal grid spacing of 2 km. However, this does not necessarily mean that all phenomena associated with convection can be resolved. In particular, phenomena on unresolved scales (such as small-scale

convergences and topography variances) can induce the forced lifting needed to initiate convection. Accordingly, parameterization to represent convective initiation is used in the LFM (Hara 2015), thereby mitigating delays in the onset of convection.

- The LFM incorporates the advanced version of the cloud microphysics scheme used in the MSM (Ikuta *et al.* 2021), with modifications for 2 km resolution.
- As described in Subsection 3.5.6, the cloud fraction used in the radiation process is diagnosed in consideration of fluctuations of temperature and water content from their grid mean values over each of the grids. As the fluctuations are expected to be smaller in higher resolution models, the width of the probability density function depicting the characteristics of the fluctuation was made smaller than that used in the MSM.
- Trace gas (CO2, N2O, CH4) concentration in the radiation process was updated with WMO (2021).
- The ocean mixed layer model introduced to the MSM in March 2022 (Subsection 3.5.10) is not used; sea surface temperature is fixed at an initial value.
- The GTOPO30 Subsection 3.5.1 data set was replaced with MERIT DEM (Yamazaki *et al.* 2017), a very accurate high-resolution topography data set.
- A turbulent orographic form drag scheme was introduced following Beljaars *et al.* (2004b). Standard deviations for topography were obtained from MERIT DEM.
- The domain decomposition and I/O server configuration in parallelization involve 138-nodes, 1656 MPI processes and 8 threads/MPI on Cray XC50. The domain is divided into 45 parts in the x-direction and 36 in the y-direction, and there are 36 I/O servers.

3.7.3 Forecast Performance

LFM forecast performance is evaluated in the same way as for MSM forecasts (Subsection 3.5.12). Figures 3.7.1 and 3.7.2 show time-series representations of threat and bias scores for LFM one-hour cumulative precipitation forecasts at a 10-mm threshold. It can be seen that the LFM tends to underestimate precipitation in the one-hour forecast range in summer and overestimate with lead times from two hours onward, with a spin-up period during the initial forecast stages.

To verify the skill of LFM convective precipitation forecasts, the Fractions Skill Score (FSS ; Subsection A.2.12) of one-hour cumulative precipitation from the LFM is compared with that from the MSM. Figure 3.7.3 shows FSS differences between the two models averaged over all initial times for July 2022 with a verification grid size of 5 km. The LFM is inferior for the thresholds less than 10 mm particularly in the one-hour forecast range, due to the spin-up effect from the initial field as seen in Figure 3.7.2. For spatial scales over 80 km, the LFM FSS is better than that of the MSM with thresholds over 15 mm. The forecast skill superiority of the LFM is remarkable for the latter forecast range. Figure 3.7.4 shows a Hovmöller diagram of FSS differences between the two models for a spatial scale of 80 km and a forecast range of nine hours averaged over all initial times for each month. For thresholds less than 5 mm, the difference is quite small for all seasons. The LFM FSS improves for thresholds over 10 mm, for summer after 2020. Thus, the LFM exhibits superior forecast skill for heavy rain events.

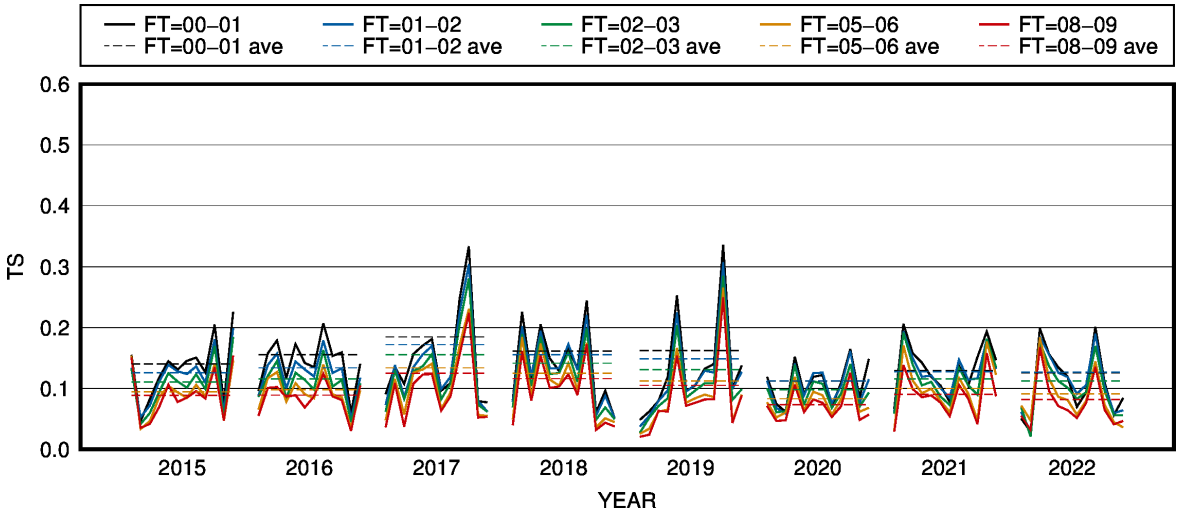


Figure 3.7.1: Monthly and annual threat scores for one-hour cumulative precipitation at the 10 mm threshold against the R/A within 10 km verification grids. The solid and dashed lines represent monthly and annual scores for each, and FT represents the forecast range (hours). The verification period is from January 2015 to December 2022.

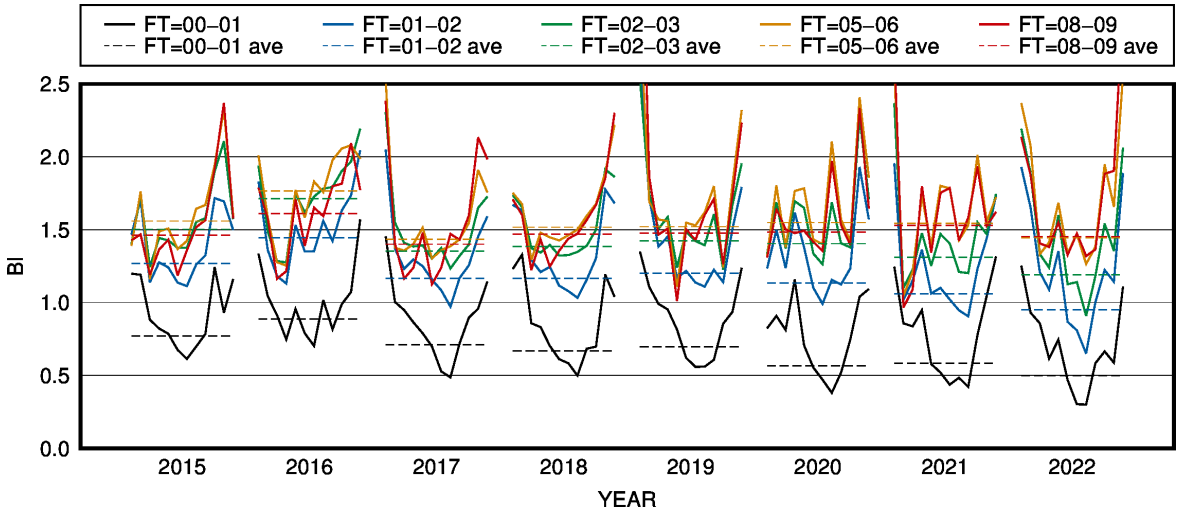


Figure 3.7.2: As per Figure 3.7.1 but for bias scores

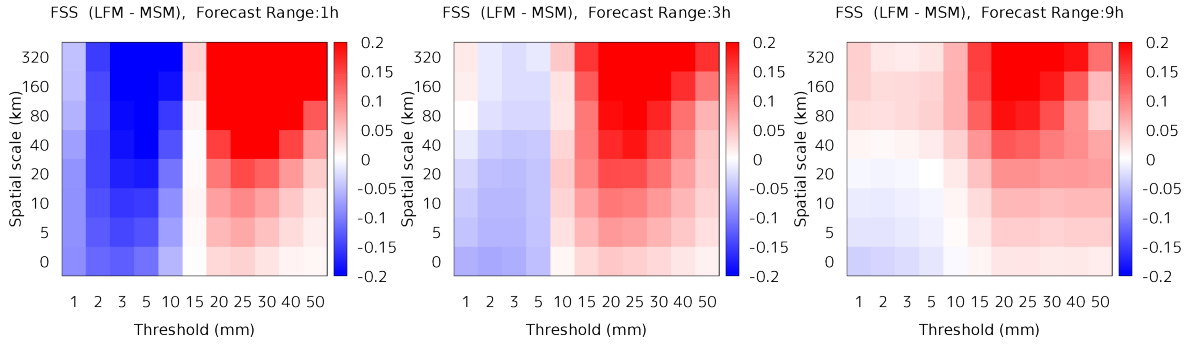


Figure 3.7.3: Monthly averaged subtraction of Fraction Skill Scores for MSM one-hour cumulative precipitation from those of the LFM. The forecast ranges are one, three and nine hours. The verification period is July 2022.

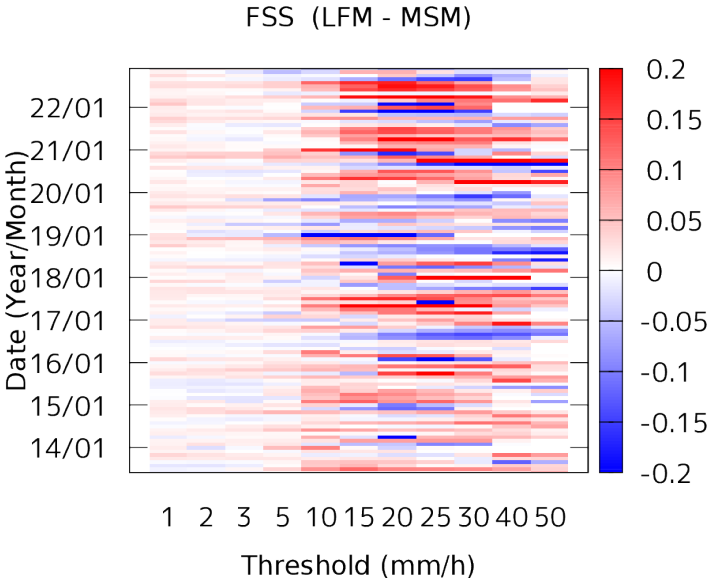


Figure 3.7.4: Monthly averaged subtraction of Fraction Skill Scores for MSM one-hour cumulative precipitation from those of the LFM. The forecast range is nine hours and the spatial scale is 80 km. The verification period is from June 2013 to December 2022.

3.8 Atmospheric Transport and Dispersion Model

3.8.1 Introduction

In July 1997, JMA was designated as a Regional Specialized Meteorological Centre (RSMC) specializing in the provision of atmospheric transport and dispersion model (ATDM) products for environmental emergency response covering Regional Association II (RA-II) of the World Meteorological Organization (WMO). RSMC Tokyo is required to provide advice on the atmospheric transport of pollutants related to nuclear facility accidents and radiological emergencies. The RSMC's ATDM products are sent to the National Meteorological Services (NMS) of WMO Member States in RA-II and to the secretariats of WMO and of the International Atomic Energy Agency (IAEA). The basic procedure of the service is defined in WMO (2019).

3.8.2 Model

3.8.2.1 Basic Model Description

The ATDM used by JMA is based on Iwasaki *et al.* (1998) with modifications developed by Kawai (2002). It involves the use of a Lagrangian approach in which tracer particles released at the temporal and spatial points of pollutant emission are displaced due to horizontal and vertical advection and diffusion and laid down through dry and wet deposition. Computation of advection, dispersion (turbulent diffusion) and deposition is based on the output of the operational numerical weather prediction (NWP) model, involving three-hourly model-level global model (GSM; see Section 3.2) outputs with temporal and spatial interpolation to tracer points. A total of 1,000,000 tracer particles are used in the operational ATDM, and time-integrated concentration and deposition are calculated using 0.5x0.5-degree latitude-longitude grids.

Horizontal velocities of tracers are estimated in accordance with Gifford (1982) as

$$\begin{aligned} u(t) &= u_{\mathrm{m}}(t) + u'(t), \\ u'(t) &= R_{\mathrm{h}} u'(t - \delta t) + \sqrt{1 - R_{\mathrm{h}}^2}\sigma G, \end{aligned} \tag{3.8.1}$$

$$\begin{aligned} v(t) &= v_{\mathrm{m}}(t) + v'(t), \\ v'(t) &= R_{\mathrm{h}} v'(t - \delta t) + \sqrt{1 - R_{\mathrm{h}}^2}\sigma G, \end{aligned} \tag{3.8.2}$$

where u and v are zonal and meridional wind speed components, and u_m and v_m are those of forecast values from the global NWP system. R_{h} is an autocorrelation of Lagrangian velocity as estimated using $e^{-\delta t/T_{\mathrm{L}}}$, where δt is the single time step length and T_{L} is the Lagrangian time scale. σ is the root mean square of horizontal velocity, which can be estimated as $(K_{\mathrm{h}}/T_{\mathrm{L}})^{1/2}$ with reference to the horizontal diffusion coefficient K_{h}. G represents random fluctuation whose statistical distributions have the Gaussian distribution function with a mean of 0 and a standard deviation of 1. The Monte Carlo method is used to determine velocities and displacements of each tracer particle. The horizontal diffusion coefficient K_{h} needs to be parameterized in consideration of the model resolution and the temporal and spatial variations of meteorological fields.

An appropriate constant value is set to reduce the computational time burden. Horizontal displacements δx and δy are given by

$$\delta x = u(t)\delta t, \tag{3.8.3}$$

$$\delta y = v(t)\delta t. \tag{3.8.4}$$

The vertical displacement δz for a single time step δt is given as

$$\delta z = w_{\mathrm{m}}(t)\delta t + \Sigma G \sqrt{2K_{\mathrm{v}}\delta t'}. \tag{3.8.5}$$

Here, w_{m} is the vertical wind speed given by the GSM forecast. The vertical diffusion coefficient depends on atmospheric vertical profiles. The time step for the integration of vertical diffusion $\delta t'$ is much shorter than

those for the integration of horizontal diffusion and advection. This shorter time step is used so that vertical displacement caused by diffusion does not exceed the thickness of the model layer. The vertical diffusion coefficient K_v is set with reference to meteorological parameters processed by the NWP model in a way analogous to the molecular diffusion coefficient estimation of Louis *et al.* (1982), and is given as follows:

$$K_v = l^2 \left| \frac{\partial c}{\partial z} \right| F(R_i), \tag{3.8.6}$$

where c is the horizontal wind speed, and the parameters l and R_i are the vertical mixing length of turbulence and the flux Richardson number, respectively. The similarity function of $F(R_i)$ is defined with reference to Louis *et al.* (1982). The mixing length is expressed as a function of the geometric height z:

$$l = \frac{\kappa z}{1 + \kappa z / l_0} \tag{3.8.7}$$

where κ is von Kármán's constant and l_0 is the maximum mixing length.

3.8.2.2 Dry and Wet Deposition

The surface tracer flux F associated with dry deposition is presented using deposition velocity $V(z_r)$ and concentration $C(z_r)$ at the reference level z_r as

$$F \equiv V(z_r) C(z_r). \tag{3.8.8}$$

For simplicity, the deposition rate is set to F/z_r following Kitada *et al.* (1986).

For wet deposition, only wash-out processes are parameterized. The wet deposition rate $\Lambda[1/\mathrm{h}]$ is approximated as a function of precipitation intensity $P[\mathrm{mm/h}]$ as predicted by the GSM with the below-cloud scavenging ratio per hour given by Kitada (1994) as

$$\Lambda \approx 0.1 P^{0.75}. \tag{3.8.9}$$

The Monte Carlo method is applied to decide which tracer particles are removed from the atmosphere at the above-mentioned dry and wet deposition rates. Noble gases such as ^{133}**Xe** are excluded from these depositing treatments.

3.8.3 Products

ATDM products are charts of 3D trajectories, time-integrated pollutant concentrations, total depositions. Sample charts are shown in Figures 3.8.1 - 3.8.7, and information on related interpretation is provided in Appendix 2.2.22 of WMO (2019).

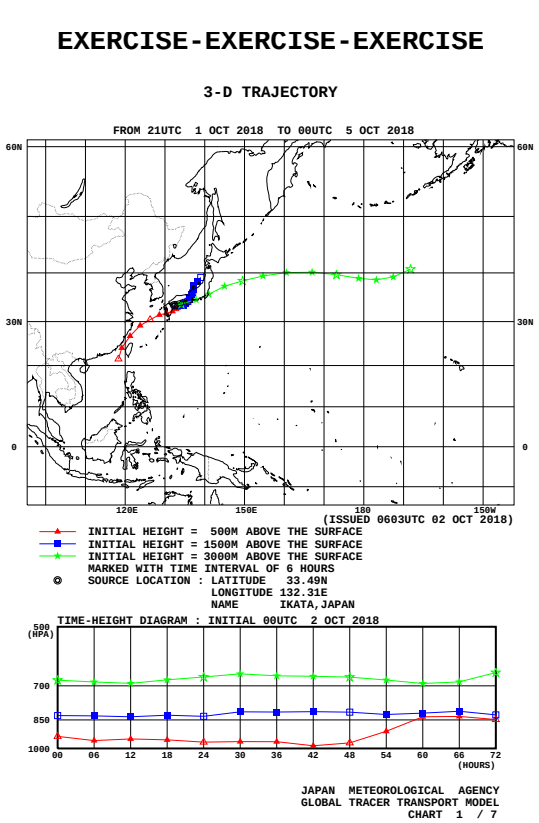


Figure 3.8.1: Example of the 3D trajectories

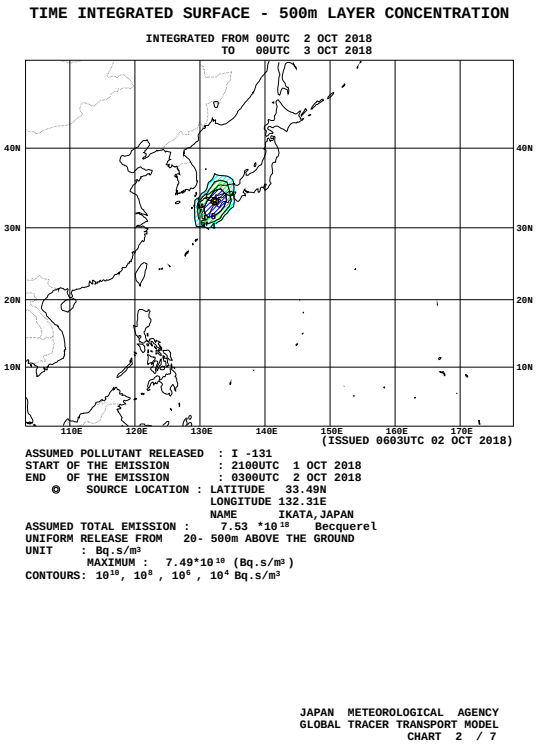


Figure 3.8.2: Example of time-integrated concentration in forecasts of up to 24 hours

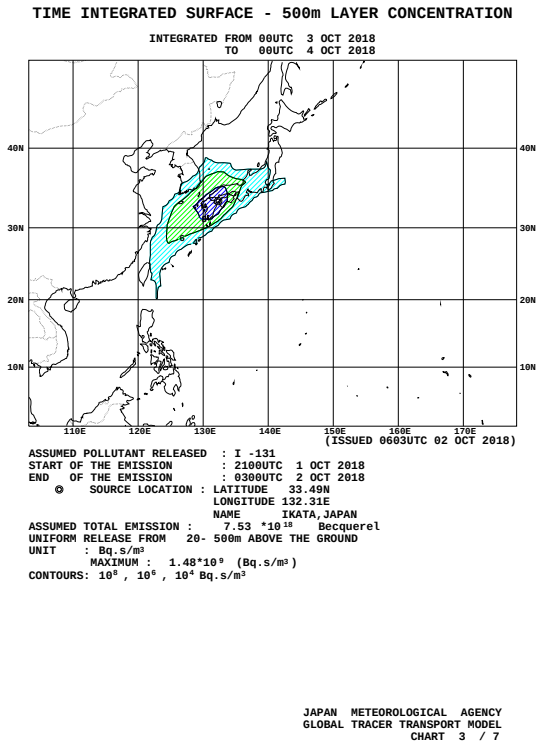


Figure 3.8.3: Example of time-integrated concentration in forecasts of up to 48 hours

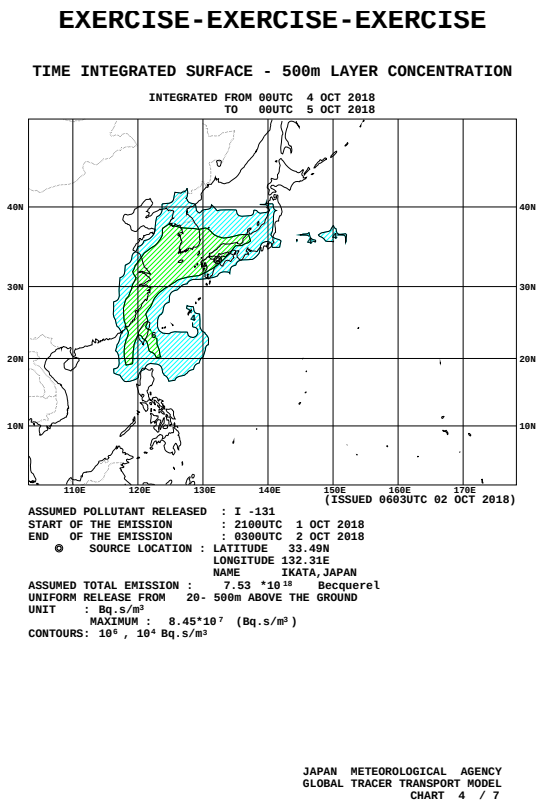


Figure 3.8.4: Example of time-integrated concentration in forecasts of up to 72 hours

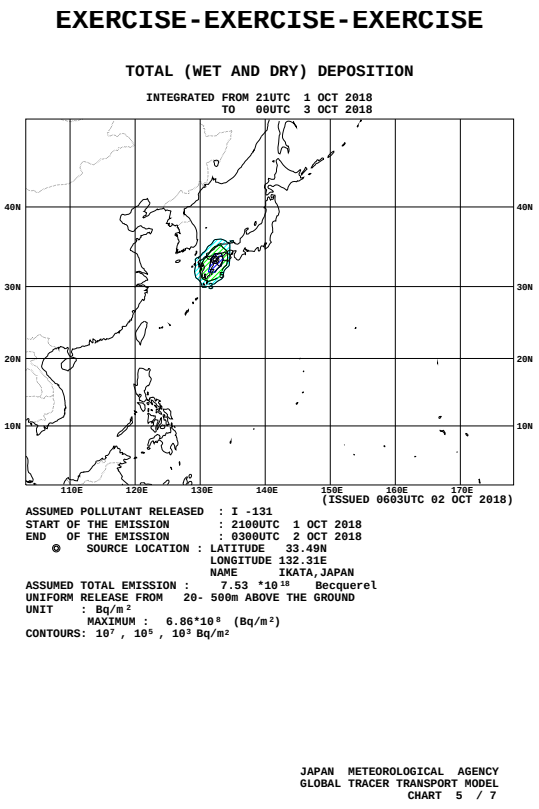


Figure 3.8.5: Example of total deposition in forecasts of up to 24 hours

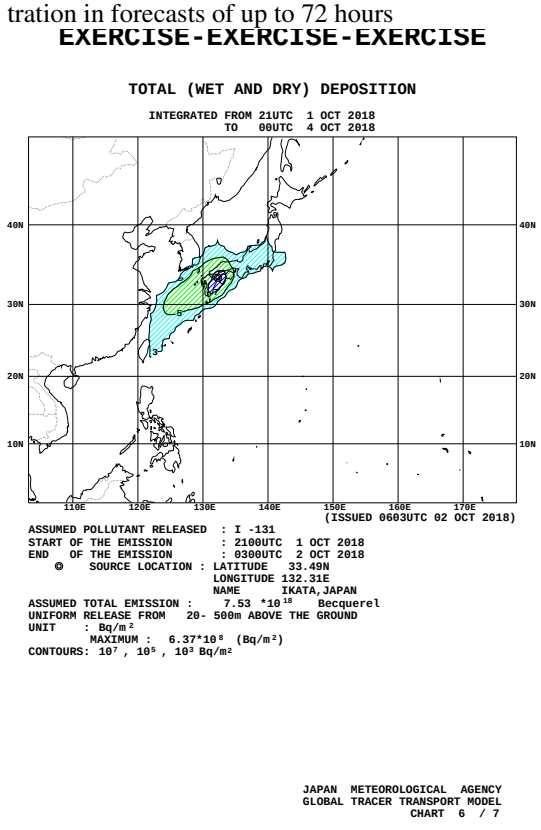


Figure 3.8.6: Example of total deposition in forecasts of up to 48 hours

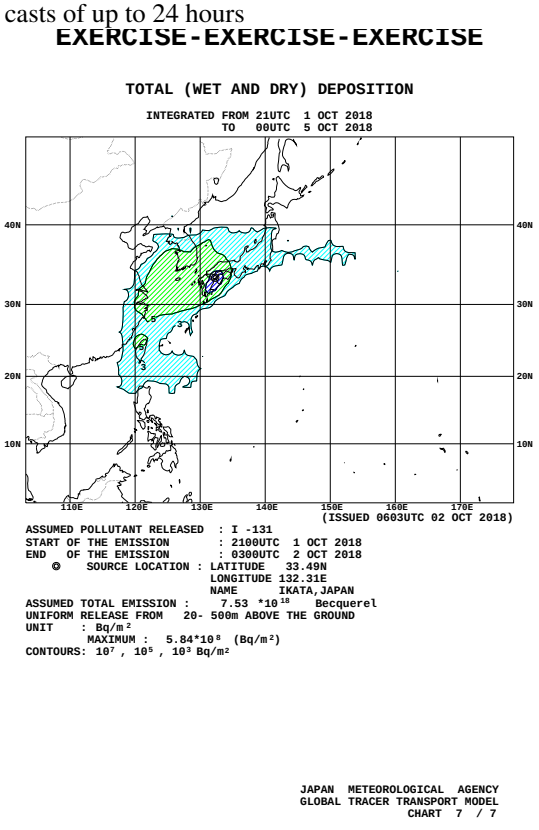


Figure 3.8.7: Example of total deposition in forecasts of up to 72 hours

3.9 Chemical Transport Model

3.9.1 Global CTM for UV Index Forecasting

3.9.1.1 Introduction

JMA provides distributions and time-series representations of UV index forecasts via its website (Figure 3.9.1 and Figure 3.9.2) based on a prediction system introduced in May 2005. UV index is calculated using a global chemical transport model (CTM) that predicts the global distribution of ozone along with a radiative transfer model. In October 2014, ozone chemistry modeling was updated with a new version of the chemistry-climate model (Deushi and Shibata 2011), which is part of the Earth System Model (MRI-ESM1; Yukimoto *et al.* 2011), and its horizontal resolution was enhanced from T42 to TL159. Nudging for total column ozone assimilation is also implemented.

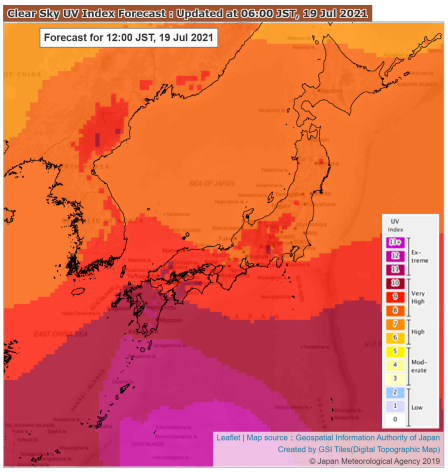


Figure 3.9.1: Clear-sky UV index forecast provided once a day (https://www.data.jma.go.jp/env/uvindex/en/).

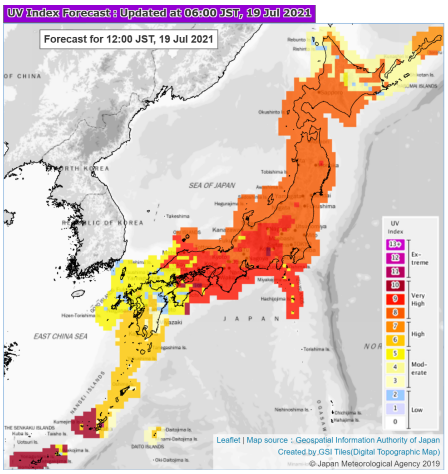


Figure 3.9.2: UV index forecast provided twice a day (https://www.data.jma.go.jp/env/uvindex/en/).

3.9.1.2 Basic Framework

The Meteorological Research Institute Chemical Climate Model version 2.1 (MRI-CCM2.1) developed by Deushi and Shibata (2011) and Yukimoto *et al.* (2019) is a global CTM used to predict distribution of atmospheric ozone and related trace gases (see Table 3.9.1 for specifications). The chemical reaction scheme is based on chemical families method, with a module incorporating 90 chemical species (64 long-lived and 26 short-lived; Table 3.9.2). In the latest module version, 266 chemical reactions are considered (191 gas-phase, 59 photolysis and 16 heterogeneous). The CTM prediction period is 120 hours and the resolution is TL159L64 (approx. 110 km horizontal-grid intervals and 64 vertical layers up to 0.01 hPa). In addition to chemical reactions, the module also treats grid scale transport with a semi-Lagrangian scheme, sub-grid scale convective transport and turbulent diffusion, dry and wet deposition and emission of trace gases from various sources.

Table 3.9.1: Global CTM specifications.

Basic equations	Eulerian model coupled with Global Spectral Model
Numerical technique	3D semi-Lagrangian transport and chemical reaction
Integration domain	Global
Grid size	TL159 (1.125°)
Vertical levels	64 (surface – 0.01 hPa)
Initial and forecast times	120 hours from 12 UTC (daily)
Boundary conditions	Similar to Global Spectral Model
Forcing data (nudging)	Global analysis (GA) and Global Spectral Model (GSM) forecasts
Observational data	Total column ozone from OMPS/NOAA

The dynamic module in the global CTM has an assimilation process in the meteorological field. Nudging (as outlined in Eq. (3.9.1)) is applied to operational global atmospheric analysis and forecasting during the CTM integration period. This starts 72 hours prior to UV index analysis and ends 48 hours after the prediction period.

3.9.1.3 Coupling of Chemical and Meteorological Parts

The chemical modules are coupled with the dynamical module with derivation of meteorological fields such as wind, temperature and precipitation as required in chemical computation (Figure 3.9.3). In the global CTM used for ozone prediction, chemical modules are directly coupled with the Atmospheric General Circulation Model (MRI-AGCM3; Mizuta *et al.* 2012) developed by the Meteorological Research Institute (MRI) of JMA as part of MRI-ESM1. Online coupling is achieved using Scup coupler software (Yoshimura and Yukimoto 2008).

CTMs generally require more computational resources than atmospheric global circulation models (AGCMs) due to the need for handling of more chemical variables and processes. Accordingly, global CTMs require lower spatial resolution in operational use, and the spatial resolution of the atmospheric model in the global CTM differs from that of the GCM used for operational meteorological analysis and forecasting (Section 3.2). There may also be differences in adopted model processes between the operational AGCM and the dynamical module of the global CTM. As a result, meteorological fields computed in the CTM are not necessarily consistent with those of operational meteorological analysis and forecasting. The nudging technique shown in Eq. (3.9.1) is often utilized optimize CTM prediction accuracy:

$$\left(\frac{\partial x}{\partial t}\right)_{nudging} = -\frac{x - x_{analysis/forecast}}{\tau} \tag{3.9.1}$$

Here, x is a meteorological variable of the CTM at a certain time t, $x_{analysis/forecast}$ is the corresponding analysis or forecast value derived from the higher-resolution operational AGCM, and τ is a relaxation time of $6 - 24$ hours. This technique enables the CTM to simulate meteorological fields realistically for the prediction period.

Table 3.9.2: Species used in the atmospheric ozone CTM.

Long-lived			
01: N_2O	02: CH_4	03: H_2O	04: NO_y
05: HNO_3	06: N_2O_5	07: Cl_y	08: O_x
09: CO	10: OClO	11: CO_2	12: Passive tracer
13: HCl	14: $ClONO_2$	15: HOCl	16: Cl_2
17: H_2O_2	18: $ClNO_2$	19: HBr	20: $BrONO_2$
21: NO_x	22: HO_2NO_2	23: ClO_x	24: BrO_x
25: Cl_2O_2	26: HOBr	27: CCl_4(CFC-10)	28: $CFCl_3$(CFC-11)
29: CF_2Cl_2(CFC-12)	30: Br_y	31: CH_3Cl	32: CH_3Br
33: CF_2ClBr (Halon1211)	34: CF_3Br (Halon1301)	35: COF_2	36: HF
37: CH_2O	38: CH_3OOH	39: C_2H_6	40: CH_3CHO
41: C_2H_5OOH	42: PAN ($CH_3C(O)OONO_2$)	43: $CH_3C(O)OOH$	44: C_3H_8
45: ACET ($CH_3C(O)CH_3$)	46: C_3H_7OOH	47: HACET ($CH_3C(O)CH_2OH$)	48: MGLY ($CH_3C(O)CHO$)
49: C_2H_4	50: GLY ALD ($HOCH_2CHO$)	51: GPAN ($HOCH_2C(O)OONO_2$)	52: GC(O)OOH ($HOCH_2C(O)OOH$)
53: C_3H_6	54: ONIT ($CH_3C(O)CH_2ONO_2$)	55: POOH (HOC_3H_6OOH)	56: C_4H_{10}
57: C_5H_8 (isoprene)	58: MACR	59: ISON	60: ISOPOOH
61: NALD	62: MACROOH	63: MPAN	64: $C_{10}H_{16}$ (terpenes)
Short-lived			
01: O(1D)	02: OH	03: Cl	04: O(3P)
05: O_3	06: HO_2	07: NO_2	08: NO
09: Br	10: N	11: ClO	12: BrO
13: NO_3	14: BrCl	15: H	16: CH_3O_2
17: $C_2H_5O_2$	18: $CH_3C(O)O_2$	19: $C_3H_7O_2$	20: $ACETO_2$ ($CH_3C(O)CH_2O_2$)
21: EO_2 ($HOC_2H_4O_2$)	22: EO (HOC_2H_4O)	23: $GC(O)O_2$ ($HOCH_2C(O)O_2$)	24: PO_2 ($HOC_3H_6O_2$)
25: $ISOPO_2$	26: $MACRO_2$		
Chemical families			
O_x = O_3 + O(3P) + O(1D)			
ClO_x = Cl + ClO			
Cl_y = ClO_x + OClO + $2Cl_2O_2$ + HCl + $ClONO_2$ + HOCl + $2Cl_2$ + $ClNO_2$ + BrCl			
NO_x = NO + NO_2 + NO_3			
NO_y = NO_x + N + HNO_3 + $2N_2O_5$ + HO_2NO_2 + $ClONO_2$ + $ClNO_2$ + $BrONO_2$ + PAN + GPAN + ONIT + ISON + NALD + MPAN			
BrO_x = Br + BrO + BrCl			
Br_y = BrO_x + HBr + HOBr + $BrONO_2$			

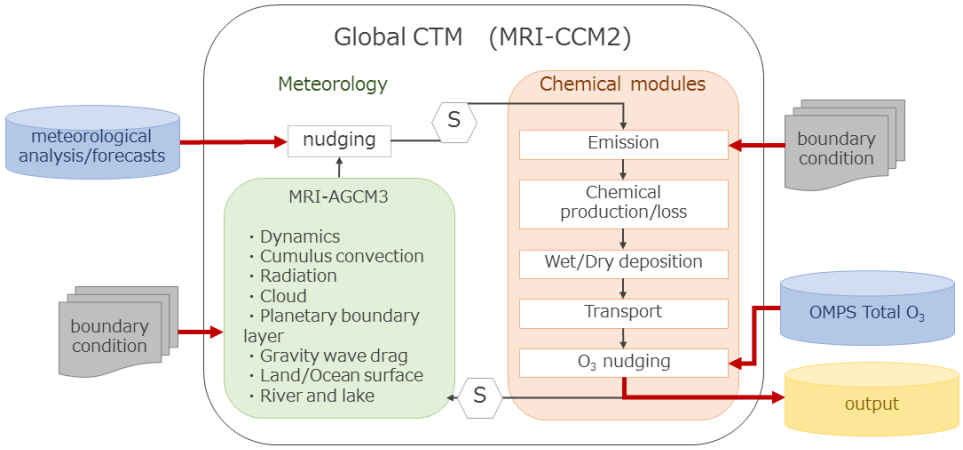


Figure 3.9.3: Global CTM ("S": Scup coupler).

3.9.1.4 Data Assimilation

The chemical module in the global CTM also has an assimilation process and total column ozone from Ozone Mapping and Profiler Suite (OMPS/NOAA-20) satellite monitoring is assimilated every hour based on

$$x_{assimilated} = x_{guess} - w(x_{guess} - x_{obs}) \tag{3.9.2}$$

Here, x is total column ozone, x_{guess} is the model guess, x_{obs} is OMPS data and $x_{assimilated}$ is the value after assimilation. The weight of the model guess with OMPS data (w in Eq. (3.9.2)) is 3/4, as determined by the ratio of the root mean square error against surface observation using a Dobson and Brewer spectrophotometer. OMPS data obtained within 72 hours prior to the UV index analysis time are assimilated into the CTM.

3.9.1.5 Verification

Figure 3.9.4 compares assimilated and observed ozone profiles for 2015. For all heights, the mean differences are within around 0.5 ppmv and the root mean square errors (RMSEs) of simulation against observation are less than 0.8 ppmv.

3.9.1.6 Radiative Transfer Model for UV Index Forecasting

The surface UV dose is calculated under clear-sky conditions using the radiative transfer model (Aoki *et al.* 2002) in an area from 122 to 149°E and from 24 to 46°N with a grid resolution of 0.25 × 0.20°. Look-up table (LUT) method is used to calculate the surface UV dose for reduced computational cost. The basic parameters of the LUT for the clear-sky UV dose are the solar zenith angle and total column ozone predicted using the CTM. The clear-sky UV index is derived from the clear-sky UV dose corrected in consideration of climatological aerosol optical depth, distance from the sun, altitude and climatological surface albedo. The UV index is derived from correction of the clear-sky UV index with cloud information from the operational weather forecast (see Table 3.9.3 for specifications).

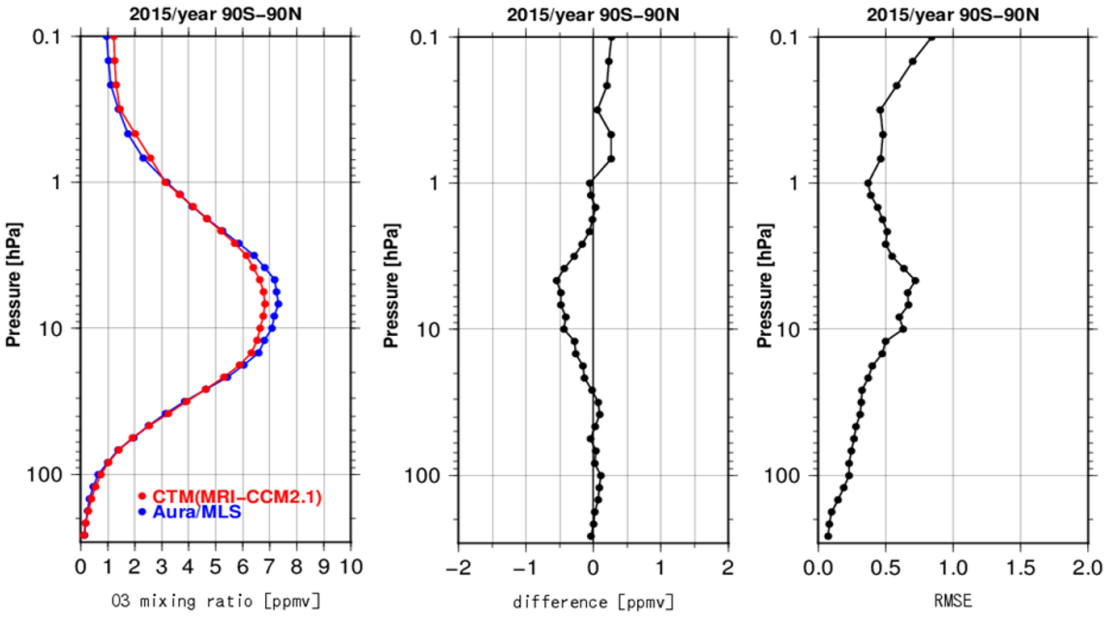


Figure 3.9.4: Annual mean vertical ozone profiles for 2015 averaged over the globe. Left: ozone mixing ratio for the CTM (red) and satellite observation (blue,Aura/MLS); middle: mixing ratio differences; right: root mean square error.

Table 3.9.3: Radiative Transfer Model specifications.

Basic equations	Radiative transfer equations for multiple scattering and absorption by atmospheric molecules and aerosols
Numerical technique	Doubling and adding
Spectral region and resolution	280 – 400 nm and 0.5 nm

The clear-sky UV index calculated using the LUT is verified against the observed UV index for clear-sky conditions (Figure 3.9.5). The mean error of the calculated clear-sky UV index is 0.1 and the RMSE is 0.4. Modeled UV indices agree closely with observation values.

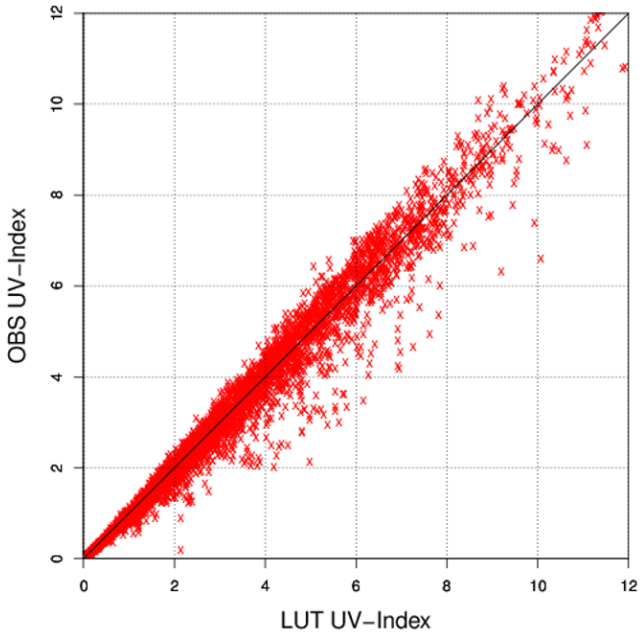


Figure 3.9.5: Relationship between calculated clear-sky UV indices and observed UV indices for clear-sky conditions at three JMA stations from 2015 to 2017.

3.9.2 Regional CTM Used for Photochemical Oxidant Information

3.9.2.1 Introduction

Based on operational predictions made using a regional CTM, JMA provides local governments with photochemical oxidant information to support the issuance of related advisories. The Agency began to utilize a global CTM (Subsection 3.9.1) for this information in August 2010, and switched to a regional CTM in March 2015. A finer-resolution local CTM nested from the regional model was also introduced for these information in March 2020. Assimilation of surface ozone concentration observation data began in March 2017 for the regional CTM, and in June 2022 for the local CTM. The calculation domain of the local CTM was expanded and the local CTM was directly nested from the global CTM in February 2023, since when the previous regional CTM has been stopped and the local CTM has been referred to by the term "new regional CTM".

3.9.2.2 Basic Framework

The regional chemical transport model for photochemical smog bulletins covering the Japan and its surrounding area was developed by Kajino *et al.* (2019, 2022) (see Table 3.9.4 for specifications). The CTM, known as ASUCA-Chem or NHM-Chem, has a horizontal resolution of 5×5 km in Lambert coordinates and a vertical resolution of 19 layers from the surface to 10 km in terrain-following coordinates. The regional CTM uses output from the operational Meso-Scale Model (MSM), known as ASUCA (Ishida *et al.* 2022), as meteorological fields to run chemical modules with offline coupling (Figure 3.9.6). In this model, 72 chemical species and 214 reactions as per SAPRC99 (Carter 2000) are treated, and major processes for atmospheric trace species such as emissions, advection, turbulent diffusion, sub-grid scale convection and dry/wet deposition are incorporated. ASUCA-Chem also implements a fully dynamic aerosol module with a three-moment bulk model using a modal-moment dynamics approach. However, in the operational version of ASUCA-Chem used for surface photochemical oxidant prediction, aerosol dynamics are fully resolved but simplified in the interests of computational efficiency.

The lateral and upper boundaries for ozone and several related species are nested from the global CTM described in Subsection 3.9.1. The natural and anthropogenic emission inventory dataset listed in Table 3.9.5 is adopted for the regional CTM. NHM-Chem source code is available subject to licensing with JMA, and further resources, including a user's manual and analysis tools, can be provided by the Meteorological Research Institute upon request.

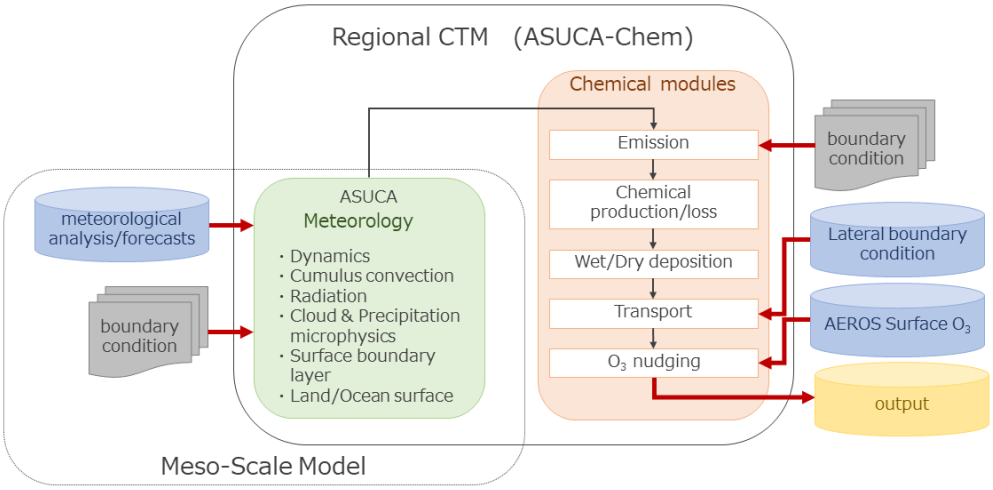


Figure 3.9.6: As per Figure 3.9.3, but for the regional CTM.

Table 3.9.4: Regional CTM specifications.

Model type	3-dimensional Eulerian chemical transport model
Area	Around Japan
Grid size	5 km
Vertical levels	19 (surface – 10 km)
Initial and forecast times	51 hours from 12 UTC (daily)
Meteorological fields	Meso-Scale Model (MSM) output
Boundary conditions	Global Chemical Transport Model (CTM)
Observational data	Surface ozone concentration from the Atmospheric Environmental Regional Observation System (AEROS)

3.9.2.3 Relaxation to Observational Data

The regional CTM is operated for a period of 51 hours with an initial time of 12 UTC. Running of the model actually starts at 19 UTC, and surface ozone concentration data (AEROS: Atmospheric Environmental Regional Observation System in Japan) for 12 – 18 UTC are assimilated via nudging similar to the right side of Eq. (3.9.1), expressed as:

$$\gamma\left(x_{obs}(s)\frac{x_{mdl}(k)}{x_{mdl}(1)} - x_{mdl}(k)\right) \tag{3.9.3}$$

Analysis of ozone concentration in the planetary boundary layer (the k-th model layer) is based on model guess concentration $x_{mdl}(k)$ and modified using surface observation $x_{obs}(s)$ with the nudging factor γ as 1×10^{-3} [s^{-1}].

3.9.2.4 Verification

Simulated surface ozone was compared with hourly in-situ observation in the calculation domain from April to September in 2022. The mean error, RMSE and correlation coefficient of one-day-ahead forecasting for spring/summer were 0.2/4.2 ppb, 13.3/14.7 ppb, and 0.69/0.71, respectively.

Table 3.9.5: Emission inventories of trace gases used in the regional CTM.

Inventory/reference			Emission source	Coverage
REAS2	Regional Emission inventory in Asia, version 2	Ohara *et al.* (2007)	Anthropogenic	East Asia
EAGrid-Japan	Asian Air Pollutant Emission Grid Database for Japan	Kannari *et al.* (2007)	Anthropogenic	Japan
GFED3	Global Fire Emission Database, version 3	Giglio *et al.* (2010)	Natural	Global
MEGAN2	Model of Emission of Gases and Aerosols from Nature, version 2	Guenther *et al.* (2006)	Natural	Global

3.10 Kosa (Aeolian Dust) Analysis Prediction Model

3.10.1 Introduction

JMA provides Kosa (Aeolian dust) forecasts via its website once a day (Figure 3.10.1) based on a prediction system introduced in January 2004. The data are provided to relevant local governments and to the Japan Meteorological Business Support Center (JMBSC), which in turn provides the data to commercial meteorological operators. Forecasts are also provided for the WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) Asia node of the China Meteorological Administration in Beijing. The model was updated to a new version based on an Earth-system model (MRI-ESM1; Yukimoto *et al.* 2011; Yukimoto *et al.* 2012) for global climate change research in November 2014, and the horizontal resolution was enhanced from TL159 to TL479 in February 2017. Dust emission flux calculation was updated to encompass the scheme of Tanaka and Chiba (2005) in November 2014. JMA's aerosol data assimilation system and Kosa analysis were introduced in January 2020.

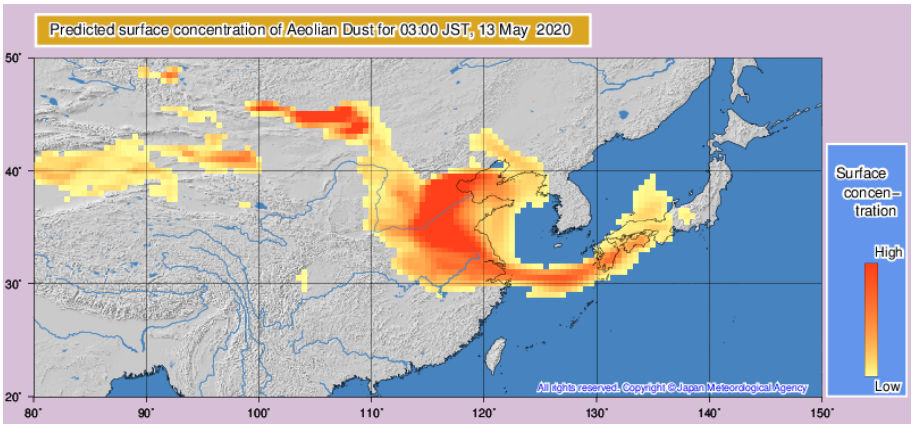


Figure 3.10.1: Prediction for surface concentration of Kosa (Aeolian dust) (https://www.data.jma.go.jp/env/kosa/fcst/en/). Data on the total amount of airborne Kosa are also provided (not shown).

3.10.2 Basic Framework

The MASINGAR (Model of Aerosol Species IN the Global AtmospheRe; Tanaka *et al.* 2003) chemical transport model used for Kosa (Aeolian dust) prediction consists of transport modules for advective transport with a semi-Lagrangian scheme and sub-grid scale eddy diffusive/convective transport, as well as other modules of surface emission, dry/wet deposition and chemical reactions (see Table 3.10.1 for specifications). It treats sulfate, black carbon, organic carbon, mineral dust and sea salt as major tropospheric aerosol species with the assumption of external mixing. The chemical module is directly coupled with MRI-AGCM3 using Scup coupler software as per the CTM for UV index forecasting (Subsection 3.9.1.3) as shown in Figure 3.10.2.

Assimilation for the meteorological field (U, V, T) involves Newtonian nudging, which is essentially the same as for the CTM for UV index forecasting (Eq. (3.9.1)). The prediction period is 96 hours, and the spatial resolution is TL479L40 (horizontal grid interval: approx. 40 km; 40 vertical layers up to 0.4 hPa in sigma-pressure hybrid coordinates).

The emission flux of mineral dust aerosol depends on meteorological, geographical and soil-surface conditions such as wind speed, land use, vegetation type, soil moisture and soil type. The emission flux F of dust (soil particles with diameter D) is expressed in proportion to the saltation flux Q:

$$F(D) = \alpha(D, d_s)Q(d_s) \tag{3.10.1}$$

Here, d_s is the diameter of saltation particles, and the proportional coefficient α depends on both D and d_s. The saltation flux Q is zero when the friction velocity on a bare surface u_* is lower than the following threshold velocity:

$$u_{*t}(d_s) = f_w \sqrt{A_N \left(\frac{\rho g d_s}{\rho_a} + \frac{\Gamma}{\rho_a d_s} \right)} \tag{3.10.2}$$

Here, A_N and Γ are empirical constants, ρ is soil particle density, ρ_a is air density, g is gravitational acceleration (Shao and Lu 2000) and f_w is a factor depending on soil moisture (Fećan *et al.* 1998). When u_* is greater than u_{*t}, the saltation flux is expressed as

$$Q(d_s) = \frac{c(d_s)\rho_a u_*^3}{g} \left(1 - \frac{u_{*t}(d_s)^2}{u_*^2} \right) \tag{3.10.3}$$

Here, c is a coefficient depending on d_s (Shao and Leslie 1997).

Table 3.10.1: Kosa (Aeolian dust) analysis and prediction model specifications.

Basic equations	Eulerian model coupled with Global Spectral Model
Numerical technique	3D semi-Lagrangian transport and dust emission calculation from surface meteorology
Integration domain	Global
Grid size	TL479 (0.375°)
Vertical levels	40 (surface – 0.4 hPa)
Initial and forecast times	96 hours from 12 UTC (daily)
Boundary conditions	Similar to Global Spectral Model
Forcing data (nudging)	Global Analysis (GA) and Global Spectral Model (GSM) forecasts
Data assimilation	2-dimensional variation (2D-Var)
Observational data for assimilation	Aerosol optical depth at 500 nm from AHI

3.10.3 Data Assimilation

JMA began 2D-Var aerosol data assimilation (Yumimoto *et al.* 2017) against aerosol optical depth (AOD) with the Himawari-8/9 geostationary satellite in January 2020. AOD data are obtained using the JAXA/EORC retrieval algorithm (Yoshida *et al.* 2018, 2021; Kikuchi *et al.* 2018) with a one-day analysis cycle to calculate initial conditions for five-day forecasting. Assimilation is performed using Himawari-8/9 AOD data with a ±1 hour target time at 06 UTC in the analysis cycle. AOD observation error is approximated as

$$\sigma_{obs} = \max(\sigma_{H08/09}, \alpha + \beta * \tau_{H08/09}) \tag{3.10.4}$$

Here, σ_{obs} is the AOD observation error, $\sigma_{H08/09}$ is the Himawari-8/9 AOD retrieval error, and α and β are coefficients estimated from statistical calculation of Himawari-8/9 AOD observation error against data from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua satellite. These estimated values are used for observation error covariance.

AOD background error is calculated using five initial forecasting data points within ±6 hours of the target assimilation time (Yumimoto *et al.* 2018). As output is every three hours, an ensemble of 25 forecast data points is used for background error covariance.

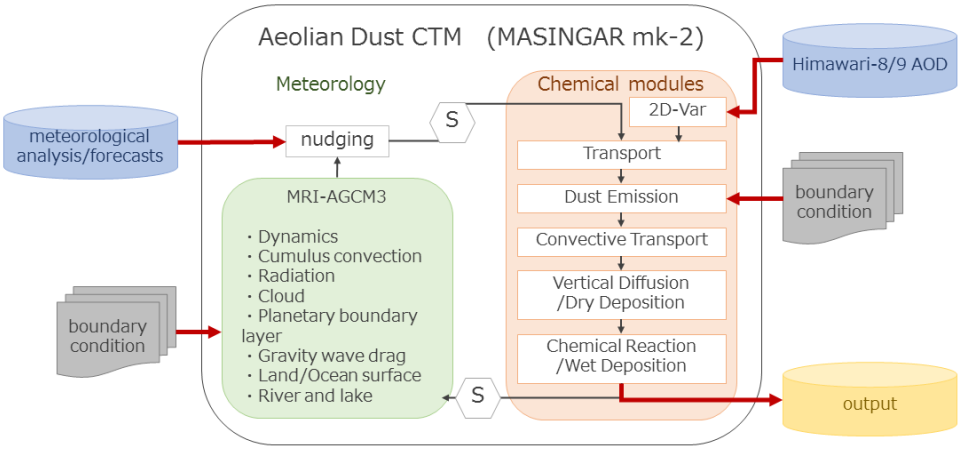


Figure 3.10.2: Kosa (Aeolian dust) CTM ("S": Scup coupler).

3.10.4 Verification

Predictions from the operational Kosa CTM are verified against surface synoptic observations (reported present weather code (ww)). The model score is calculated using categorical verification as outlined in Appendix A (Table 3.10.2). The threshold surface dust concentration value for Kosa prediction is currently 90 μg/m^3, and the threat score for Kosa prediction with/without 2D-Var after 24 hours in the Japan area is 0.28/0.26 averaged over spring (2016 – 2018).

Table 3.10.2: Verification indices for Kosa (Aeolian dust) CTM categorical prediction (*ww*: present weather code; ρ_s^{dust}: predicted surface dust concentration).

	Observed (ww = 06 – 09, 30 – 35, 98)	Not Observed
Forecast ($\rho_s^{dust} \geq 90\ \mu$g/m^3)	FO	FX
Not Forecast ($\rho_s^{dust} < 90\ \mu$g/m^3)	XO	XX

3.11 Verification

JMA verifies the output of Numerical Weather Prediction (NWP) model forecasts against observation and/or analysis outcomes, and the results are used as reference in research and development regarding NWP models. GSM and Global EPS verification results are exchanged between JMA and other NWP centers via the Lead Centre for Deterministic NWP Verification (LC-DNV) and the Lead Centre for EPS Verification. The standard verification procedures are defined in the Manual on the Global Data-processing and Forecasting System (WMO 2019). This section summarizes operational verification for the GSM and the Global EPS.

The specifications of GSM verification against analysis values are shown in Table 3.11.1, and the scores used for verification are presented in Appendix A.1. GSM forecast performance, including typhoon forecasting, is described in Subsection 3.2.14.

Radiosonde data and synoptic surface observation (SYNOP) data are used for verification against observation values. The specifications of verification against radiosonde data for the GSM are shown in Table 3.11.2. All radiosonde data passing quality control are used in verification. Stations from which radiosonde data are used in verification are selected on the basis of recommendations from the Commission for Basic Systems. The specifications of verification against SYNOP data for the GSM are shown in Table 3.11.3.

The specifications of verification regarding Global EPS output for One-week Forecasting are shown in Table 3.11.4. The probabilistic forecast for verification is defined as the ratio of the number of ensemble members in an event to the ensemble size for every grid. The verification results for the Global EPS are described in Subsection 3.3.6.

Table 3.11.1: Operational verification against analysis for the GSM

Verification grid	$1.5° \times 1.5°$ (latitude/longitude)	
Elements	Extra-tropics:	Geopotential height (Z) and temperature (T) at 850, 500, 250 and 100 hPa; Wind at 925, 850, 700, 500, 250 and 100 hPa; Relative humidity (RH) at 850 and 700 hPa; Mean sea-level pressure (MSLP)
	Tropics:	Z, T and wind at 850 and 250 hPa; RH at 850 and 700 hPa
Scores	Wind:	Root mean square vector wind error and mean error of wind speed
	Other elements:	Mean error, root mean square error, anomaly correlation, S1 score (only for MSLP), mean absolute error, root mean square forecast and analysis anomalies, standard deviations of forecast and analysis fields
Forecast range (initial time)	Up to 264 hours (00 and 12 UTC)	
Forecast steps	Every 12 hours	
Areas	NH extra-tropics (90°N – 20°N), SH extra-tropics (20°S – 90°S), Tropics (20°N – 20°S), North America (25°N – 60°N, 50°W – 145°W), Europe/North Africa (25°N – 70°N, 10°W – 28°E), Asia (25°N – 65°N, 60°E – 145°E), Australia/New Zealand (10°S – 55°S, 90°E – 180°E), Northern polar region (90°N – 60°N) and Southern polar region (90°S – 60°S)	

Table 3.11.2: Operational verification against radiosonde data for the GSM

Verification grid	Nearest model grid point to the observation location	
Elements	Extra-tropics:	Z and T at 850, 500, 250 and 100 hPa; Wind at 925, 850, 700, 500, 250 and 100 hPa; RH at 850 and 700 hPa
	Tropics:	Z, T and wind at 850 and 250 hPa; RH at 850 and 700 hPa
Scores	Wind:	Root mean square vector wind error and mean error of wind speed
	Other elements:	Mean error, mean absolute error and root mean square error
Forecast range (initial time)	Up to 264 hours (00 and 12 UTC)	
Forecast steps	Every 12 hours	
Areas	NH extra-tropics, SH extra-tropics, Tropics, North America, Europe/North Africa, Asia, Australia/New Zealand, Northern polar region and Southern polar region	

Table 3.11.3: Operational verification against SYNOP data for the GSM

Verification grid	Nearest model grid point to the observation location
Elements	T at 2 m, wind speed at 10 m, wind direction at 10 m and 24-hour precipitation
Scores	T at 2 m, wind speed at 10 m and wind direction at 10 m: Mean error, mean absolute error and root mean square error Contingency tables for following thresholds wind speed at 10 m: 5, 10 and 15 m/s 24-hour precipitation: 1, 10 and 50 mm
Forecast range (initial time)	Up to 264 hours (00 and 12 UTC)
Forecast steps	24-hour precipitation: Every 24 hours Other elements: Every 6 hours up to 72-hour forecast; Every 12 hours up to end of the forecast

Table 3.11.4: Operational verification of the Global EPS for One-week Forecasting

	Deterministic verification	Probabilistic verification	
Analysis	Global analysis on $1.5^\circ \times 1.5^\circ$ grid		
Forecast	Ensemble mean	Probability	
Climatology	Climatological fields and standard deviations are calculated from common climatology provided from LC-DNV. The climatological probability is given by the monthly frequency derived from analysis fields.		
Elements	Z at 500 hPa; T at 850 hPa; u wind component (U) and v wind component (V) at 850 and 250 hPa; MSLP	Anomalies of Z at 500 hPa, T at 2 m, T at 850 hPa and MSLP with thresholds of ±1, ±1.5 and ±2 climatological standard deviation; Wind speed at 850 hPa with thresholds of 10, 15 and 25 m/s; U and V at 850 and 250 hPa with thresholds of 10th, 25th, 75th and 90th percentile points with respect to the defined climatology; 10 m wind speed with thresholds 10 and 15 m/s; 24-hour precipitation with thresholds 1, 5, 10 *and* 25 mm/24 hours	Z at 500 hPa; T at 2 m and 850 hPa; U and V at 850 and 250 hPa; Wind speed at 10 m and 850 hPa; MSLP; 24-hour precipitation
Scores	Root mean square error, and anomaly correlation	Reliability table	Continuous ranked probability score
Forecast range (initial time)	Up to 264 hours (00 and 12 UTC)		
Forecast steps	Every 24 hours		
Areas	NH extra-tropics, SH extra-tropics and Tropics		

3.12 Atmospheric Transport Model for Volcanic Ash

3.12.1 Introduction

Since April 1997, JMA has provided information on volcanic ash clouds to airlines, civil aviation authorities and related organizations in its role as the Volcanic Ash Advisory Centre (VAAC) Tokyo. JMA introduced the Global Atmospheric Transport Model (JMA-GATM) in December 2013 to create Volcanic Ash Advisories (VAAs). JMA also launched its Volcanic Ash Fall Forecast (VAFF) product based on the Regional Atmospheric Transport Model (JMA-RATM) in March 2008 (Shimbori *et al.* 2009) and updated it in spring 2015 (Hasegawa *et al.* 2015).

In March 2021, the new Atmospheric Transport Model (JMA-ATM; Shimbori and Ishii 2021) has been developed by the Meteorological Research Institute to unify the JMA-GATM and JMA-RATM, and then implemented in the JMA's supercomputer system connected to the Volcanic ash Advisory and ash fall Forecast distribution System (VAFS).

3.12.2 Initial Condition

3.12.2.1 Emission Source Parameter Model

In the case of creating VAFFs, a volcanic plume in the shape of an inverted cone is adopted as the initial condition of the JMA-ATM. The initial plume is based on information from observational reports, including eruption time and plume height, and on the duration of volcanic ash emission. For volcanic ash grain size, log-normal distribution is adopted. The probability density function of the diameter D is given by

$$f(D) = \frac{1}{\sqrt{2\pi\sigma_D^2}} \exp\left[-\frac{\log^2(D/D_\mathrm{m})}{2\sigma_D^2}\right] \tag{3.12.1}$$

Here, D_m is the mean diameter and σ_D is the standard deviation of distribution. The number of tracers, diameter the range and these parameters D_m, σ_D are shown in Table 3.12.1.

The vertical distribution of volcanic particles in the eruption plume is calculated according to Suzuki (1983). The probability that a tracer with diameter D is released from a height of z above vent level is given as follows:

$$P(D, z) = AY(D, z)e^{-Y(D,z)} \tag{3.12.2}$$

Here, A is the normalization constant and $Y(D, z)$ is expressed as

$$Y(D, z) = \beta \frac{W(z) - w_\mathrm{t}(D, 0)}{w_\mathrm{t}(D, 0)} \tag{3.12.3}$$

with release constant β (set as 0.017). $w_\mathrm{t}(D, 0)$ is the terminal velocity at the height of the volcano summit as calculated using Eq. (3.12.20). $W(z)$ is the vertical velocity in the eruption column at height z, which can be represented as

$$W(z) = W_0\left(1 - \frac{z}{H}\right) \tag{3.12.4}$$

where $W_0 = \sqrt{H/H_0}\,[\mathrm{m\,s^{-1}}]$ is initial velocity, H is column height and $H_0 = 0.22\,\mathrm{m}$. The total amount of volcanic ejecta M is given as (e.g. Koyaguchi 2008)

$$M = K_M H^4 T_M \tag{3.12.5}$$

where T_M is the duration of eruption and K_M is a constant set to $193\,\mathrm{kg\,km^{-4}s^{-1}}$.

3.12.2.2 Initialization for Volcanic Ash Clouds

In the case of creating VAAs, the boundary of volcanic ash observed by a meteorological satellite is adopted as the initial condition of the JMA-ATM. While the ash cloud-top height can be obtained from satellite imagery,

cloud-base height is estimated using vertical wind shear. To import the shape of the observed volcanic ash cloud, an initialization method for the ash clouds has been implemented in the operational system (Ishii *et al.* 2021). This method has a low computational cost, and therefore, is suitable for creating initial conditions immediately after an eruption occurs. The specifications of tracer particles for VAAC operation are also shown in Table 3.12.1.

Table 3.12.1: JMA-ATM specifications for VAAs and VAFFs

	VAA	VAFF (Scheduled)	VAFF (Preliminary)	VAFF (Detailed)
Number of Tracer Particles	40,000	100,000		250,000
Diameter of Tracer Particles	1 – 10 μm (for ash cloud) 10 – 100 μm (for plume)	0.65 μm – 96 mm		
Mean Diameter	7.81 μm (for ash cloud) 31.6 μm (for plume)	0.25 mm		
Standard Deviation	0.5 (for ash cloud) 1.0 (for plume)	1.0		
Forecast Time	18 hours	18 hours	1 hour	6 hours
Time Step	600 seconds (Euler)	200 seconds (4-stage Runge-Kutta)		

3.12.3 Model

3.12.3.1 Basic Framework

The JMA-ATM adopts an offline Lagrangian scheme, the time tendency of tracer variables is calculated in each process and integrated at the last time step in order that dynamical and physical processes are commutative at each time step. The location (longitude λ, latitude φ, altitude z) of each tracer after the time step Δt is given by

$$\lambda(t+\Delta t) = \lambda(t) + \left(f_\lambda^{\mathrm{hadv}} + f_\lambda^{\mathrm{hdif}}\right)\Delta t \tag{3.12.6}$$

$$\varphi(t+\Delta t) = \varphi(t) + \left(f_\varphi^{\mathrm{hadv}} + f_\varphi^{\mathrm{hdif}}\right)\Delta t \tag{3.12.7}$$

$$z(t+\Delta t) = z(t) + \left(f_z^{\mathrm{vadv}} + f_z^{\mathrm{vdif}} + f_z^{\mathrm{fall}} + f_z^{\mathrm{ddep}} + \sum_\alpha f_z^{\mathrm{wout}(\alpha)}\right)\Delta t \tag{3.12.8}$$

Here, f represents tendency of each physical process as shown in the following Subsections 3.12.3.2 to 3.12.3.5.

3.12.3.2 Horizontal and Vertical Advection

In the JMA-ATM, tracers are transported by mean wind velocity $\left(\overline{U}, \overline{V}, \overline{W}\right)$ at each tracer location as derived by interpolating forecast gridded data from the GSM (see Section 3.2), MSM (see Section 3.5) or LFM (see Section 3.7). The time tendencies of transport process are given by

$$f_\lambda^{\mathrm{hadv}} = \frac{\overline{U}}{R_\mathrm{E}\cos\varphi} \quad (\varphi \neq \pm 90^\circ) \tag{3.12.9}$$

$$f_\varphi^{\mathrm{hadv}} = \frac{\overline{V}}{R_\mathrm{E}} \tag{3.12.10}$$

$$f_z^{\mathrm{vadv}} = \overline{W} \tag{3.12.11}$$

where R_E is the radius of earth.

3.12.3.3 Horizontal and Vertical Diffusion

The effects of horizontal dispersion are represented using horizontal wind perturbation $(U'(t),\ V'(t))$ from mean wind $\left(\overline{U}(t),\ \overline{V}(t)\right)$. $(U'(t),\ V'(t))$ are given by

$$U'(t) = R_{U'}(\Delta t)U'(t-\Delta t) + \sqrt{1-R_{U'}(\Delta t)^2}\sigma_{U'}\Gamma \tag{3.12.12}$$

$$V'(t) = R_{V'}(\Delta t)V'(t-\Delta t) + \sqrt{1-R_{V'}(\Delta t)^2}\sigma_{V'}\Gamma \tag{3.12.13}$$

Here, Γ represents a random number which obeys Gaussian distribution with a mean of 0 and a standard deviation of 1. The autocorrelation coefficients of Lagrangian velocity $(R_{U'},\ R_{V'})$ are given by

$$R_{U'}(\Delta t) = R_{V'}(\Delta t) = e^{-\Delta t/t_{\mathrm{Lh}}} \tag{3.12.14}$$

using the Lagrangian time scale t_{Lh}, and their standard deviations $(\sigma_{U'},\ \sigma_{V'})$ are

$$\sigma_{U'} = \sigma_{V'} = \sqrt{\frac{K_{\mathrm{h}}}{t_{\mathrm{Lh}}}} \tag{3.12.15}$$

The parameters are set to $K_{\mathrm{h}} = 5.864{\times}10^3\ \mathrm{m^2 s^{-1}}$ (for VAAs) or $5.864{\times}10^4\ \mathrm{m^2 s^{-1}}$ (for VAFFs), $t_{\mathrm{Lh}} = 5.0{\times}10^4$ s, and the initial condition of horizontal wind perturbation is $U'(0) = V'(0) = 0.253\Gamma\ \mathrm{m\,s^{-1}}$ following Kawai (2002).

The time tendencies of horizontal dispersion process $\left(f_\lambda^{\mathrm{hdif}},\ f_\varphi^{\mathrm{hdif}}\right)$ are determined by assigning $(U',\ V')$ to Eqs. (3.12.9) - (3.12.10) rather than $\left(\overline{U},\ \overline{V}\right)$.

Meanwhile, vertical dispersion is represented as atmospheric vertical turbulence. The diffusion coefficient K_{v} is given using vertical shear of mean horizontal wind $\overline{\boldsymbol{U}_{\mathrm{h}}} = \left(\overline{U},\ \overline{V}\right)$ as

$$K_{\mathrm{v}} = l^2\left|\frac{\partial\overline{\boldsymbol{U}_{\mathrm{h}}}}{\partial z}\right| F_{\mathrm{v}}\left(R_{\mathrm{f}}\right) \tag{3.12.16}$$

Here, $F_{\mathrm{v}}\left(R_{\mathrm{f}}\right)$ represents atmospheric stability as a function of the flux Richardson number R_{f} given by the level 2 scheme of Mellor and Yamada (1974, 1982). The mixing length l is given by

$$l = \frac{\kappa z'}{1+\kappa z'/l_0} \tag{3.12.17}$$

Here, z' is height above ground surface, κ is von Kármán constant (set to 0.4) and l_0 is the maximum mixing length given by Holtslag and Boville (1993)

$$l_0 = \begin{cases} l_{\mathrm{FA}} + (l_{\mathrm{PBL}} - l_{\mathrm{FA}})\exp\left(1-\dfrac{z'}{h_{\mathrm{PBL}}}\right) & (z' > h_{\mathrm{PBL}}) \\ l_{\mathrm{PBL}} & (z' \le h_{\mathrm{PBL}}) \end{cases} \tag{3.12.18}$$

The maximum mixing length in the free atmosphere l_{FA}, that in the planetary boundary layer (PBL) l_{PBL} and height of PBL h_{PBL} are set to 30 m, 100 m, 1000 m, respectively.

The time tendency of vertical dispersion process is given by

$$f_{\mathrm{z}}^{\mathrm{vdif}} = W' = \frac{\sum'\sqrt{2K_{\mathrm{v}}\Delta t'}\Gamma}{\Delta t} \qquad (\Delta t = \Sigma'\Delta t') \tag{3.12.19}$$

Here, $\Delta t'\ (\le \Delta t)$ is the limitation of time step to prevent the displacement by vertical dispersion from becoming excessively large (Iwasaki *et al.* 1998).

3.12.3.4 Gravitational Fallout

In the JMA-ATM, it is assumed that volcanic ash particles fall at their terminal velocity w_t, as determined from the equation for balance between gravitational force and air resistance force:

$$w_t(D, z) = \sqrt{\frac{4C_C(\rho_p - \rho_a)gD}{3C_a\rho_a}} \tag{3.12.20}$$

where ρ_p is particle density, ρ_a is air density and C_a is the drag coefficient in consideration of the shape parameter $F = (a_2 + a_3)/2a_1$ (where a_1, a_2 and a_3 are particle principal axes, with a_1 as the longest) as given by Suzuki (1983):

$$C_a = \frac{24}{Re}F^{-0.32} + 2\sqrt{1.07 - F} \tag{3.12.21}$$

where Re is the Reynolds number represented as $Re = \rho_a w_t D/\eta_a$ with air viscosity η_a. C_C is the Cunningham correction factor, which is used to account for the reduction of drag on small particles and is expressed as

$$C_C = 1 + Kn\left[1.257 + 0.400\exp\left(-\frac{1.100}{Kn}\right)\right] \tag{3.12.22}$$

with the Knudsen number $Kn = 2\mathrm{MFP}/D$ based on the mean free path of air MFP. η_a and MFP are given as follows:

$$\eta_a(z) = \eta_0\left[\frac{1 + C_S/T_0}{1 + C_S/T(z)}\right]\left[\frac{T(z)}{T_0}\right]^{1/2} \tag{3.12.23}$$

$$\mathrm{MFP}(z) = \mathrm{MFP}_0\frac{\eta_a(z)}{\eta_0}\left[\frac{p(z)}{p_0}\right]^{-1}\left[\frac{T(z)}{T_0}\right]^{1/2} \tag{3.12.24}$$

where $p(z)$ is air pressure at height z, $T(z)$ is air temperature at z, and C_S is the Sutherland constant of air (117 K). η_0 (18.18 μPa s) and MFP_0 (0.0662 μm) are the standard values for the reference atmosphere ($T_0 = 293.15$ K and $p_0 = 1013.25$ hPa).

The density of volcanic ash particles ρ_p is defined as a function of diameter D:

$$\rho_p(D) = \frac{\rho_{\min} + a\rho_{\max}D}{1 + aD} \tag{3.12.25}$$

where $a = 5 \times 10^3\ \mathrm{m}^{-1}$, $\rho_{\max}$ is the density for coarse tephra and is set to $1 \times 10^3\ \mathrm{kg\,m}^{-3}$ as per the density of pumice stone, $\rho_{\min}$ is the density for fine ash and is set to $2.4 \times 10^3\ \mathrm{kg\,m}^{-3}$.

The time tendency of gravitational fallout process is given by

$$f_z^{\mathrm{fall}} = -w_t \tag{3.12.26}$$

3.12.3.5 Dry and Wet Deposition

Dry deposition works on tracers within the surface boundary layer. The dry deposition rate P_d is computed from the depth of the surface boundary layer L_r and dry deposition velocity V_d as follows:

$$P_d = \begin{cases} \dfrac{V_d\Delta t}{L_r} & (z' \le L_r) \\ 0 & (z' > L_r) \end{cases} \tag{3.12.27}$$

V_d is represented in consideration of the aerodynamic resistance r_a (Kitada *et al.* 1986)

$$V_d = \frac{1}{r_a} \tag{3.12.28}$$

$$r_a = \frac{1}{\kappa u_*}\ln\frac{z' - d}{z_0} \tag{3.12.29}$$

$$u_* = \kappa U_{\rm h} \left| \ln \frac{z' - d}{z_0} \right. \tag{3.12.30}$$

where $U_{\rm h}$ is horizontal wind speed. Roughness length z_0 and zero-plane displacement d are determined from vegetation type or sea surface temperature.

For tracers judged to be deposited based on the probability Eq. (3.12.27), the time tendency of the dry deposition process is given by

$$f_{\rm z}^{\rm ddep} = -\frac{z'}{\Delta t} \tag{3.12.31}$$

which means the tracers are deposited on the surface while one time step.

Wet deposition involves a washout process (below-cloud scavenging) representing the deposition of tracers via precipitation and a rainout process (in-cloud scavenging) representing removal of tracers via their roles as cloud condensation nuclei. In the JMA-ATM for operational use, only the washout process is considered. The wet deposition probability associated with the type of precipitation (α = r(rain), s(snow), g(graupel)) is given by

$$P_{\rm w(\alpha)} = \begin{cases} 1 - e^{-\Lambda_{\rm w(\alpha)} \Delta t} & (z' < L_{\rm b}) \\ 0 & (z' \geq L_{\rm b}) \end{cases} \tag{3.12.32}$$

where $L_{\rm b}$ is the height of cloud base and $\Lambda_{\rm w(\alpha)}$ is the wet deposition rate given by

$$\Lambda_{\rm w(\alpha)} = A_\alpha R_\alpha^{B_\alpha} \tag{3.12.33}$$

where R_α [mm h^{-1}] is precipitation intensity. The below-cloud scavenging coefficients are set to $A_{\rm r} = A_{\rm s} = A_{\rm g} = 2.98 \times 10^{-5}$ s^{-1}, $B_{\rm r} = 0.75$ and $B_{\rm s} = B_{\rm g} = 0.30$.

The time tendency of washout process is given by

$$f_{\rm z}^{\rm wout(\alpha)} = -\frac{z'}{\Delta t} \tag{3.12.34}$$

in the same way as for the dry deposition process.

3.12.3.6 Time Integration

Two types of time integral methods, the Euler method (as shown in Eqs. (3.12.6) - (3.12.8)) and the four-stage Runge-Kutta (RK4) method are implemented for the JMA-ATM. The Euler method integrates time tendencies of processes which have dependency on time step Δt such as dispersion processes and deposition processes. While the RK4 method is used for transport processes and gravitational fallout process for VAFFs. Following the RK4 method, a variable ϕ_i after the time step Δt is calculated by

$$\phi_i(t + \Delta t) = \phi_i(t) + \frac{1}{6} \sum_k \left(d_i^{k(1)} + 2 d_i^{k(2)} + 2 d_i^{k(3)} + d_i^{k(4)} \right) \Delta t \tag{3.12.35}$$

$d_i^{k(1)}$, $d_i^{k(2)}$, $d_i^{k(3)}$ and $d_i^{k(4)}$ are specified by the time tendency f_i^k of each process k as

$$d_i^{k(1)}(t) = f_i^k \left(t,\, \phi_i(t) \right) \tag{3.12.36}$$

$$d_i^{k(2)}(t) = f_i^k \left(t + \frac{1}{2}\Delta t,\, \phi_i(t) + \frac{1}{2} d_i^{k(1)} \Delta t \right) \tag{3.12.37}$$

$$d_i^{k(3)}(t) = f_i^k \left(t + \frac{1}{2}\Delta t,\, \phi_i(t) + \frac{1}{2} d_i^{k(2)} \Delta t \right) \tag{3.12.38}$$

$$d_i^{k(4)}(t) = f_i^k \left(t + \Delta t,\, \phi_i(t) + d_i^{k(3)} \Delta t \right) \tag{3.12.39}$$

3.12.4 Products

3.12.4.1 Volcanic Ash Advisory

VAAC Tokyo issues VAAs in text, graphical and IWXXM (XML/GML) formats as defined in ICAO (2018). VAAs include information on the forecast height and area of ash clouds 6, 12 and 18 hours ahead of observation times based on JMA-ATM results. The forecast is normally updated every six hours (at 00, 06, 12 and 18 UTC) when ash clouds are identified in satellite imagery. If notable changes occur in ash clouds, updates are provided as often as needed. Sample VAAs are shown in Figures 3.12.1 and 3.12.2.

FVFE01 RJTD 130600
VA ADVISORY
DTG: 20210813/0600Z
VAAC: TOKYO
VOLCANO: FUKUTOKU-OKA-NO-BA 284130
PSN: N2417 E14129
AREA: JAPAN
SUMMIT ELEV: -29M
ADVISORY NR: 2021/3
INFO SOURCE: HIMAWARI-8
AVIATION COLOUR CODE: NIL
ERUPTION DETAILS: VA EMISSIONS CONTINUING
OBS VA DTG: 13/0520Z
OBS VA CLD: SFC/FL540 N2413 E14136 - N2251 E13850 - N2243 E13455 - N2336 E13423 - N2455 E13655 - N2442 E14134 MOV W 50KT
FCST VA CLD +6 HR: 13/1120Z SFC/FL560 N2420 E14141 - N2116 E13432 - N2121 E13006 - N2157 E12908 - N2348 E13229 - N2529 E13959
FCST VA CLD +12 HR: 13/1720Z SFC/FL550 N2415 E14153 - N1904 E12509 - N1926 E12334 - N2441 E13143 - N2646 E13913
FCST VA CLD +18 HR: 13/2320Z SFC/FL560 N2413 E14153 - N1623 E11912 - N1648 E11823 - N2520 E12821 - N2745 E13859
RMK: NIL
NXT ADVISORY: 20210813/0900Z=

Figure 3.12.1: Sample VAA in text form

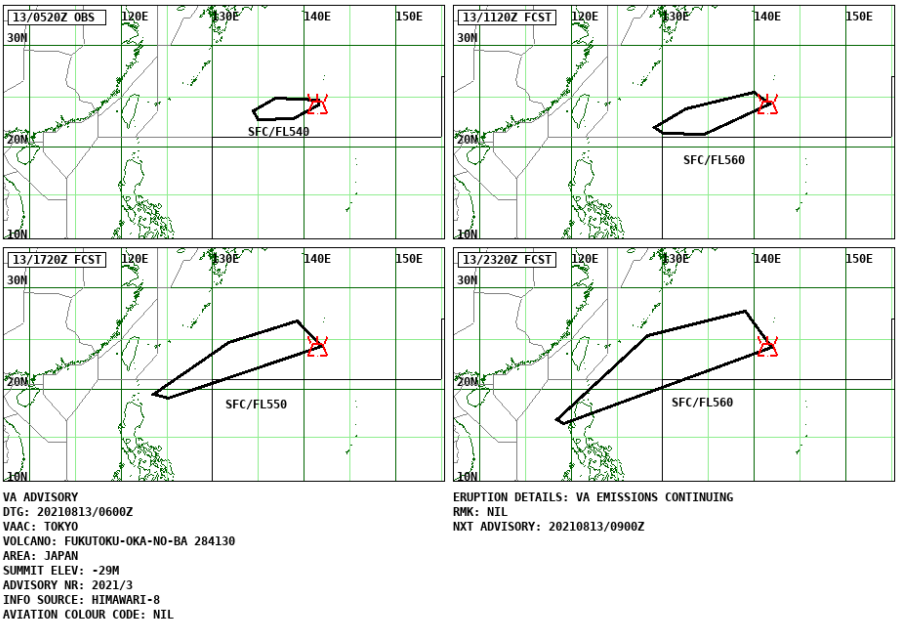


Figure 3.12.2: Sample VAA in graphical form

3.12.4.2 Volcanic Ash Fall Forecast

JMA provides three types of forecasts sequentially: VAFFs (Scheduled) are issued periodically based on an assumed eruption for active volcanoes, VAFFs (Preliminary) are brief forecasts up to one hour ahead issued within 5 - 10 minutes of an actual eruption, and VAFFs (Detailed) are more accurate forecasts up to six hours ahead every hour issued within 20 - 30 minutes of an actual eruption. The amount of volcanic ash fall and maximum particle size at grid points with spacing of 0.02 degrees are calculated based on tracers falling to the surface for use in VAFFs. The amount of ash fall is shown in three qualitative categories combined with recommended action to be taken by the general public (Hasegawa *et al.* 2015). VAFFs include graphical information on ash quantities and lapilli pieces with dimensions of 1 cm or more as well as text information on the direction of ash drift, municipalities affected and precautions for disaster prevention. Sample VAFFs are shown in Figures 3.12.3 (Scheduled), 3.12.4 (Preliminary) and 3.12.5 (Detailed).

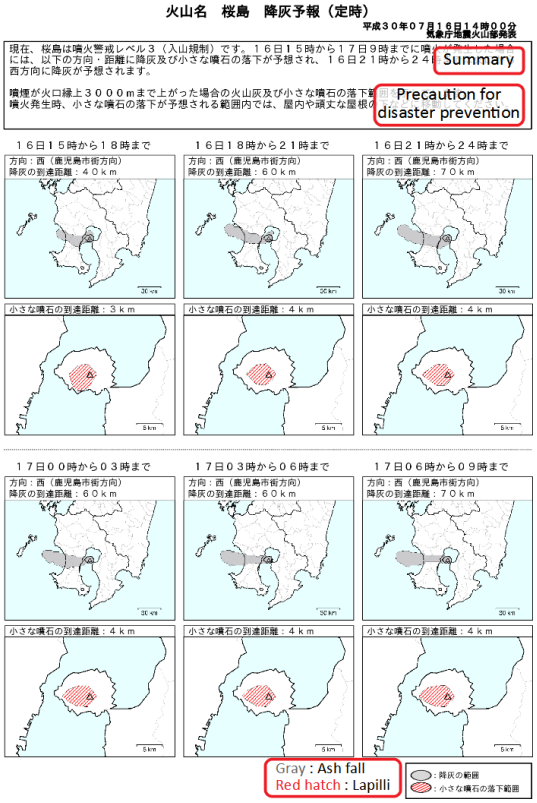


Figure 3.12.3: Sample VAFF (Scheduled)

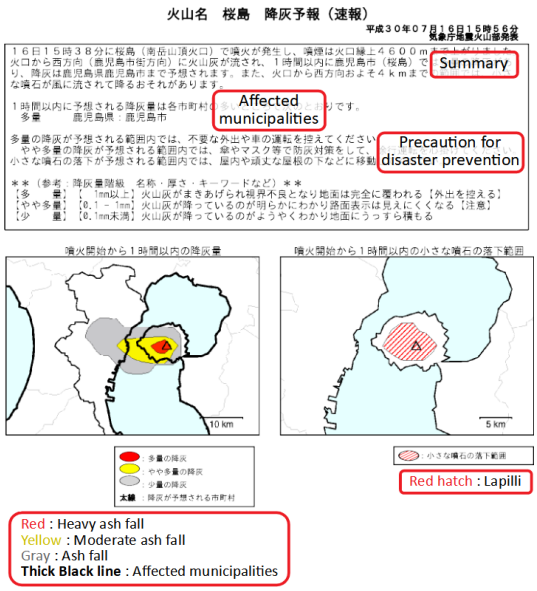


Figure 3.12.4: Sample VAFF (Preliminary)

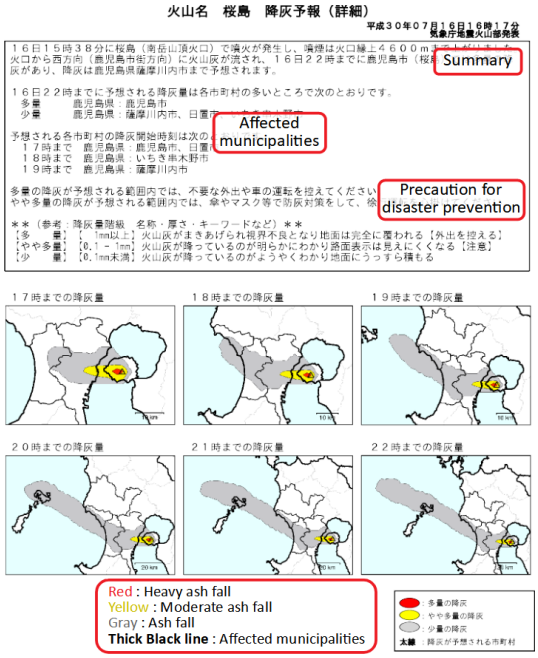


Figure 3.12.5: Sample VAFF (Detailed)

Chapter 4

NWP Application Products

4.1 Summary

JMA provides application products based on both direct and derived NWP results to its own local observatories as well as to corporate operators and related organizations both domestically and internationally.

Direct NWP results have long been delivered in facsimile chart form, and today's telecommunication infrastructure and sophisticated visualization systems now also enable provision of gridded data products. Derived products include very-short-range forecasting of precipitation, half-hourly analysis of three-dimensional temperature and wind distribution, various kinds of guidance for short-range forecasting and aviation services, statistics from ensemble prediction systems, and atmospheric angular momentum functions.

The specifications of these products and their utilization by JMA are described here.

4.2 Weather Chart Services

Facsimile chart provision is a conventional service operated to disseminate the results of NWP in graphical form. Under this service, JMA facsimile charts are sent to national meteorological services via the Global Telecommunication System (GTS) and to ships via the shortwave radio transmission (call sign:JMH).

Table 4.2.1 and Figure 4.2.1 give summaries of weather charts readily accessible by international users (i.e., those provided through GTS and JMH).

The development of the Web complements and supports innovation in these services, and a number of related projects are under way worldwide. JMA takes part in international initiatives such as the Project on the Provision of City-Specific Numerical Weather Prediction (NWP) Products to Developing Countries via the Internet in WMO Regional Association II (RA II) and the Severe Weather Forecasting Programme (SWFP) involving WMO RAs II and V. The Agency's own projects in this regard include the JMA Pilot Project on EPS Products and SATAID Services on the WMO Information System.

Table 4.2.1: List of facsimile charts provided through GTS and radio facsimile JMH. Symbols for vertical level: Surf: surface, Trop: tropopause, numbers (850, 700, ... 100): level of pressure in hPa; Symbols for contours: D: dewpoint depression ($T - T_d$), E: precipitation (over the past 12 h for 24 h forecast, and over the past 24 h for others), H: geopotential height, J: wave height, O: vertical velocity (ω), P: MSL pressure, T: temperature, W: wind speed (isotachs), Z: vorticity; Symbols for other drawings: a: wind arrow from gridded data, b: observation plots, d: hatch for area $T - T_d < 3$ K, g: arrow for prevailing wave direction, j: jet axis, m: wave period, s: daily mean sea surface temperature, t: temperature numbers, x: streamlines; Symbols for dissemination and temporal specialty: ': sent to GTS, *: sent to JMH, ¶: only for 00 UTC, §: only for 12 UTC.

Model	Area (see Figure 4.2.1)	Forecast Time							
		Analysis	12h	24h	36h	48h	72h	96h 120h	144h 168h 192h
GSM	A' (Far East)	500 (H, Z)' 850 (T; a)+700 (O)'		500 (T)+700 (D)'* 500 (H, Z)'* 850 (T; a)+700 (O)'* Surf(P, E; a)'*					
	C (East Asia)	300 (H, W; a, t, b)'¶ 500 (H, T; a, b)'* 700 (H, T; b, d)' 850 (H, T; b, d)'*		Surf(P, E)'		500 (H, Z)' 850 (T, a)+700 (O)'§ Surf(P, E)'*		Surf(P, E)'*§	
	O (Asia)							500 (H, Z)'§ Surf(P)+850 (T)'§	
	Q (Asia-Pacific)	200 (H, W; t, a, j)+Trop(H)' 250 (H, W; t, a)'		250 (H, W; t, a)' 500 (H, W; t, a)'					
	W (West Pacific)	200 (x)' 850 (x)'		200 (x)' 850 (x)'		200 (x)' 850 (x)'			
	D (N Hem.)	500 (H, T)'§							
Ocean Wave	X (Japan)	Surf(J; b, g, m)'*		Surf(J; b, g, m)'*					
	C" (NW Pacific)	Surf(J; g, m)'*	Surf(J; g, m)'*			Surf(J; g, m)'*			
Sea Surface Temperature	C"2 (NW Pacific)	Surf(s)'*¶							

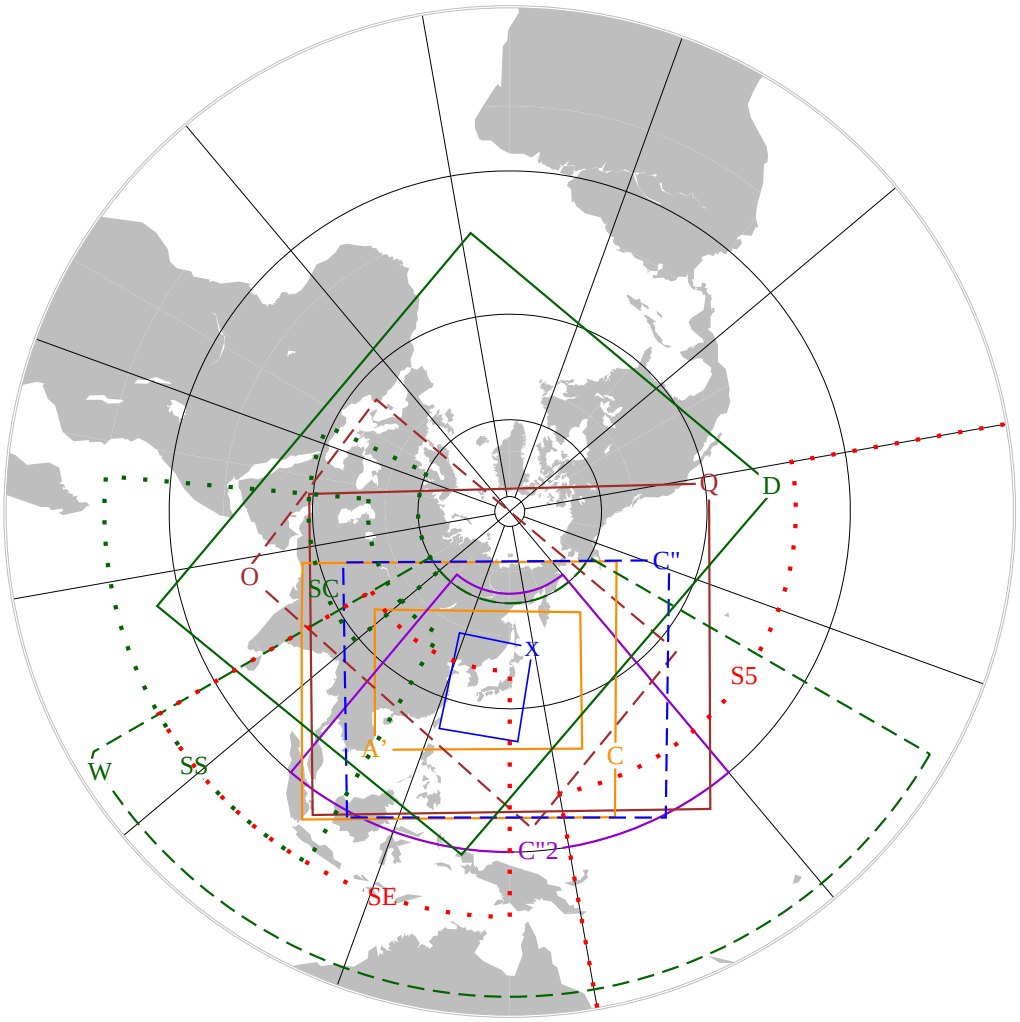


Figure 4.2.1: Areas for charts disseminated through GTS and radio facsimile JMH (symbols A', C, C", C"2, D, O, Q, W and X). The dotted boxes labeled SE, SS, SC and S5 are areas of SWFP products for Southeast Asia, South Asia, Central Asia, and WMO Region V, respectively (for information).

4.3 Gridded Data Products

As part of JMA's general responsibility in meteorological information service provision, gridded data products are distributed to domestic and international users. In line with the requirements of the WMO Information System (WIS), this data service utilizes both dedicated and public (i.e., Internet) network infrastructure.

The dedicated infrastructure consists of an international part called GTS, together with domestic parts inside JMA (including the Meteorological Satellite Center and the Meteorological Research Institute) and provision to government agencies and the Meteorological Business Support Center, which is in charge of managed services for general users including those in the private sector.

The portal to JMA's international services over the Internet is the website of the Global Information System Centre (GISC) Tokyo[1]. Currently, the international service for gridded data products includes the GSM, the Global Ensemble Forecast (One-week) and the Ocean Wave Model as listed in Table 4.3.1 and Table 4.3.2.

[1] https://www.wis-jma.go.jp

Table 4.3.1: List of gridded data products transmitted through GTS and the GISC Tokyo website. Symbols for content: D: dewpoint depression ($T - T_d$), E: precipitation (from initial time), E_6: precipitation over the past 6h, G: prevailing wave direction, H: geopotential height, J: wave height, M: wave period, O: vertical velocity (ω), P: MSL pressure, R: relative humidity, T: temperature, U: eastward wind speed, V: northward wind speed, W: vertical wind shear, X: stream function, Y: velocity potential, Z: vorticity, μ: average over ensemble, σ: standard deviation over ensemble. The symbols °, *, ¶, §, ♯, ♭, †, ‡are notes on availability, as detailed in the table.

Model	GSM	GSM	GSM
Service Channel	GTS and GISC	GTS and GISC	GTS and GISC
Code form	GRIB Edition 1	GRIB Edition 1	GRIB Edition 1
Area	Whole Globe	20°S–60°N 60°E–160°W	Whole Globe
Resolution	1.25° × 1.25°	1.25° × 1.25°	2.5° × 2.5°
Contents 10, 20 hPa	H, U, V, T	H, U, V, T	H*, U*, V*, T*
30, 50, 70hPa	H, U, V, T	H, U, V, T	H°, U°, V°, T°
100 hPa	H, U, V, T, W♯	H, U, V, T	H°, U°, V°, T°
150 hPa	H, U, V, T, W♯	H, U, V, T	H*, U*, V*, T*
200 hPa	H, U, V, T, X, Y, W♯	H§, U§, V§, T§, X, Y	H, U, V, T
250 hPa	H, U, V, T, W♯	H, U, V, T	H°, U°, V°, T°
300 hPa	H, U, V, T, R, O, W♯	H, U, V, T, D	H, U, V, T, D*‡
400 hPa	H, U, V, T, R, O, W♯	H, U, V, T, D	H*, U*, V*, T*, D*‡
500 hPa	H, U, V, T, R, O, Z, W♯	H§, U§, V§, T§, D§, Z	H, U, V, T, D*‡
600 hPa	H, U, V, T, R, O		
700 hPa	H, U, V, T, R, O, W♯	H§, U§, V§, T§, D§, O	H, U, V, T, D
850 hPa	H, U, V, T, R, O, X, Y	H§, U§, V§, T§, D§, O, X, Y	H, U, V, T, D
925 hPa	H, U, V, T, R, O	H, U, V, T, D, O	
1000 hPa	H, U, V, T, R, O	H, U, V, T, D	H, U*, V*, T*, D*‡
Surface	P, U, V, T, R, E†, E_6♭†	P¶, U¶, V¶, T¶, D¶, E¶	P, U, V, T, D*‡, E†
Forecast time range (*from–until/interval*)	0–84h/6h †: except for analysis ♯: 0-36h/6h, ♭: 6-36h/6h	0–84h/6h	0–72h/24h *: Analysis only
Extension on 12UTC	96–192h/12h	§: 96–192h/24h ¶: 90–192h/6h	96–192h/24h °: 96–120h/24h
Initial times	00, 06, 12, 18UTC	00, 06, 12, 18UTC	00UTC and 12UTC ‡: 00UTC only

Model		Global Ensemble Forecast (One-Week)	Ocean Wave Model
Service Channel		GTS and GISC	GTS and GISC
Code form		GRIB Edition 1	GRIB Edition 2
Area		Whole Globe	75°S–75°N, 0°E–359.5°E
Resolution		2.5° × 2.5°	0.5° × 0.5°
Contents	250 hPa	μU, σU, μV, σV	
	500 hPa	μH, σH	
	850 hPa	μU, σU, μV, σV, μT, σT	
	1000 hPa	μH, σH	
	Surface	μP, σP	J, M, G
Forecast time range		0–192h/12h	0–84h/6h
Extension on 12UTC		(none)	96–192h/12h
Initial times		00UTC and 12UTC	00,06,12,18UTC

Table 4.3.2: List of gridded data products transmitted through the GISC Tokyo website. Symbols for content: C_L: low cloud amount, C_M: middle cloud amount, C_H: high cloud amount, E: precipitation (from initial time), H: geopotential height, N: total cloudiness, O: vertical velocity (ω), Di: relative divergence, P: MSL pressure, P_S: surface pressure, R: relative humidity, T: temperature, U: eastward wind speed, V: northward wind speed, X: stream function, Y: velocity potential, Z: vorticity, W: wind speed, G: gusts, μ: average over ensemble, σ: standard deviation over ensemble, ρ: probability of ensemble prediction results (parentheses represent probability thresholds). The symbol †is a note on availability, as detailed in the table.

Model	GSM	GSM
Service Channel	GISC	GISC
Code form	GRIB Edition 2	GRIB Edition 2
Area	Whole Globe and also 5°S–90°N, 30°E–165°W	Whole Globe
Resolution	0.5° × 0.5° (0.25° × 0.25° for surface)	1.25° × 1.25°
Contents		
10, 20, 30, 50, 70, 100, 150hPa	H, U, V, T, R, O	H, U, V, T
200 hPa	H, U, V, T, R, O, X, Y	H, U, V, T, X, Y
250 hPa	H, U, V, T, R, O	H, U, V, T, Z ,Di
300, 400 hPa	H, U, V, T, R, O	H, U, V, T, R, O
500 hPa	H, U, V, T, R, O, Z	H, U, V, T, R, O, Z
600 hPa	H, U, V, T, R, O	H, U, V, T, R, O
700 hPa	H, U, V, T, R, O	H, U, V, T, R, O, Z, Di
800 hPa	H, U, V, T, R, O	
850 hPa	H, U, V, T, R, O, X, Y	H, U, V, T, R, O, X, Y
900 hPa	H, U, V, T, R, O	
925 hPa	H, U, V, T, R, O	H, U, V, T, R, O, Z, Di
950, 975 hPa	H, U, V, T, R, O	
1000 hPa	H, U, V, T, R, O	H, U, V, T, R, O
Surface	P, U, V, T, R, E†, P_S, N, C_L, C_M, C_H	U, V, T, R, P, E†
Forecast time range (*from–until/interval*)	0–132h/3h †: except for analysis	0–132h/6h
Extension on 00 and 12UTC	138–264h/6h	144–264h/12h
Initial times	00, 06, 12, 18UTC	00, 06, 12, 18UTC

Model		Global Ensemble Forecast (One-Week)
Service Channel		GISC
Code form		GRIB Edition 2
Area		Whole Globe
Resolution		1.25° × 1.25°
Contents	250hPa	μU, σU, μV, σV
	500hPa	μH, σH
	850hPa	μU, σU, μV, σV, μT, σT, μW, σW, ρT(±1, ±1.5, ±2)
	1000 hPa	μH, σH
	Surface	μP, σP, ρE(1, 5, 10, 25, 50, 100 mm/24h), ρW(10, 15, 25 m/s), ρG(10, 15, 25 m/s)
Forecast time range		0–264h/12h
Initial times		00UTC and 12UTC

4.4 Very-short-range Forecasting of Precipitation

JMA has operated a fully automated system for analysis and very-short-range forecasting of precipitation since 1988 to provide the following products for monitoring and forecasting of local severe weather conditions:

1. Radar/Raingauge-Analyzed Precipitation (R/A; a type of precipitation analysis)
2. Very-Short-Range Forecasting of Precipitation (VSRF; a type of precipitation forecast)
3. Extended VSRF (ExtVSRF; a type of extended precipitation forecast)

The data in these products show a close correlation with rainfall amounts observed using raingauges. From R/A and VSRF, indices with close ties to landslides, flooding and inundation are produced and used to issue advisories and warnings for these phenomena. The products are provided to local meteorological offices, local governments and broadcasting stations which have responsibility for disaster mitigation. This section outlines how the products are created.

Table 4.4.1: Specifications of Radar/Raingauge-Analyzed Precipitation (R/A), Very-Short-Range Forecasting of Precipitation (VSRF) and Extended VSRF (ExtVSRF)

	R/A	VSRF	ExtVSRF
Spatial resolution	1 km	1 km	5 km
Update interval	10 min.	10 min.	1 hour
Analysis/Forecast element	1 hour accumulated rainfall amount	1 hour accumulated rainfall amount	1 hour accumulated rainfall amount
Forecast time	-	Up to 6 hours ahead	From 7 to 15 hours ahead
Forecast interval	-	1 hour	1 hour
Time required to execute	About 5 min. after observation time	About 8 min. after observation time	About 18 min. after observation time

4.4.1 Analysis of Precipitation (R/A)

R/A enables estimation of accurate one-hour precipitation amounts based on precipitation intensity as observed using radars and rainfall amounts observed using raingauges. It involves the use of data from 46 radars (20 JMA units and 26 operated by other organizations) to cover large areas at a higher spatial resolution than the raingauge network as well as data from up to 10,000 raingauges (1,300 AMeDAS units and 8,700 operated by other organizations) to determine actual amounts of precipitation. JMA uses the Z-R relationship to convert the radar reflectivity factor to precipitation intensity. The Kdp-R relationship (Bringi and Chandrasekar 2001) is also clarified for dual-polarized JMA radar, and is used to estimate high-intensity precipitation in ten-minute precipitation amounts close to analysis times.

The one-hour cumulative precipitation amounts estimated using radar observation usually differ from those observed with raingauges. Radar amount calculation is based on cumulative precipitation intensity over an hour, and values are calibrated using raingauge data to enable more accurate estimation (Makihara 2000). A schematic diagram of this procedure is shown in Figure 4.4.1. First, the radar data are quality controlled to remove ground and sea clutter, bright band and weak echoes unrelated to rainfall on the ground. Satellite imagery and NWP gridded data are used for this purpose. Then, the primary calibration factor over the entire detection range of each radar is calculated by comparing radar precipitation to that of neighboring radars and raingauge data with differences in radar beam height taken into account. Next, the secondary calibration factor is calculated by comparing radar precipitation calibrated with the primary factor to raingauge data at grids where raingauges are located to evaluate local heavy precipitation more accurately. For grids with no raingauges, factor calculation is based on weighted interpolation of values for surrounding grids where raingauges are present. A nationwide composite chart of all radar calibrated precipitation data is created using the

maximum value method, in which the largest value is selected if a grid has several data from observation by multiple radars.

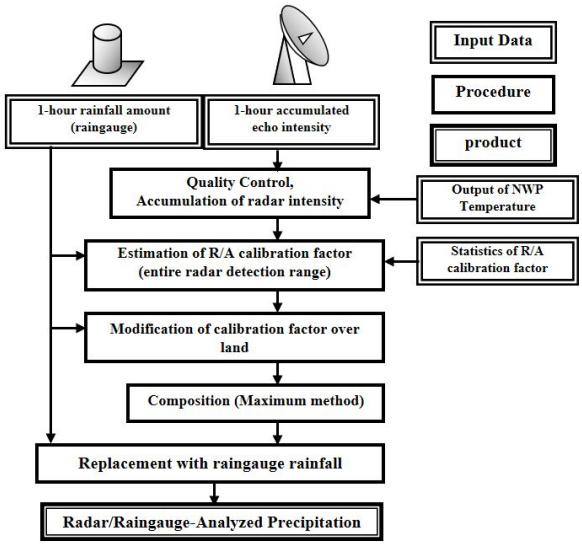


Figure 4.4.1: Flow of Radar/Raingauge-Analyzed Precipitation

4.4.2 Forecasting of Precipitation up to 6 hours ahead (VSRF)

VSRF, which employs the calibrated precipitation intensity determined in the course of R/A as the initial value and is formulated from extrapolation and model forecasts, is a superior estimate of precipitation.

A schematic diagram of the related procedure is shown in Figure 4.4.2. Two methods are used for VSRF. One is the extrapolation of movements of analyzed precipitation systems (i.e., extrapolation forecasts; referred to here as EX6). In the course of extrapolation, the growth and decay of precipitation systems caused by orographic effects and echo intensity trends are taken into account. The other method involves precipitation forecasts of the MSM and LFM, which are available after around two hours from the initial time for the MSM and one hour for the LFM. EX6 is more skillful than MSM and LFM forecasts, although the forecast time is short and skill rapidly diminishes. Meanwhile, the skill of MSM and LFM forecasts degrades gradually and is comparable to that of EX6 when the forecast time reaches a few hours. To produce better model forecasts (referred to here as BLD), JMA introduced a blending technique involving weighted averaging of MSM forecasts and LFM forecasts. This is similar to the merging technique outlined below (see Subsection 4.4.2.2)

The merging technique essentially involves weighted averaging of the EX6 and the BLD. As merging weights are set close to zero for the BLD in the first hour, the products are similar to EX6 output. Thereafter, merging weights for the BLD increase with forecast time. These are determined by comparing the skills of the BLD and the EX6 with R/A. The forecast time steps are two or five minutes, and forecast precipitation is accumulated to produce hourly forecasts up to six hours ahead.

4.4.2.1 Processes assumed in EX6

Extrapolation vectors (i.e., the movement vectors of precipitation systems) are evaluated using a generalized cross-correlation method involving comparison of precipitation system locations at the initial time with those at 0.5, 1, 2 and 3 hours before.

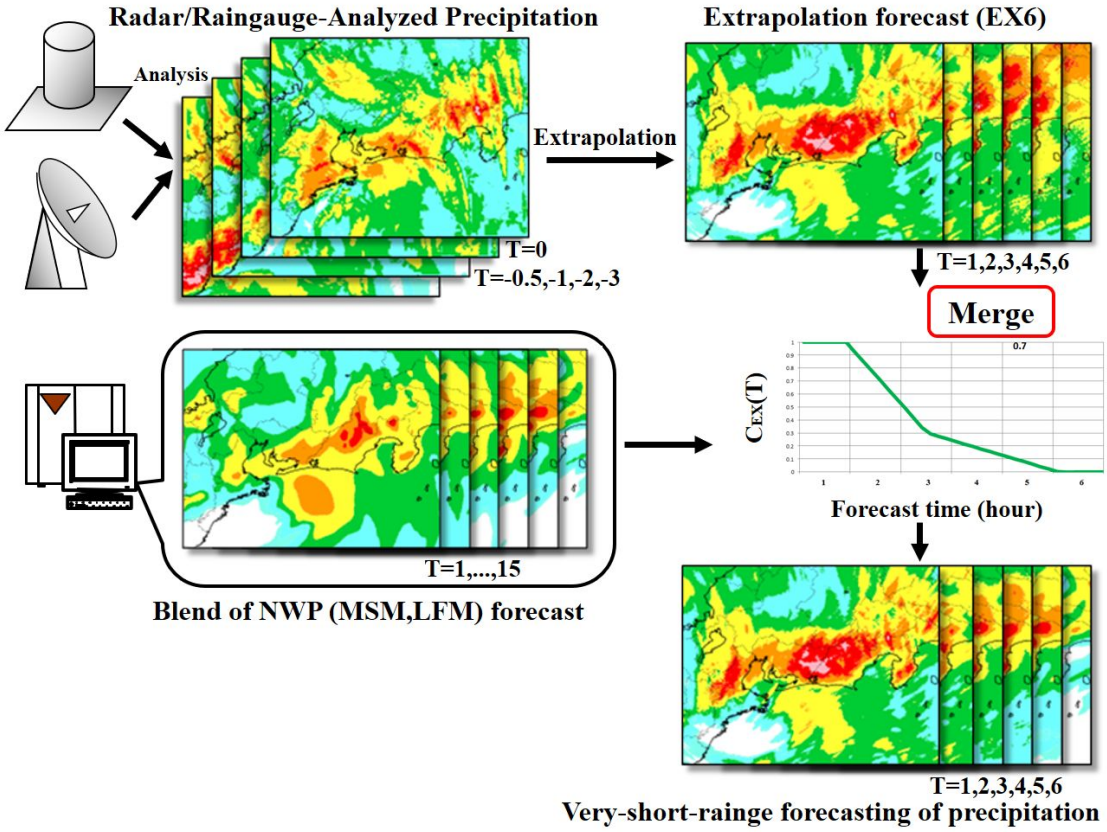


Figure 4.4.2: Schematic representation of very-short-range precipitation forecasting

As the seeder-feeder mechanism is assumed to work in regions of orographic updraft, precipitation systems are allowed to grow in the course of extrapolation over such regions. Precipitation systems that have passed over mountains higher than the echo top height are decayed when the following conditions are met:

1. Orographic downslope motion of the rain system is expected from the low-level wind of the MSM,

2. The direction of rain system movement or that of the 700 hPa wind of the MSM is largely parallel to that of the 900 hPa wind of the MSM.

Echo intensity trends can also be determined by comparing the current area average of echo intensity to a past one. Movement vectors for intensity trends are calculated in addition to extrapolation vectors. The vectors move echo intensity trends, which in turn change forecast precipitation.

4.4.2.2 Merging Technique

First, the relative skill of the EX6 and the BLD are estimated. The EX6 from three hours before is verified against the current analysis. For the BLD forecast, the latest available data are verified against the current analysis. The relative reliability coefficient C_{RR} is defined as follows:

$$C_{RR} = \min\left(1, \frac{D_{EX}}{D_{BLD}}\right) \tag{4.4.1}$$

where D_{EX} is the two-dimensional pattern distance, or the two-dimensionally extended Levenshtein distance, between the EX6 and the analysis, and D_{BLD} is the two-dimensional pattern distance between the BLD forecast and the analysis.

The relative weight of the extrapolation forecast $C_{EX}(T)$ is then determined using C_{RR} and the function $C(T, BR)$, where BR [2] is blend reliability and T denotes the forecast time as indicated in the merge process of Figure 4.4.3

$$C_{EX}(T) = 1 - C_{RR} \cdot (1 - C(T, BR)) \tag{4.4.2}$$

Finally, the merged forecast $R_{MRG}(T)$ is calculated with the following equation:

$$R_{MRG}(T) = C_{EX}(T) \cdot R_{EX}(T) + (1 - C_{EX}(T)) \cdot R_{BLD}(T) \tag{4.4.3}$$

where $R_{EX}(T)$ denotes extrapolation forecasting of precipitation at the forecast time T and $R_{BLD}(T)$ denotes the BLD forecast of precipitation from the latest initial time at the same valid time T.

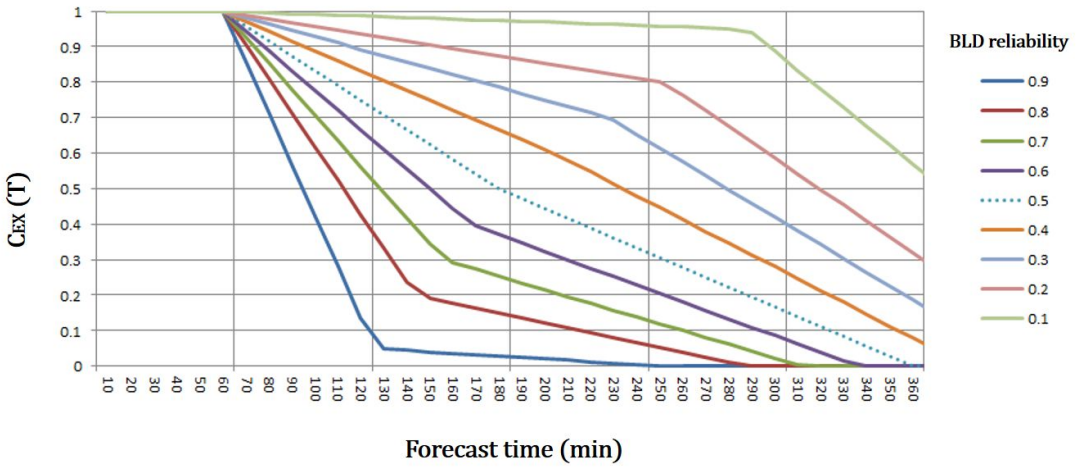


Figure 4.4.3: Dependence of $C_{EX}(T)$ on forecast time and BLD reliability

4.4.2.3 Example and Verification Score of R/A and VSRF

R/A and VSRF examples are shown in Figure 4.4.4. R/A for the Kinki region in the central western area of Japan for 14:50 UTC on 23 August 2018 is shown in the left panel (a), and the three-hour VSRF forecast for the same valid time (i.e., initial time 11:50 UTC 23 on August 2018) is shown in the right panel (b). The intense rain band is well forecasted.

VSRF accuracy has been statistically verified with the critical success index (CSI)[3]. Forecasts are compared with precipitation analysis after both fields are averaged in 5×5km grids. Indices from 1- to 6-hour forecasts for July 2018 are shown in Figure 4.4.5.

It can be seen that scores deteriorate with longer forecast times. EX6 maintains its superiority to BLD up to three hours, but the relationship reverses after this time. VSRF exhibits the best performance for all forecast times.

[2] $BR \equiv Score(BLD)/(Score(BLD) + Score(EX6))$

[3] The CSI is the number of correct "Yes" forecasts divided by the total number of occasions on which the event was forecast and/or observed. It is also cited as the threat score (see Subsection A.2.9).

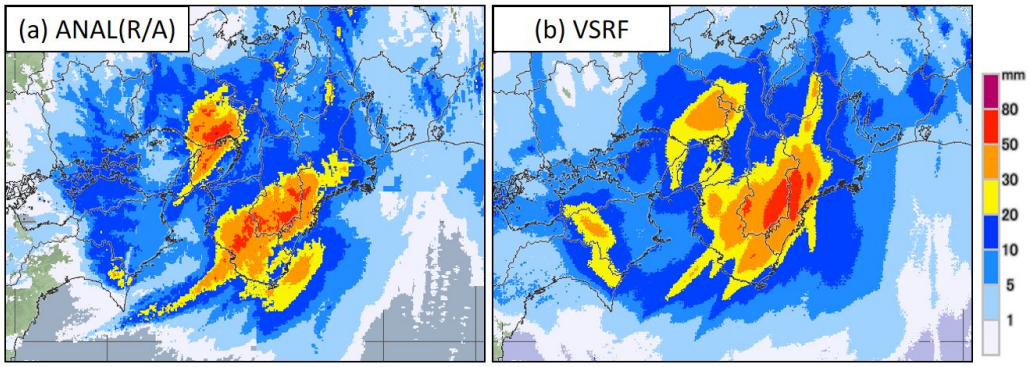


Figure 4.4.4: (a) the Radar/Raingauge-Analyzed Precipitation for 11:50 UTC on 23 August 2018, and (b) 3-hour VSRF forecast of precipitation for the same valid time.

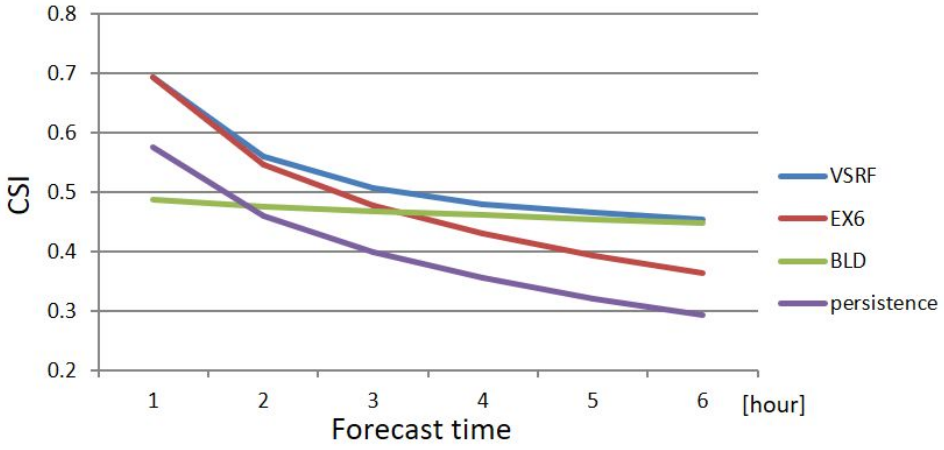


Figure 4.4.5: CSI of very-short-range forecasting (VSRF) of precipitation averaged within 5 × 5km grids over land for July 2018, together with that of EX6, BLD and persistence forecasting. The threshold value is 1.0mm.

4.4.3 VSRF Forecast Range Extension to 15 hours (ExtVSRF)

4.4.3.1 Basic Concept of ExtVSRF

In June 2018, JMA launched an extended VSRF forecast called ExtVSRF to support early judgement on the need for evacuation and other measures by clarifying the tendency of rainfall toward dawn when heavy rain falls in the evening. The extended forecast facilitates understanding of overall precipitation distribution as a trend, and was developed as a separate product from VSRF.

The flow of ExtVSRF calculation is shown in Figure 4.4.6. The forecast is derived from a combination of MSM precipitation amount forecasts, MSM Guidance for mean and maximum precipitation amounts and LFM Guidance for maximum precipitation amounts, and is not merged with the EX6 because the latter's precision from 7 to 15 hours ahead is significantly poorer than that produced by the combination of these guidance forecasts.

The latest available guidance forecasts for mean precipitation amounts and maximum precipitation amounts are divided into two groups, and are verified with current analysis using the fraction skill score (FSS)[4]. The forecast with the best score from each group is chosen and mixed with the weighted average.

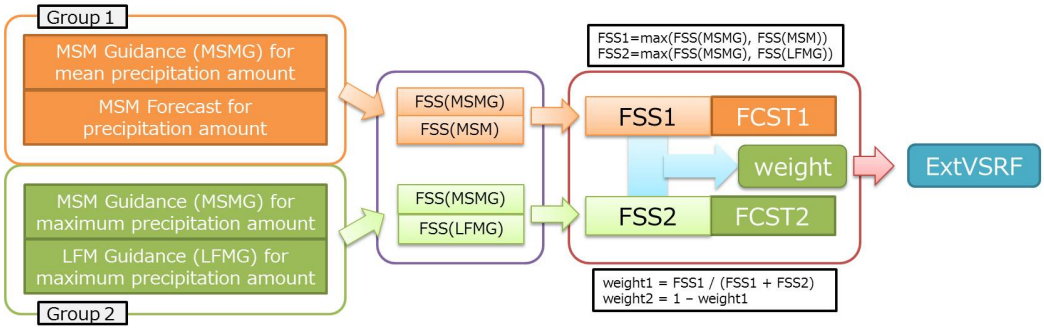


Figure 4.4.6: Extended VSRF flow chart

4.4.3.2 ExtVSRF Verification Score and Example

An example of R/A and ExtVSRF is shown in Figure 4.4.7. The R/A for the western area of Japan at 23 UTC 19 June 2018 (08 JST 20 June 2018) is shown in the left panel (a), and the 14-hour forecast of ExtVSRF at the same valid time (i.e., initial time 09 UTC 19 June 2018 (18 JST 19 June 2018)) is shown in the right panel (b). The intense rain band is well forecast, and it is understandable that this morning rainfall will be heavy at the moment of previous early evening.

The accuracy of ExtVSRF has been statistically verified with FSS data. Forecasts are compared with precipitation analysis in 5 × 5km grids. Indices from 1- to 15-hour forecasts for July/October 2017 and January 2018 are shown in Figure 4.4.8 together with those of VSRF and MSM Guidance forecasts for mean precipitation amounts. It can be seen that ExtVSRF is superior to MSM Guidance for all forecast times.

4.5 Half-hourly Analysis

Half-hourly analysis provides gridded data on three-dimensional temperature and wind every 30 minutes to support continuous atmospheric monitoring/forecasting. Online meteorological imagery is also provided to the aviation sector.

[4] In FSS, verification incorporates consideration of position gaps and indicates the accuracy of precipitation distribution (see Subsection A.2.12).

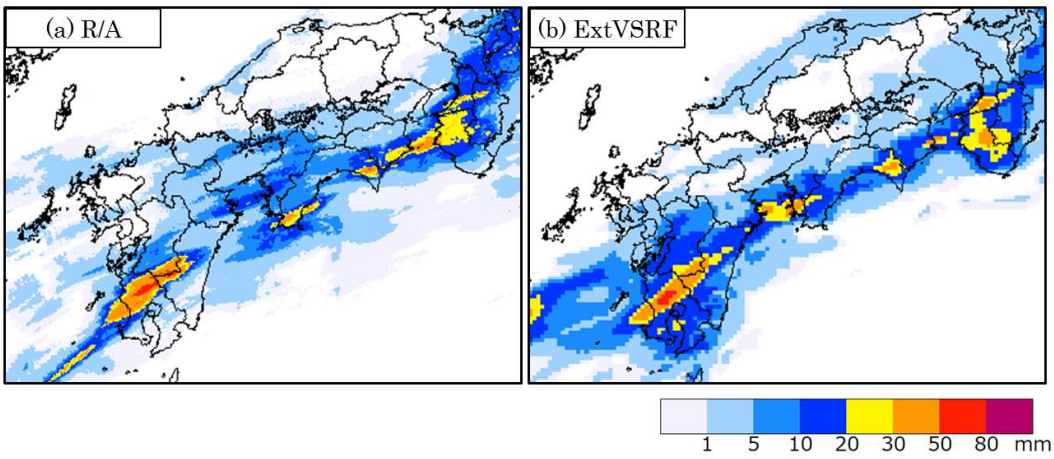


Figure 4.4.7: (a) Radar/Raingauge-Analyzed Precipitation at 23 UTC on 19 June 2018, and (b) 14-hour ExtVSRF forecast of precipitation for the same valid time.

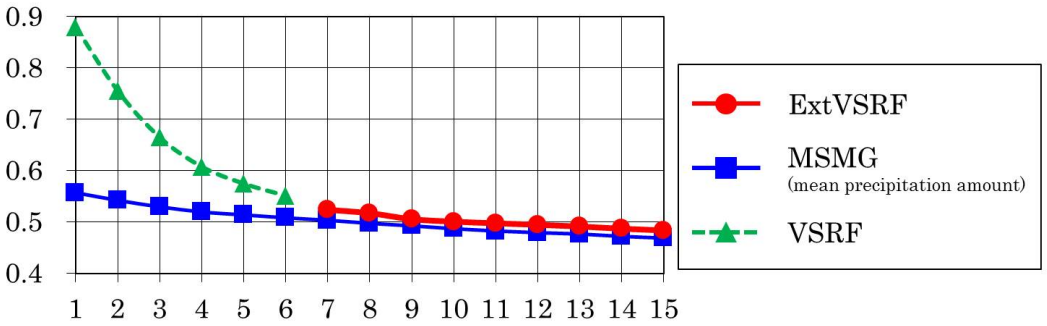


Figure 4.4.8: FSS of the Extended VSRF (ExtVSRF) for June 2017, October 2017 and January 2018 together with that of VSRF and MSM Guidance forecast for mean precipitation amount. The threshold value is 1 mm.

The configuration of the half-hourly analysis system, which is based on 3D-Var, is given in Table 4.5.1. First-guess values are from Local Forecast Model (LFM; Section 3.7) data with a 3- or 3.5-hour forecast depending on the analysis time. The analysis covers Japan and its surrounding area at a horizontal resolution of 2 km with 1,581 × 1,301 grid points. A total of 76 vertical layers are defined for hybrid terrain-following coordinates (Subsection 3.5.3) with the top of the domain at 21,801 m.

Observations assimilated in analysis are from the Automated Meteorological Data Acquisition System (AMeDAS; a nationwide network of surface stations monitoring variables such as wind and temperature), Wind Profiler (wind), Weather Doppler radar (radial velocity), AIREP, AMDAR (wind and temperature), and AMV/wind data from Himawari-9. Data cut-off is 18 minutes past the analysis time to enable product distribution within 30 minutes.

In order to obtain a good fit to surface observations on land, the 3D-Var in the half-hourly analysis adopts a short background error correlation distance and a small observation error on the surface. Thus, the surface field on land typically has large local increments. Outlined below are modifications of the 3D-Var scheme and additional post-processing performed to address this situation.

Table 4.5.1: Half-hourly analysis specifications

Analysis time	Every 30 minutes (on the hour and the half hour)
Analysis scheme	3D-Var
Data cut-off time	18 minutes after analysis time
First guess	3- or 3.5-hour forecast by the LFM
Domain configuration	Japan and its surrounding area Lambert projection: 2 km at 60°N and 30°N, 1581 × 1301 Grid point (1, 1) is at the northwest corner of the domain. Grid point (1121, 901) is at 140°E, 30°N.
Vertical coordinate	z-$z^\star$ hybrid
Vertical levels	76 levels up to 22 km
Analysis variables	Wind, temperature, surface wind and surface temperature
Observations	AMeDAS, Wind Profiler, Weather Doppler radar (radial velocity), AIREP, AMDAR, and AMVs from Himawari-9
Post-processing	Surface filtering (followed by adjustment of the increment within the boundary layer)
Product distribution	Within 30 minutes of analysis time

- The control variables at the bottom level are treated as uncorrelated to those at the other levels in 3D-Var analysis because the large analysis increments at the surface need to be adjusted independently with a surface filter as described below.
- After 3D-Var analysis, a surface filter is applied to the surface temperature and wind fields to attenuate surface increments over the sea with distance from the coastline and to reduce excessive increments in sea regions near the coastline within the range of correlation from land observations. Analysis increments on the surface and in the upper air are inconsistent in this respect.
- After application of the surface filter, the increments on the surface and in the upper air are merged in each vertical column within the boundary layer of the first guess to make the surface and upper-air increments consistent. The weight of the surface increment attenuates with height above the ground, and approaches zero around the top of the boundary layer.

4.6 Guidance for Short-range Forecasting

4.6.1 Overview

JMA provides extensive forecast guidance to support the issuance of warnings, advisories, information and weather forecasts. Related objectives include reduction of systematic errors in NWP output variables (e.g., temperature and wind) and derivation of values not directly calculated in NWP (e.g., the probability of thunderstorm and weather categories). Table 4.6.1 lists the guidance provided for short-range forecasting (up to 84 hours).

To cope with frequent model upgrades, JMA uses methods that allow ongoing adjustment of statistical equations. These adaptive approaches, which are based on Kalman filtering and a neural network, replaced the previous non-successive multivariate regression method in 1996 and have since been applied to various guidance values.

The Kalman filtering and neural network used in the guidance system are outlined in Subsection 4.6.2 and Subsection 4.6.3, respectively, and utilization of guidance at forecasting offices is summarized in Subsection 4.6.4.

4.6.2 Guidance Based on Kalman Filtering

4.6.2.1 Kalman Filtering

As a statistical post-processing method for NWP output, JMA developed guidance using Kalman filtering (KF) on the basis of earlier work conducted by Persson (1991) and Simonsen (1991). KF evolves coefficients based on the following equations:

$$y_\tau = \boldsymbol{c}_\tau \boldsymbol{X}_\tau + v_\tau \tag{4.6.1}$$

$$\boldsymbol{X}_{\tau+1} = \boldsymbol{A}_\tau \boldsymbol{X}_\tau + \boldsymbol{u}_\tau \tag{4.6.2}$$

where y represents a predictand (i.e., the target of forecasting), $\boldsymbol{c}$ represents predictors ($1 \times n$ matrix), $\boldsymbol{X}$ represents coefficients ($n \times 1$ matrix) with a covariance matrix $\boldsymbol{Q}$ ($n \times n$ matrix), v represents observation noise with variance D, $\boldsymbol{u}$ represents system noise ($n \times 1$ matrix) with a covariance matrix $\boldsymbol{U}$ ($n \times n$ matrix), and $\boldsymbol{A}$ ($n \times n$ matrix) describes the evolution of the coefficients in time. The subscript τ for each variable represents the time step.

Eq. (4.6.1) relates to observation, and is a linear expression relating the predictand and predictors. Eq. (4.6.2) is a system exprssion denoting the time evolution of the coefficients. In guidance, the time evolution matrix $\boldsymbol{A}$ can be treated as a unit matrix:

$$\boldsymbol{A}_\tau \equiv \boldsymbol{I} \tag{4.6.3}$$

The objective of KF is to determine the most likely estimation of the coefficients $\boldsymbol{X}_{\tau+1/\tau}$, whose subscript denotes an estimate at $\tau + 1$ based on observation at τ. In contrast, the single subscripts in Eq. (4.6.1) and

Table 4.6.1: Parameters of the guidance products for short-range forecasting

Parameters	Target		Model	Forecast hour	Method*
Categorized weather over 3 hours (fair, cloudy, rainy, sleety, snowy)	Grids	20km	GSM	FT=6, 9, . . . , 84	NN
		5km	MSM	FT=3, 6, . . . , 39/78**	
Mean precipitation amount over 3 hours	Grids	20km	GSM	FT=6, 9, . . . , 84	KF & FBC
		5km	MSM	FT=3, 6, . . . , 39/78	
		5km	MEPS	FT=3, 6, . . . , 39	
Maximum precipitation amount over 1, 3 hours	Grids	20km	GSM	FT=6, 9, . . . , 84	NN
		5km	MSM	FT=3, 6, . . . , 39/78	
		5km	MEPS	FT=3, 6, . . . , 39	
Maximum precipitation amount over 24, 48 and 72 hours (48, 72 hours:GSM only)	Grids	20km	GSM	FT=27, 30, . . . , 84	MLR
		5km	MSM	FT=24, 27, . . . , 39/78	
		5km	MEPS	FT=24, 27, . . . , 39	
Mean precipitation amount over 1 hour	Grids	5km	LFM	FT=1, 2, . . . , 10	LAF & FBC
Maximum precipitation amount over 1 hour	Grids	5km	LFM	FT=1, 2, . . . , 10	SLR & PPM
Probability of heavy rain over 3 hours > 100mm/3h, 150mm/3h	Grids	5km	MSM	FT=3, 4, . . . , 39/51***	LR
			MEPS	FT=3, 4, . . . , 39	
Probability of precipitation over 6 hours > 1mm/6h	Grids	20km	GSM	FT=9, 15, . . . , 81	KF
		5km	MSM	FT=6, 12, . . . , 36/39/75****	
Maximum temperature in the daytime (09-18 local time)	Points	AMeDAS	GSM	Today to 3 days after	KF
			MSM		
			MEPS	Today and tomorrow	
Minimum temperature in the morning (00-09 local time)	Points	AMeDAS	GSM	Today to 3 days after	KF
			MSM		
			MEPS	Today and tomorrow	
Time-series temperature	Points	AMeDAS	GSM	FT=3, 4, . . . , 84	KF
			MSM	FT=1, 2, . . . , 39/78	
			LFM	FT=1, 2, . . . , 10	
			MEPS	FT=1, 2, . . . , 39	
Temperature	Grids	5km	GSM	FT=3, 4, . . . , 84	KF
		5km	MSM	FT=1, 2, . . . , 39/78	
		2km	LFM	FT=1, 2, . . . , 10	
		5km	MEPS	FT=1, 2, . . . , 39	
Time-series humidity	Points	SYNOP	GSM	FT=3, 4, . . . , 84	KF
			MSM	FT=1, 2, . . . , 39/78	
Wind speed and direction	Points	AMeDAS	GSM	FT=3, 6, . . . , 84	KF & FBC
			MSM	FT=1, 2, . . . , 39/78	
			MEPS	FT=1, 2, . . . , 39	
Maximum wind speed and direction over 3hours	Points	AMeDAS	GSM	FT=3, 6, . . . , 84	KF & FBC
			MSM	FT=3, 6, . . . , 39/78	
			MEPS	FT=3, 6, . . . , 39	
Daily minimum humidity	Points	SYNOP	GSM	Today to 3 days after	NN
			MSM		
Snowfall amount over 6,12 hours	Points	AMeDAS	GSM	FT=9, 12, . . . , 84	NN & FBC
			MSM	FT=6, 9, . . . , 39/78	
Snowfall amount over 24 hours	Points	AMeDAS	GSM	FT=27, 30, . . . , 84	NN & FBC
			MSM	FT=24, 27, . . ., 39/78	
Snowfall amount over 3,6 hours	Grids	2.5km	LFM	FT=3, 4, . . . , 10	DM
Snowfall amount over 3,6,12 hours	Grids	5km	GSM	FT=6, 9, . . . , 84	DM
			MSM	FT=3, 6, . . . , 39/78	
			MEPS	FT=3, 6, . . . , 39	
Snowfall amount over 24 hours	Grids	5km	GSM	FT=27, 30, . . . , 84	DM
			MSM	FT=24, 27,. . . , 39/78	
			MEPS	FT=24, 27,. . . , 39	
Probability of thunderstorm over 3 hours	Grids	20km	GSM	FT=6, 9, . . . , 84	LR & LAF
			MSM	FT=6, 9, . . . , 39/78	
			MEPS	FT=6, 9, . . . , 39	
Visibility	Grids	20km	GSM	FT=3, 6, . . . , 84	DM
		5km	MSM	FT=3, 6, . . . , 39/78	

∗ NN: neural network, KF: Kalman filter, FBC: frequency bias correction, MLR: multiple linear regression, DM: diagnostic method, LR: logistic regression, LAF: lagged averaged forecast method, SLR: simple linear regression, PPM: perfect prognosis method

∗∗ 39 hours (03, 06, 09, 15, 18, 21 UTC initial times), 78 hours (00, 12 UTC initial times)

∗∗∗ 39 hours (03, 06, 09, 15, 18, 21 UTC initial times), 51 hours (00, 12 UTC initial times)

∗∗∗∗ 36 hours (03, 09, 15, 21 UTC initial times), 39 hours (06, 18 UTC initial times), 75 hours (00, 12 UTC initial times)

Eq. (4.6.2) indicate the variables are stochastic variables at τ. $\boldsymbol{X}_{\tau+1/\tau}$ is determined from the previous estimate $\boldsymbol{X}_{\tau/\tau-1}$ and the forecast error:

$$\boldsymbol{X}_{\tau+1/\tau} = \boldsymbol{X}_{\tau/\tau} \tag{4.6.4}$$
$$= \boldsymbol{X}_{\tau/\tau-1} + \boldsymbol{K}_{\tau}(y_{\tau} - \boldsymbol{c}_{\tau}\boldsymbol{X}_{\tau/\tau-1}) \tag{4.6.5}$$

where $\boldsymbol{K}_{\tau}$ is the Kalman gain and can be written as follows:

$$\boldsymbol{K}_{\tau} = \boldsymbol{Q}_{\tau/\tau-1}\boldsymbol{c}_{\tau}^{\mathrm{T}}(\boldsymbol{c}_{\tau}\boldsymbol{Q}_{\tau/\tau-1}\boldsymbol{c}_{\tau}^{\mathrm{T}} + D_{\tau})^{-1} \tag{4.6.6}$$

$\boldsymbol{Q}$, the covariance of $\boldsymbol{X}$, is updated as follows:

$$\boldsymbol{Q}_{\tau+1/\tau} = \boldsymbol{Q}_{\tau/\tau} + \boldsymbol{U}_{\tau} \tag{4.6.7}$$
$$= \boldsymbol{Q}_{\tau/\tau-1} - \boldsymbol{K}_{\tau}\boldsymbol{c}_{\tau}\boldsymbol{Q}_{\tau/\tau-1} + \boldsymbol{U}_{\tau} \tag{4.6.8}$$

Eq. (4.6.4) and Eq. (4.6.7) are derived from Eq. (4.6.2) and Eq. (4.6.3), respectively.

Finally, the forecast value is calculated with the updated coefficients and predictors at $\tau + 1$;

$$y_{\tau+1/\tau} = \boldsymbol{c}_{\tau+1}\boldsymbol{X}_{\tau+1/\tau} \tag{4.6.9}$$

For some forecast parameters (such as temperature), the predictand y is defined as the difference between NWP output and observations, while for others (such as precipitation amount), y is the observation itself.

In the forecast guidance system with KF, D_{τ} in Eq. (4.6.6) and $\boldsymbol{U}_{\tau}$ in Eq. (4.6.8) are treated as empirical parameters for control of the adaptation rate.

4.6.2.2 Frequency Bias Correction

With KF, the most likely estimation of the predictand that minimizes the expected root mean square error is obtained. However, the output has a tendency toward lower-than-actual frequency for forecasting of relatively rare events such as strong wind and heavy rain. To compensate for this unfavorable characteristic, a frequency bias correction scheme is applied to the KF output.

The basic concept involves multiplying the estimation of KF, y, by a correction factor $F(y)$ to obtain the final output y^b:

$$y^b = y \cdot F(y)$$

To determine $F(y)$, certain thresholds t^i are set to span the given observation data. The corresponding thresholds f^i for the forecast dataset are then calculated so that the number of observation data smaller than t^i approximates to that of forecast data smaller than f^i. Finally, the correction factors are computed as follows:

$$F(f^i) = t^i / f^i$$
$$F(y) \text{ for } f^i < y < f^{i+1} \text{ is linearly interpolated between } F(f^i) \text{ and } F(f^{i+1}).$$

Since KF is an adaptive method, f^i is also updated each time the observation y_{τ} corresponding to the estimates of KF $y_{\tau/\tau-1}$ is available. The update procedure is as follows:

$$f^i_{\tau+1} = \begin{cases} f^i_{\tau}(1+\alpha) & \text{if} \quad y_{\tau} < t^i \quad \text{and} \quad y_{\tau/\tau-1} > f^i \\ f^i_{\tau}(1-\alpha) & \text{if} \quad y_{\tau} > t^i \quad \text{and} \quad y_{\tau/\tau-1} < f^i \\ f^i_{\tau} & \text{otherwise} \end{cases}$$

where α is an empirical parameter used to determine the adaptation rate.

4.6.2.3 Guidance Based on Kalman Filtering (3-hour Precipitation Amount)

In this subsection, we will introduce the guidance of 3-hour precipitation amount derived from the NWP model (GSM) output. The predictand is the observed 3-hour cumulative precipitation amount averaged within a 20×20km square. The following parameters derived from NWP output are used as predictors.

1. NW85: NW − SE component of wind speed at 850hPa
2. NE85: NE − SW component of wind speed at 850hPa
3. SSI: Showalter's stability index (between 850 and 500hPa)
4. OGES: Orographic precipitation index
5. PCWV: Precipitable water content × Wind speed at 850hPa × Vertical p-velocity at 850hPa
6. QWX: Σ (Specific humidity × Vertical p-velocity × Relative humidity) between 1000 and 300hPa
7. EHQ: Σ (Excess from reference humidity × Specific humidity × Depth of wet layer) between 1000 and 300hPa
8. DXQV: Precipitation index in winter synoptic pattern
9. FRR3: Mean precipitation amount over three hours based on the GSM

In the precipitation forecasts shown in Figure 4.6.1, the GSM predicted no precipitation in area M, while actual observation showed light precipitation as also indicated by the guidance. Examination of the coefficients for the predictors shows an orographic effect relating to the OGES enhanced precipitation amount in this area.

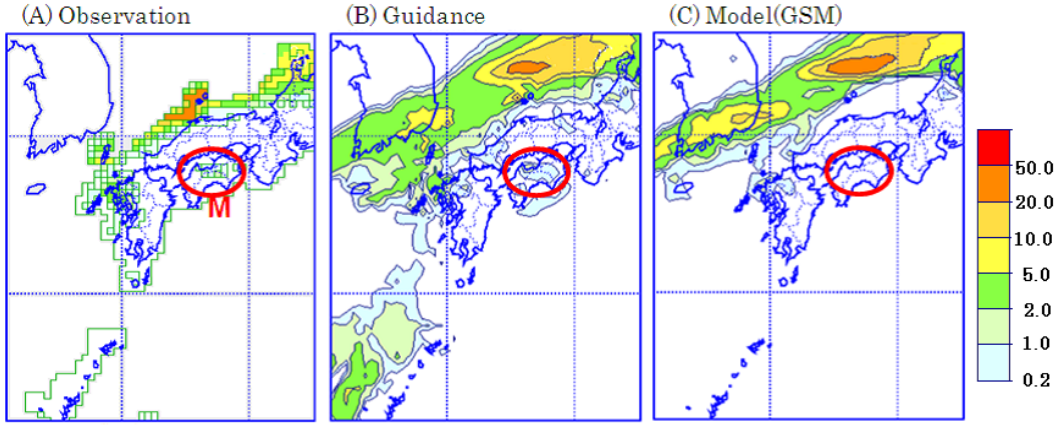


Figure 4.6.1: Mean precipitation amount over three hours: (A) observation, (B) guidance forecast, (C) NWP model (GSM) forecast

4.6.3 Guidance Based on a Neural Network

4.6.3.1 Neural Network

The neural network (NN) approach involves machine learning to support analysis of non-linear relationships between predictors and predictands (Yanagino and Takada 1995). In this method, multiple layers of neurons are linked to construct a hierarchical neural network as shown in Figure 4.6.2. The first layer is called the input layer and the last layer is called the output layer. Those between them are called the hidden layers.

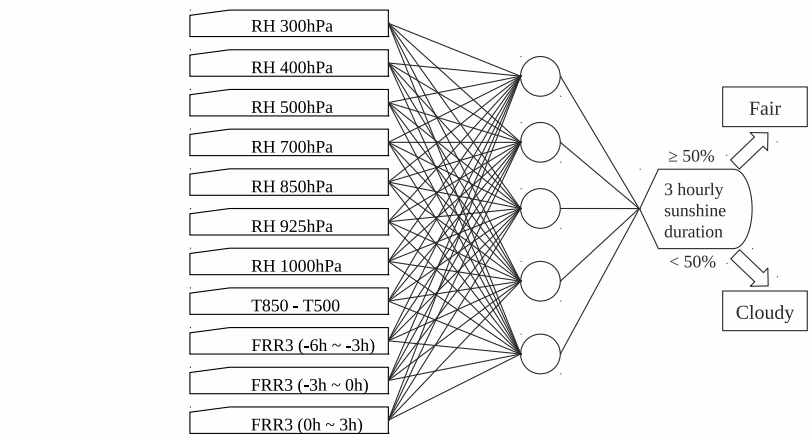


Figure 4.6.2: Neural network for fair/cloudy determination. RH: relative humidity, FRR3: precipitation over 3 hours, T850: temperature at 850hPa, T500: temperature at 500hPa

A signal put into the input layer is propagated to the next layer via inter-neuron connections. The signal is subjected to simple processing by the neurons of the receiving layer prior to its propagation to the next layer. This process is repeated until the signal reaches the output layer.

A schematic diagram of a neuron is shown in Figure 4.6.3. The input of each neuron is a weighted sum of the outputs of other neurons, and the output is a function of its input. This is called an activation function, and a sigmoid function as shown in Figure 4.6.4 is usually used for this purpose.

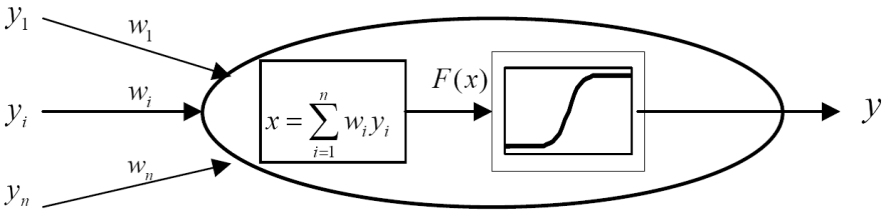


Figure 4.6.3: Schematic representation of a neuron

NN weights are iteratively adjusted using numerous sets of input/output data. The most popular adjustment method involves the back-propagation of error algorithm described as follows:

1. First, weights are initialized with randomized values.
2. The network calculates output values using a given set of input values.
3. Weights are adjusted to make the NN output close to the supervisor data (correct values of the output variable).
4. The processes of 2 and 3 are iterated until the error measure falls below a specified value or the number of iterations reaches a specified maximum.

Even in NN, the output has a tendency toward lower-than-actual frequency for forecasting of relatively rare events such as heavy snowfall. To compensate for this unfavorable characteristic in snowfall amount guidance,

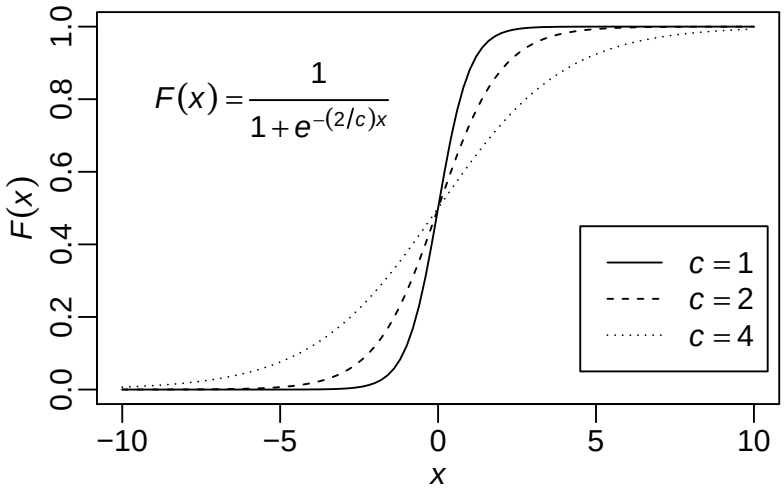


Figure 4.6.4: Examples of the sigmoid function

the frequency bias correction scheme is applied to the NN output using the same method for KF output (see Subsection 4.6.2.2).

4.6.3.2 Example of Guidance Based on a Neural Network (Categorized Weather)

In the forecast guidance system, a neural network model is constructed at each grid or observation point from sets of NWP output and observed weather elements. Categorized weather is one of the forecast guidance parameters to which the NN is applied. Figure 4.6.5 shows an example of output categorized weather guidance, in which an NN model is used to derive sunshine duration, which is in turn used to determine non-precipitation weather categories (fair or cloudy). The NN is constructed for each AMeDAS station, and output values (three-hourly sunshine durations) are interpolated to grid points. Precipitation weather categories (rain, sleet, snow) are determined from precipitation amount guidance and precipitation type guidance. The constitution of the sunshine duration NN model is shown in Figure 4.6.2, and its characteristics are summarized as follows:

1. The model incorporates a three-layer feed-forward neural network.

2. A linear activation function is used in the output layer and sigmoid activation functions are used in the hidden layer.

3. In the learning processes, NWP output is used as input data and sunshine durations observed at each AMeDAS point are used as supervisor data.

4. The weights of the network are modified when the observation data are obtained.

4.6.4 Utilization of Guidance at Forecasting Offices

Guidance is provided to observatories for drafting of weather forecasts in editing software (see Figure 4.6.6). Time-series of weather categorization, six-hour probability of precipitation (PoP) and temperature are revised on the display in consideration of current weather conditions and empirical expertise for forecast production. The processed data are then composed to create a forecast bulletin and provided to users.

An algorithm incorporating the steps shown below is used to draft weather forecast bulletins automatically.

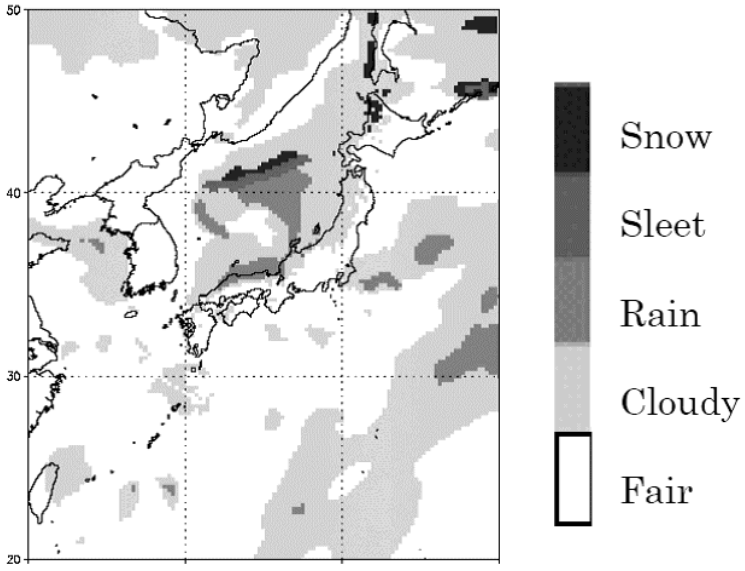


Figure 4.6.5: Sample output of categorized weather guidance

Table 4.6.2: Sample of the algorithm for weather forecast bulletin drafting

Sequence of 3-hourly categorized weather*								Draft of a weather forecast bulletin
0 - 3	3 - 6	6 - 9	9 -12	12-15	15-18	18-21	21-24	
F	F	F	F	C	F	F	F	Fair
R	R	R	R	R	S	S	S	Rain, snow from the evening
C	R	F	R	C	F	R	C	Cloudy, occasional rain
C	R	C	C	C	C	R	C	Cloudy, rain in the morning and the evening

* F:Fair C:Cloudy R:Rain S:Snow

1. Three-hourly dominant weather categories are derived from the majority of weather categorization on grids in the forecast area.
2. The weather forecast bulletin for the day is derived from the sequence of three-hourly dominant weather categories over the forecast area. Examples of the algorithm are shown in Table 4.6.2.

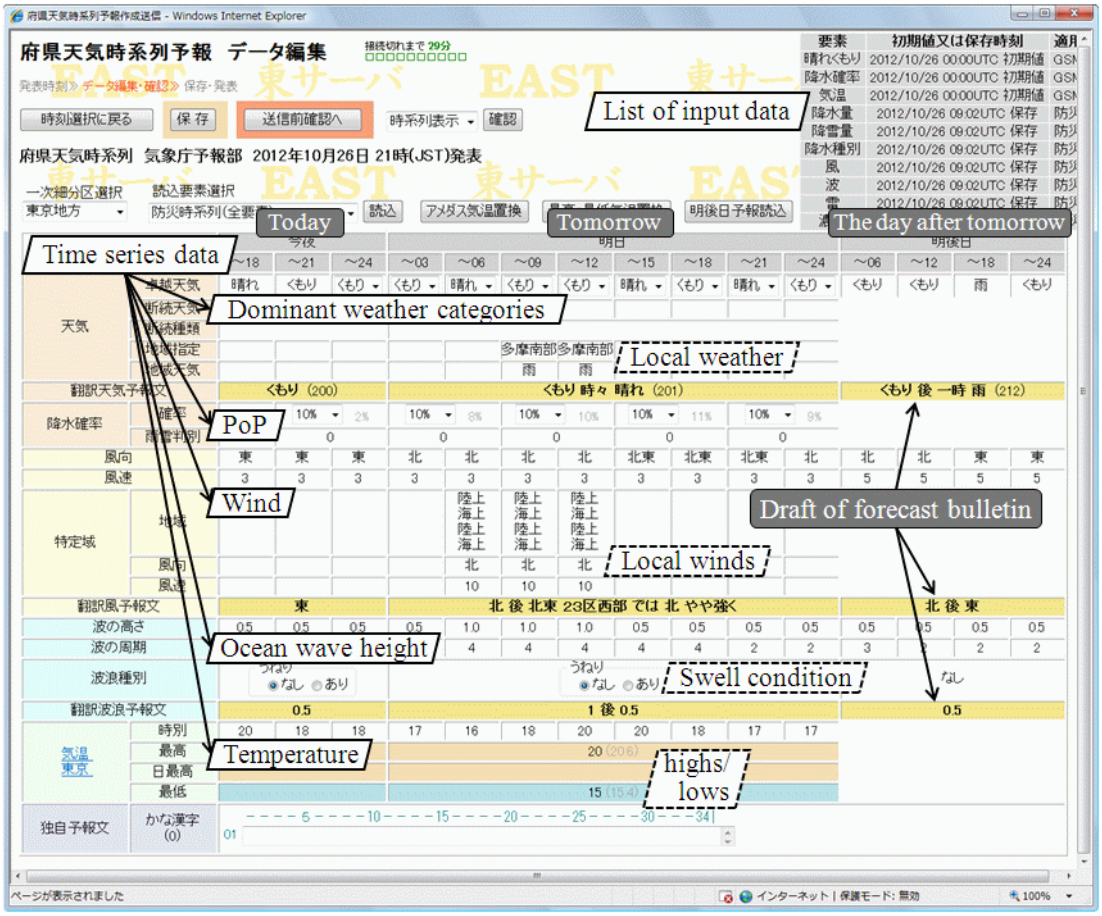


Figure 4.6.6: Sample data entry screen in forecast editing software

4.7 Application Products for Aviation Services

4.7.1 Aerodrome Forecast Guidance

Aerodrome forecast guidance (TAF guidance) is derived from MSM output eight times a day and from MEPS four times a day, providing hourly predictions up to 39 hours ahead (78 hours at 00 and 12UTC for MSM initial times) as detailed in Table 4.7.1.

Table 4.7.1: Parameters of TAF guidance. (for 90 airports in Japan)

Parameters	TAF guidance
Visibility	Minimum and mean visibility during hourly periods Probability of minimum visibility < 5km and 1.6km during 3-hour periods
Cloud	Cloud density and 3-layer height at minimum ceiling during hourly periods Probability of minimum ceiling < 1,000 and 600ft during 3-hour periods
Weather	Categorized weather every hour
Temperature	Maximum temperature in the daytime, minimum temperature in the morning and temperatures every hour
Wind	Wind speed and direction every hour Wind speed and direction of hourly maximum peak wind
Gust	Probability of gust during 3 hours Gust speed and direction of hourly maximum peak gust
Thunderstorm	Probability of thunder during 3 hours
Snow	Snowfall amount during 3 hours

4.7.1.1 Visibility

Visibility guidance (VIS) is based on statistical NWP interpretation for individual airports using the following MSM output:

- No precipitation: $(1 - RH)^{1/2}$, $Qc^{1/2}$, where RH is surface relative humidity ($0 \sim 1$), Qc is cloud water content near surface(kg/kg).
- Rain: $RR^{1/2}$, $(1 - RH)^{1/2}$, $Qc^{1/2}$, where RR is precipitation amount over an hour (mm).
- Snow: $RR^{1/2}$, $(1 - RH)^{1/2}$, $\min(VV, 15) \times \max(\min(T, 0), -10)$, where VV is surface wind speed (m/s), T is surface temperature (°C).

Selection of the relevant equation content is based on the category predicted in weather guidance (described later), with coefficients adapted via Kalman filtering (see Subsection 4.6.2.1) with predictors and METAR report data. Frequency bias correction (see Subsection 4.6.2.2) is then applied to equation output.

4.7.1.2 Cloud

TAF cloud guidance involves statistical interpretation of NWP output. First, cloud amounts in each of 38 layers (0, 100, ..., 1000, 1500, ..., 5000, 6000, ..., 10000, 12000, ..., 30000 ft) are calculated using an NN (see Subsection 4.6.3), and the lowest three cloud layers are then extracted as with METAR reports. The input data (predictors) are relative humidity, the lapse rate between the surface and 925hPa, and the precipitation amount. Frequency bias correction (see Subsection 4.6.2.2) is applied to outputs from NN.

Guidance on minimum ceiling probability is based on statistical interpretation of NWP output, with predicting over three-hour periods for values below 1,000 and 600ft. The predictors are precipitation amount over three hours (total precipitation or snow water equivalent depending on the airport), the lapse rate between the surface and 925hPa, relative humidity, the E-W component of wind speed, the S-N component of wind speed, cloud amount, cloud ice content and cloud water content. Logistic regression (Agresti 2002) was introduced for prediction of minimum ceiling probability in December 2010.

4.7.1.3 Weather

Weather guidance predicts weather conditions based on a diagnostic method for the interpretation of MSM output into categories (fine, cloudy, rainy, snowy and precipitation intensity) (JMA 1997). To determine the precipitation type (rain or snow), hourly temperature guidance is used instead of MSM temperature.

4.7.1.4 Wind and Temperature

Wind and temperature guidance are calculated using the same methods as guidance for short-range forecasting (see Section 4.6).

4.7.1.5 Gust Winds

Wind gust guidance predicts the probability of wind gust during three-hour periods as well as the speed and direction of the hourly maximum peak gust. Wind gust guidance uses Kalman filtering, frequency bias correction and logistic regression to predict each variable. Predictors are gust speed calculated by the MSM, surface wind speed, maximum wind speed in the boundary layer, vertical wind shear between the surface and the boundary layer, SSI, and vertical p-velocity at 925hPa. TAF gust guidance was introduced in December 2012.

4.7.1.6 Thunderstorms

Probability of thunderstorm (PoT) guidance predicts the probability of thunderstorm during three-hour periods around airports. PoT is predicted using logistic regression. Six predictors are selected from twelve potential predictors, in which three predictors, SSI, CAPE, and precipitation amounts over the three-hour period, are always selected. PoT guidance was introduced in May 2007.

4.7.1.7 Snow

Snowfall amounts at airports are calculated from the grid-type snowfall amount guidance shown in Table 4.6.1. Snowfall amount guidance for TAF was introduced in October 2015.

4.7.2 Products for Domestic Area Forecast

4.7.2.1 Gridded Values of Significant Weather

For these values, JMA applies variables including turbulence, icing, CB clouds, tropopause height and vertical wind shear derived from MSM and LFM outputs along with common meteorological variables such as temperature, wind and humidity for domestic area forecasts and SIGMET information. These purely aviation-oriented datasets are called SIGGVs (gridded values of significant weather), with the specifications listed in Table 4.7.2. Vertical wind shear (VWS), which is an indicator of clear air turbulence (CAT), is calculated in kt/1000ft. TBindex (Kudo 2011) is a combined measure of multiple turbulence indices for predicting CAT, mountain waves and cloud-related turbulence. CB cloud amount and top height calculations are based on the parcel method. The icing parameter is an indicator of aircraft icing derived from an empirical equation incorporating temperature and dew-point temperature. As illustrated in Figure 4.7.1, SIGGVs derived from MSM and LFM outputs are visualized on terminals at aviation forecast offices. SIGGVs from MSM data are distributed in binary form in addition to the fax charts detailed below.

Table 4.7.2: SIGGV specifications

Forecast time	T=0-39, 1 hourly for MSM, T=0-10, 1 hourly for LFM
Grid coordinate	Polar Stereographic, 111 × 93 for MSM, 349 × 293 for LFM
Parameters	U, V, W, T, RH, Psea, Rain, CB cloud amount, CB top height, Tropopause height, Icing, VWS, TBindex

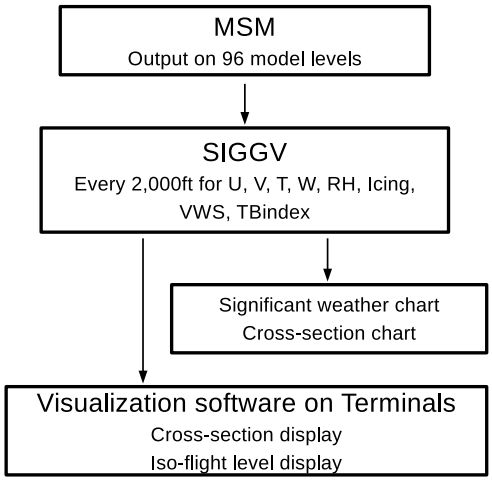


Figure 4.7.1: Data flow of MSM-derived products for domestic area forecasting

4.7.2.2 Domestic Significant Weather Chart

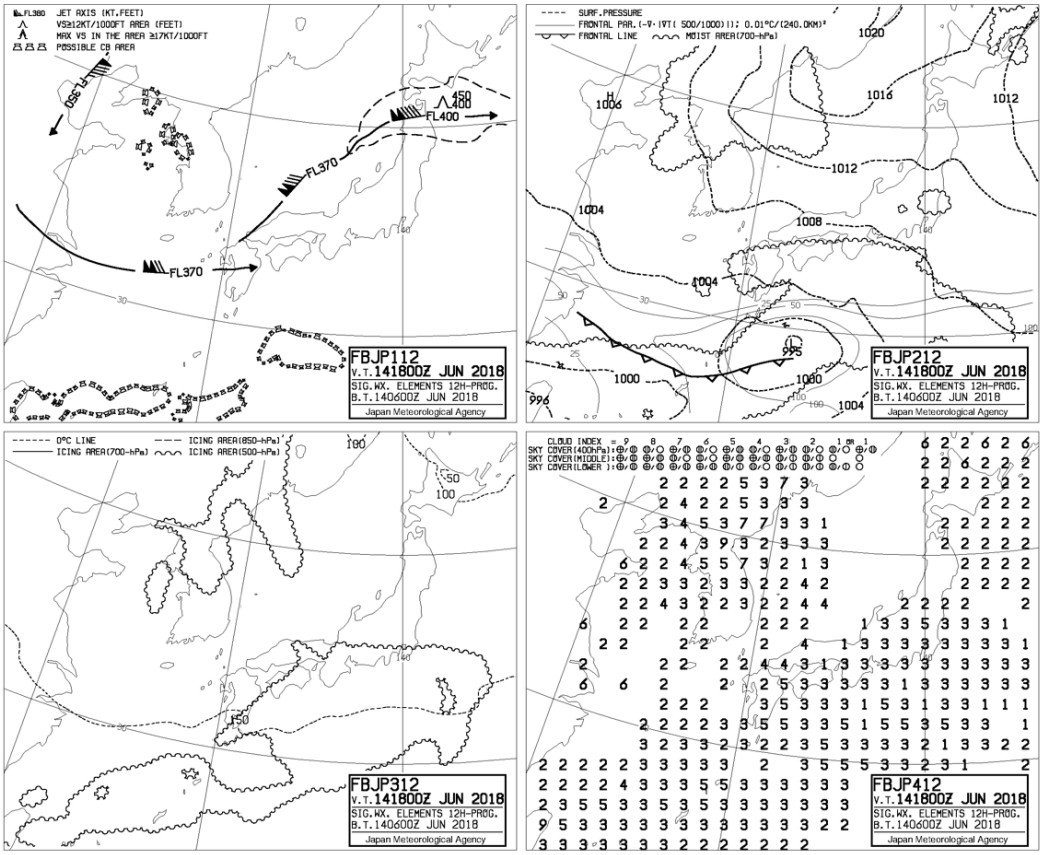


Figure 4.7.2: Sample domestic significant weather chart

This chart shows 12- hour forecast fields of the parameters listed below in four panels: (Figure 4.7.2)

- Upper-left:
  - Jet stream axes.
  - Possible CAT areas.
  - Possible CB areas.
- Lower-left:
  - Contours of 0°C height.
  - Possible icing areas at 500, 700 and 850hPa based on the -8 D method (Godske 1957)
- Upper-right:
  - Contours of sea level pressure.
  - Moist areas at 700 hPa.
  - Front parameters $DDT = -\nabla_n|\nabla_n T|$, where T is mean temperature below 500hPa and ∇_n denotes the horizontal gradient perpendicular to the isotherms.

– "NP fronts" drawn along the maxima of DDT.

- Lower-right:
  - Cloud indices indicating low, middle and upper cloud amounts.

4.7.2.3 Domestic Cross-section Chart

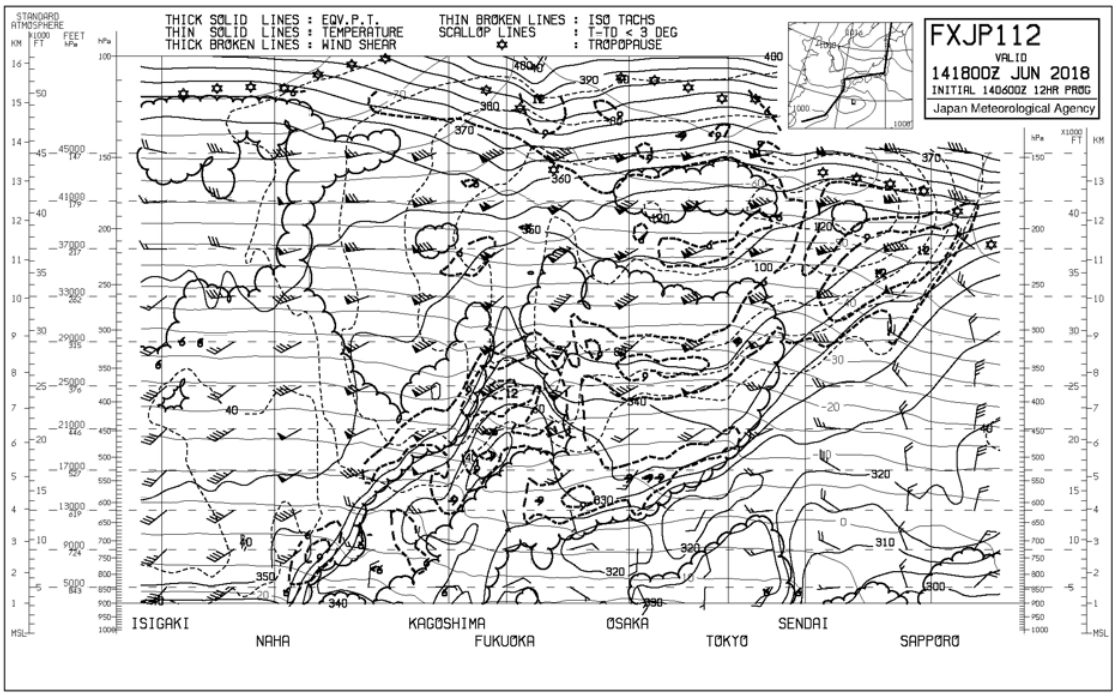


Figure 4.7.3: Sample domestic cross-section chart. Only the lower part of the fax, corresponding to the 12-hour forecast, is shown.

This chart shows 6- and 12-hour forecast fields along the major domestic route, which illustrates temperature, equivalent potential temperature, wind barbs and isotachs, moist areas, vertical wind shear and tropopause height (Figure 4.7.3).

4.7.3 Products for International Area Forecast

Global Grid Point Values are derived from the GSM four times a day and distributed in thinned GRIB code, a format compatible with products from the World Area Forecast Centers (WAFCs). In addition to the parameters included in WAFC products, TBindex, Icing, VWS and Cb cloud top height are derived using the method used for domestic SIGGV (see Subsection 4.7.2).

JMA produces 13 significant weather (SIGWX) charts and 18 wind and temperature (WINTEM) charts based on significant weather data provided by WAFCs.

4.8 Ensemble Prediction System Products

4.8.1 EPS Products for One-week Forecasting

To assist forecasters in issuing one-week weather forecasts, ensemble mean products are made from EPS output.

An example of an ensemble chart showing average mean sea level pressure and precipitation is shown in Figure 4.8.1.

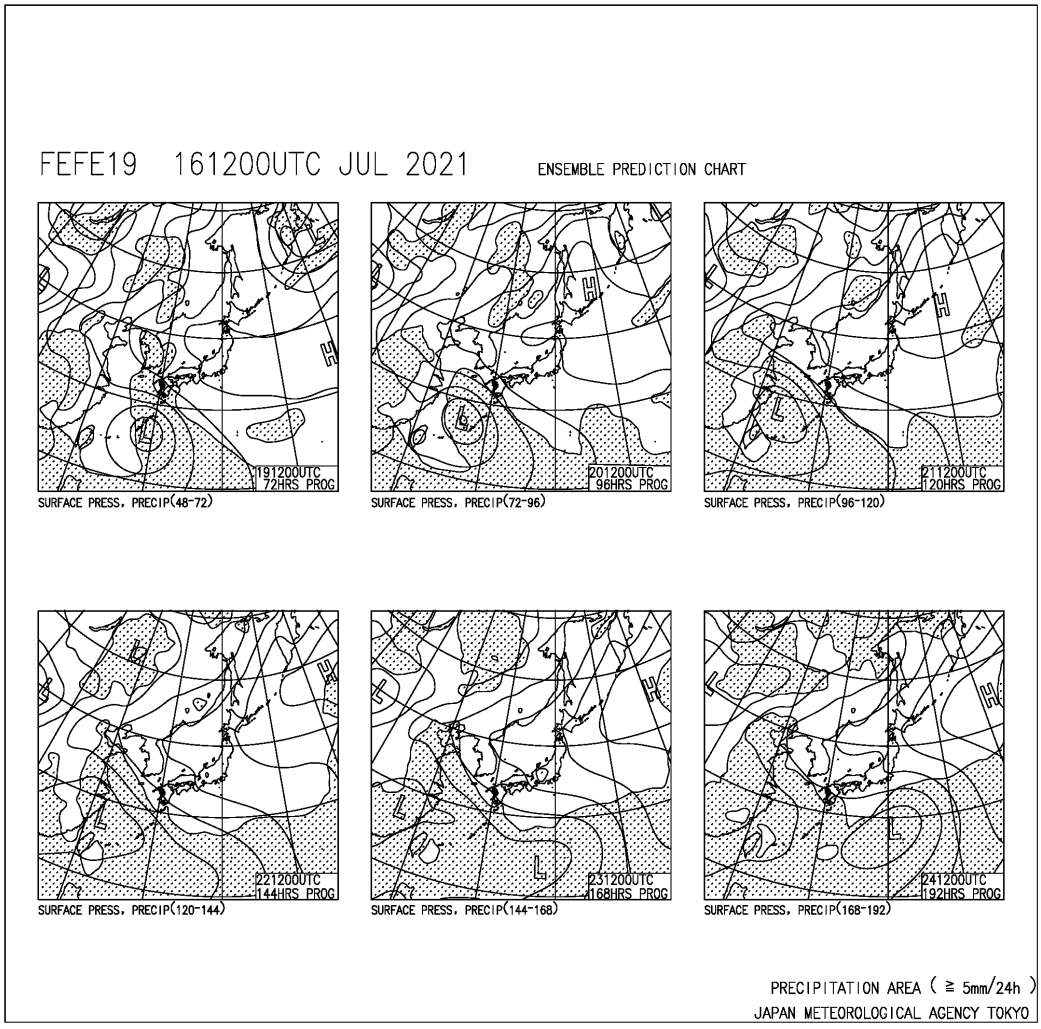


Figure 4.8.1: Ensemble prediction chart showing average mean sea level pressure and precipitation from day 2 to day 7. This schematic representation is produced by averaging with all members.

4.8.2 EPSs Products for One-month and Seasonal Forecasting

4.8.2.1 Standard Products

The operational forecast and hindcast products from the Global and Seasonal EPSs detailed below are provided via the World Meteorological Center (WMC) Tokyo page on the Tokyo Climate Center (TCC) website[5]. Details of map products are shown in Table 4.8.1 and Table 4.8.2.

- Ensemble mean maps
- Ensemble spread maps (for operational forecasts only)
- Verification maps and scores
- SST Index time-series representations

Table 4.8.1: Map products for one-month forecasts on the TCC website

	Forecast time	Parameters
Ensemble mean	Averages of days 3 – 9, 10 – 16, 17 – 30, 3 – 30	Geopotential height at 500 hPa and related anomaly Temperature at 850 hPa and related anomaly Sea level pressure and related anomaly Stream function at 200 hPa and related anomaly Stream function at 850 hPa and related anomaly Velocity potential at 200 hPa and related anomaly Precipitation and related anomaly Temperature at 2 m and related anomaly Sea surface temperature (initial anomaly)

Table 4.8.2: Map products for three-month and warm/cold-season forecasts on the TCC website

	Forecast time	Parameters
Ensemble mean, related anomaly and spread	Three-month forecasts: averages of first/second/third months and all three Warm/cold season forecasts: Three-months averages(JJA and DJF)	Geopotential height at 500 hPa, related anomaly and spread Temperature at 850 hPa, related anomaly and spread Sea level pressure, related anomaly and spread Stream function at 200 hPa, related anomaly and spread Stream function at 850 hPa, related anomaly and spread Wind (u, v) anomaly at 850 hPa Velocity potential at 200 hPa, related anomaly and spread Precipitation, related anomaly and spread Temperature at 2 m, related anomaly and spread Sea surface temperature and related anomaly

4.8.2.2 Gridded Datasets

The operational forecast and hindcast gridded datasets of the Global and Seasonal EPSs detailed below are provided to registered users via the WMC Tokyo web-page (see Table 4.8.3 and Table 4.8.4 for details, with further information on Global EPS for one-month Prediction in Table 3.3.1). The product is initialized at 12 UTC based on a 50-member lagged ensemble with two initial times. Table 3.4.1 gives further details on Seasonal EPS for six-month Prediction.

- Global EPS

[5] https://www.data.jma.go.jp/tcc/tcc/products/model/index.html

  - Daily mean ensemble statistics (for operational forecasts only)
  - Daily mean forecast of individual ensemble members

- Seasonal EPS
  - Monthly and 3-monthly mean ensemble statistics (for operational forecasts only)
  - Monthly mean forecast of individual ensemble members

Table 4.8.3: Global 1.25-degree gridded data products (Global EPS for one-month Prediction)

Details		Levels (hPa)	Ensemble mean / members	Anomaly
Ensemble statistics	Sea level pressure*	-	○	○
	Daily mean precipitation	-	○	○
	Temperature*	Surf, 850, 700	○	○
	Relative humidity	850	○	
	Geopotential height*	500, 100	○	○
	Wind (u,v)	850, 200	○	
	Stream function	850, 200	○	○
	Velocity potential	200	○	○
Individual ensemble members	Sea level pressure*	-	○	
	Daily mean precipitation	-	○	
	Temperature*	Surf, 1000, 850, 700, 500, 300, 200, 100	○	
	Relative humidity	1000, 850, 700, 500, 300	○	
	Geopotential height*	1000, 850, 700, 500, 300, 200, 100	○	
	Wind (u,v)	1000, 850, 700, 500, 300, 200, 100	○	
	Stream function	850, 200	○	
	Velocity potential	200	○	

∗ Geopotential height, sea level pressure and temperature are calibrated by subtracting the systematic error from the direct model output.

Table 4.8.4: Global 1.25-degree gridded data products (Seasonal EPS for six-month Prediction)

Details		Levels (hPa)	Ensemble mean / members	Anomaly	Spread
Ensemble statistics	Sea level pressure*	-	○	○	○
	Daily mean precipitation	-	○	○	○
	Sea surface temperature*	-	○	○	
	Sea ice concentration	-	○	○	
	Temperature*	Surf, 850	○	○	○
	Geopotential height*	500	○	○	○
	Wind (u,v)	850, 200	○	○	○
Individual ensemble members	Sea level pressure*	-	○	○	
	Daily mean precipitation	-	○	○	
	Sea surface temperature*	-	○	○	
	Sea ice concentration	-	○	○	
	Temperature*	Surf, 850, 500, 200	○	○	
	Relative humidity	850	○	○	
	Geopotential height*	850, 500, 300, 200, 100	○	○	
	Wind (u,v)	850, 500, 200	○	○	

∗ Geopotential height, sea level pressure, temperature and sea surface temperature are calibrated by subtracting the systematic error from the direct model output.

4.8.2.3 El Ninõ Outlook

Outlooks of sea surface temperature deviations in the Pacific and the Indian Ocean produced from the Seasonal EPS are provided via the El Ninõ Monitoring and Outlook web-page on the TCC website (Figure 4.8.2).

4.8.2.4 Probabilistic Forecast Products for Seasonal Forecasts

To support NMHSs (National Meteorological and Hydrological Services) in their production of seasonal forecasts, probabilistic forecast products with three-categories (e.g., above-, near-, and below-normal) are produced from the Seasonal EPS and provided via the WMC Tokyo web-page (Figure 4.8.3).

4.8.2.5 Forecast Products in Support of Early Warnings for Extreme Weather Events

To support NMHSs in their production of early warnings for extreme weather events, Extreme Forecast Index (EFI) warning maps and meteograms for major cities are produced from the Global EPS and provided to registered users via the WMC Tokyo web-page (Figure 4.8.4 and Figure 4.8.5). The EFI is a measure of EPS forecast deviation from climatological probability distribution (Lalaurette 2003). JMA uses a revised version of the EFI to add weight for the tails of probability distribution (Zsótér 2006).

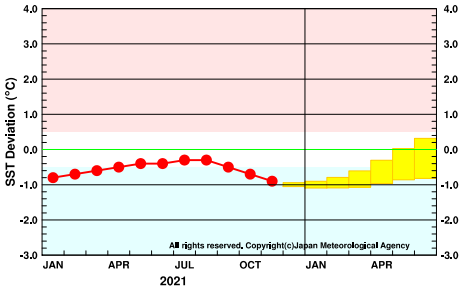


Figure 4.8.2: Five-month running mean of SST deviation for NINO.3. Red dots indicate analysis values, and yellow boxes indicate values predicted by the Seasonal EPS. Each box denotes the range in which the value is expected to be included with a probability of 70% or more.

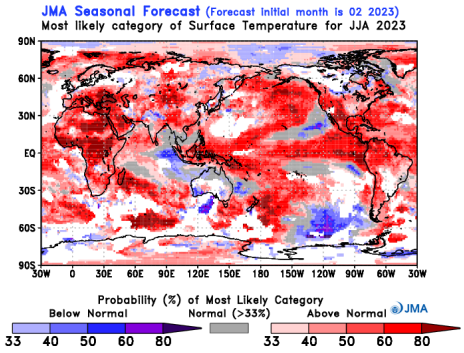


Figure 4.8.3: Probabilistic forecast map of surface air temperature for seasonal forecasting. Probabilities are estimated using numerical guidance with application of the Model Output Statistics (MOS) technique based on hindcast results.

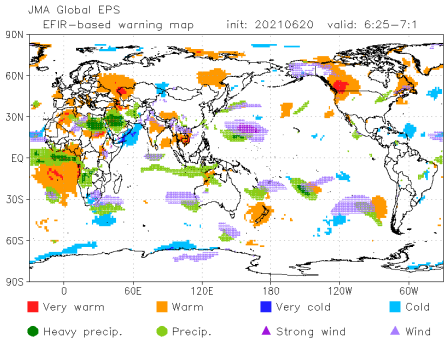


Figure 4.8.4: EFI warning map for temperature, precipitation and wind. Pale symbols indicate grids where the EFI is above 0.5, and dark symbols indicate grids where the EFI is above 0.8.

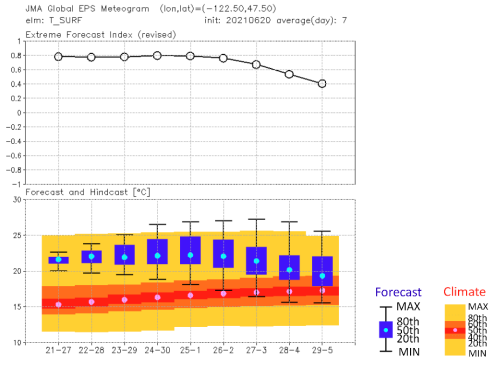


Figure 4.8.5: Meteogram of the Global EPS for temperature. The upper graph is an EFI timeseries representation, and the lower one is a timeseries representation of forecast ranges with cold color box-and-whisker plots and climatological values with warm shading.

4.9 Atmospheric Angular Momentum Functions

Atmospheric Angular Momentum (AAM) functions were proposed to support evaluation of the earth's rotational variations based on precise estimation of variations in atmospheric angular momentum. For monitoring of atmospheric effects associated with the earth's rotation, JMA sends AAM products to NCEP (a sub-bureau of the International Earth Rotation Service (IERS)) via GTS. The AAM functions are expressed as follows (Barnes *et al.* 1983):

$$\chi_1 = -1.00\left[\frac{r^2}{(C-A)g}\right]\int P_S \sin\phi\cos\phi\cos\lambda\, dS - 1.43\left[\frac{r}{\Omega(C-A)g}\right]\iint (u\sin\phi\cos\lambda - v\sin\lambda)\, dPdS, \tag{4.9.1}$$

$$\chi_2 = -1.00\left[\frac{r^2}{(C-A)g}\right]\int P_S \sin\phi\cos\phi\sin\lambda\, dS - 1.43\left[\frac{r}{\Omega(C-A)g}\right]\iint (u\sin\phi\sin\lambda + v\cos\lambda)\, dPdS, \tag{4.9.2}$$

$$\chi_3 = -0.70\left[\frac{r^2}{Cg}\right]\int P_S \cos^2\phi\, dS - 1.00\left[\frac{r}{\Omega Cg}\right]\iint u\cos\phi\, dPdS. \tag{4.9.3}$$

In Eq. (4.9.1) to Eq. (4.9.3), P is pressure, $\int dS$ is the surface integral over the globe, (ϕ, λ) are latitude and longitude, u and v are the eastward and northward components of wind velocity, P_S is surface pressure, g is the mean acceleration of gravity, r is the mean radius of the earth, C is the polar moment of inertia of the solid earth, A is the equatorial moment of inertia, and Ω is the mean angular velocity of the earth.

The functions χ_1 and χ_2 represent equatorial components, and the function χ_3 is the axial component. All components are non-dimensional. The first term of each one is a pressure term related to the redistribution of air masses. The second is a wind term related to the relative angular momentum of the atmosphere.

Variations in AAM functions calculated from JMA global analysis data have been reported to correspond closely to variations in the earth's rotation. Figure 4.9.1 shows seasonal variations in observed earth rotation and atmospheric relative angular momentum (the wind term of χ_3) calculated by the National Astronomical Observatory of Japan (Naito and Kikuchi 1992).

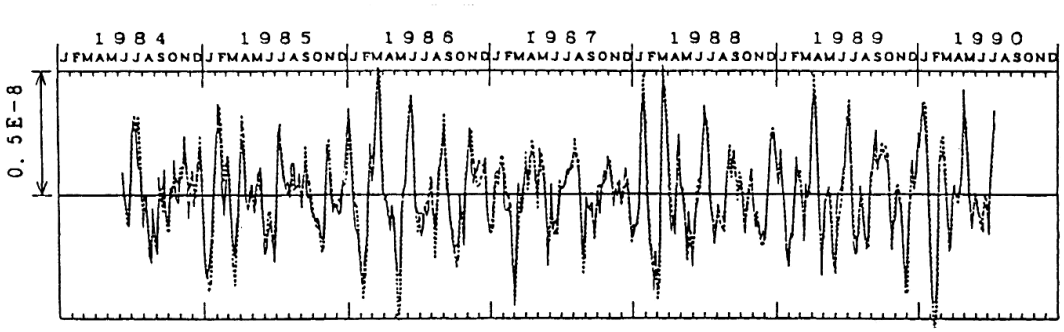


Figure 4.9.1: Seasonal variations in observed earth rotation (solid line) and calculated atmospheric angular momentum (broken line). Both sets of data are 150 days' high-pass filtered.

AAM functions calculated from JMA global analysis data at 00, 06, 12 and 18UTC have been provided operationally since early 1993. AAM functions calculated from JMA global 8-day forecast data at 12UTC are now also provided.

AAM functions calculated for a test period between 21 June and 30 September 1992 are shown in Figure 4.9.2, where days 1 - 102 correspond to this period. Each term of the AAM functions is multiplied by 10^7.

The broken lines show 6-hourly values of the functions (i.e., the difference from the period mean values), and the solid lines show band-pass filtered values for periods of 5-10 days. Oscillation with a 5-10 day period is notable in each term of each component, implying a corresponding oscillation with a similar period in the global-scale atmosphere.

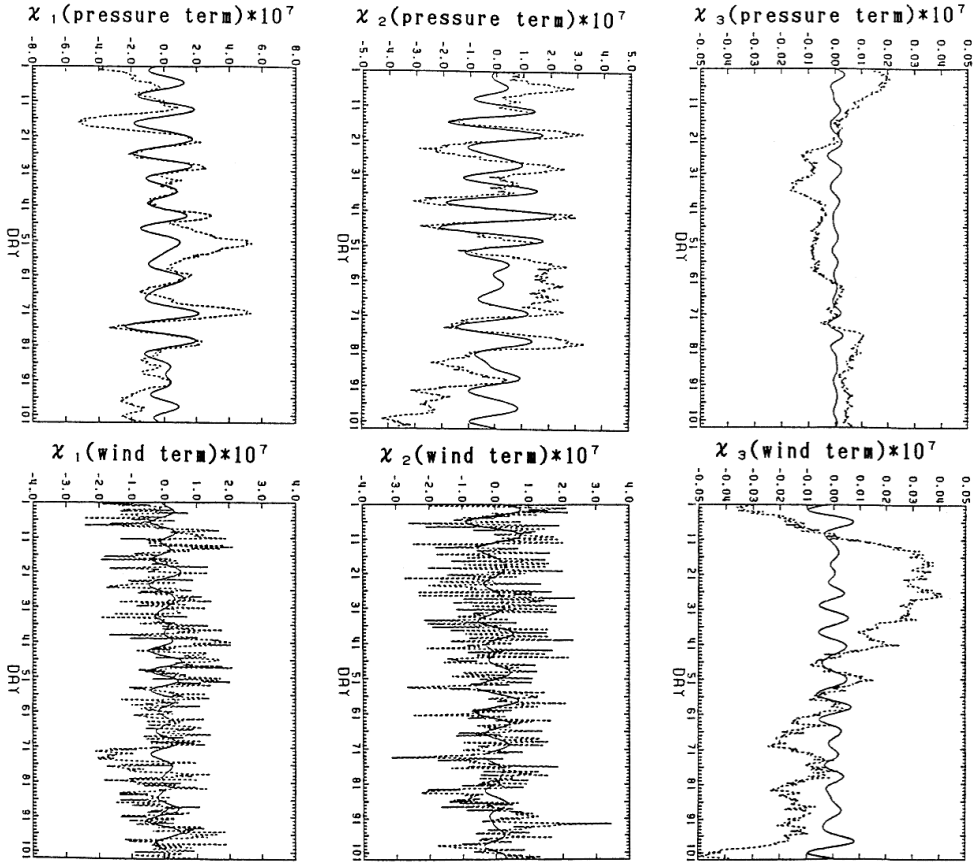


Figure 4.9.2: Pressure terms (top) and wind terms (bottom) of AAM functions. The left panels are the χ_1 component, the center ones are χ_2 and the right ones are χ_3. Days 1 - 102 correspond to 21 June - 30 September 1992. The broken lines show 6-hourly values of the functions, and the solid lines show band-pass filtered values for periods of 5-10 days. Each value is multiplied by 10^7.

Chapter 5

Ocean Models

5.1 Summary

(i) JMA operates two high-resolution MGDSST and HIMSST analyses based on satellite remote sensing data and *in-situ* observation data to provide real-time information on sea surface temperature. The former covers the global area with a horizontal resolution of 0.25°, and the latter covers the western North Pacific with a higher resolution of 0.1°. The Agency also conducts analysis based only on *in-situ* observation data (COBE-SST and COBE-SST2) for consistent monitoring of long-term ocean variations in phenomena such as El Niño events and global warming. COBE-SST2 has been operated as an update to COBE-SST since May 2021.

(ii) JMA has operated cascading ocean data assimilation systems since 2008 based on a unified ocean data assimilation framework known as the Multivariate Ocean Variational Estimation/Meteorological Research Institute Community Ocean Model (MOVE/MRI.COM) developed by the Agency's Meteorological Research Institute (MRI). Two such systems (MOVE-G3 and MOVE-JPN) are currently in operation. MOVE-G3 analyzes the global ocean to monitor El Niño and initializes the oceanic part of the coupled seasonal prediction system. MOVE-JPN is used for the analysis and prediction of oceanographic sub-mesoscale events and sea ice conditions over coastal areas around Japan.

(iii) JMA operates the Global Wave Model (GWM), the Coastal Wave Model (CWM) and the Wave Ensemble System (WENS) as ocean wave models. All three are based on the MRI-III, the third-generation wave model developed by MRI. The WENS was developed to provide probabilistic information on ocean waves in medium-range forecasts. The Shallow-water Wave Model (SWM) is also operated in trial mode to predict ocean waves in bays and nearshore areas around Japan with a higher horizontal resolution of 1 arc minute.

(iv) JMA operates the Japan Area storm surge model, the Japan area storm surge probabilistic forecast system and the Asia area storm surge ensemble prediction system. The Japan area storm surge model covers the Japan region and predicts storm surges generated by tropical and extra-tropical cyclones for the Japanese coast. JMA began operating the Japan area storm surge probabilistic forecast system based on the Japan area storm surge model in 2022. The Asia area storm surge ensemble prediction system is for the Asian region and is developed within the WMO Storm Surge Watch Scheme framework. The system was updated from the Asia area storm surge model using the GSM and all members of GEPS in 2022. Horizontal storm surge maps and time-series charts for selected points are issued to Typhoon Committee Members.

(v) JMA's oil spill prediction model was introduced in 1999. Operation is triggered when a large-scale oil spill occurs offshore. The effects of transport by sea surface winds, ocean waves and sea surface currents, turbulent diffusion, evaporation, and emulsification are considered.

5.2 Sea Surface Temperature Analysis

5.2.1 Merged Satellite and *In-situ* Data Global Daily Sea Surface Temperature (MGDSST)

Daily sea surface temperatures (SSTs) in the global ocean on a grid of $0.25 \times 0.25°$ are objectively analyzed to support ocean information services (Kurihara *et al.* 2006). The data are also used as boundary conditions for atmospheric short- and medium-range prediction models (see Section 3.2 and Section 3.3) and as observational data in the ocean data assimilation system (see Section 5.3). SST data obtained from satellite infrared sensors (Suomi-NPP/VIIRS, NOAA-20/VIIRS) and microwave sensors (GCOM-W/AMSR2) are used together with *in-situ* SST observations. Many *in-situ* data are obtained through the Global Telecommunication System, but domestic organizations also provide large amounts of data by e-mail.

Satellite-derived SST anomalies (SSTAs) from daily SST climatologies are decomposed into large/medium/small scales with cutoff wavelengths of 580, 143 and 55km, respectively, and into long- and short-time scales with cutoff periods of 53 and 10 days, respectively. The medium scale is intended to represent SST signals caused by mesoscale (eddy-scale) phenomena, and the small scale to represent front signals. Long-time scale signals represent intra-seasonal variations, and short-time scale signals represent variations influenced by atmospheric conditions such as tropical cyclones. Short-time signals are further processed by division into 27-53- and 10-27-day components, and only AMSR2 data are used for the latter. Signals varying with a period shorter than 10 days are cut off due to their significant data noise content.

The large scale and long-time scale components of SSTAs from satellites are calibrated with *in-situ* SSTAs using Poisson's equation (Reynolds 1987). Space-time optimum interpolation (OI) is applied to each component, and a zero value is adopted as the first guess. Space-time correlation coefficients and RMS values of the first guess error and satellite observation errors are statistically estimated a priori from satellite data using the method of Kuragano and Kamachi (2000). The daily SST is the sum of components of interpolated SSTAs and daily climatologies (Figure 5.2.1).

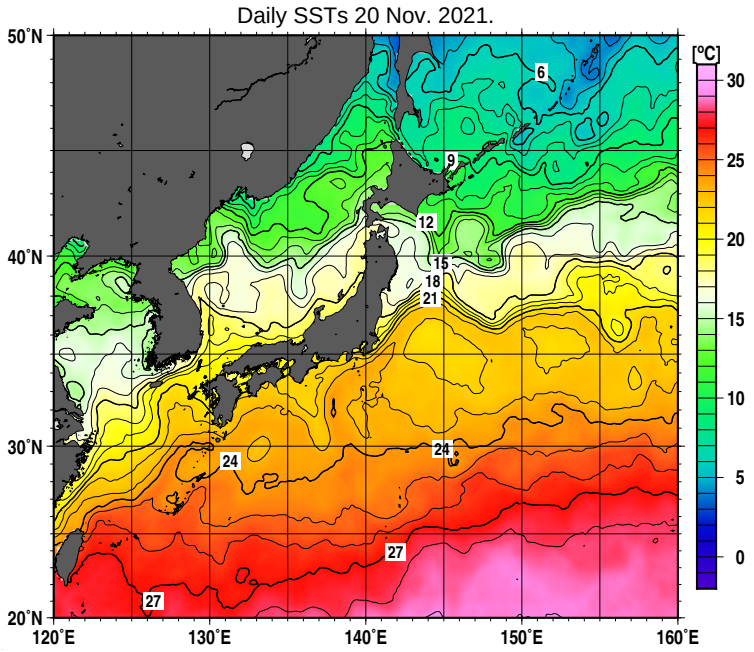


Figure 5.2.1: Sample SST analysis

5.2.2 HIgh-resolution Merged satellite and *In-situ* data Sea Surface Temperature (HIMSST)

SST analysis with high resolution is expected to provide optimal information variables such as boundary conditions for NWP models. Regional daily high-resolution ($0.1 \times 0.1^\circ$) SST analysis for the western North Pacific (HIMSST) began in November 2016. In this work, the HIMSST analysis framework is based on that of MGDSST, in which satellite-derived SSTs are decomposed into several spatio-temporal components and analyzed via optimum interpolation (OI).

In addition to the satellite data used in MGDSST, JMA Himawari-9 L3 SSTs produced with a $0.02 \times 0.02^\circ$ horizontal grid resolution and 10-minute intervals are incorporated into HIMSST analysis. JMA Himawari-9 L3 SSTs are calculated using a quasi-physical algorithm (Kurihara *et al.* 2016) and masked using JMA's cloud mask product (Imai and Yoshida 2016). More frequent geostationary meteorological satellite observation supports shorter time-scale components with periods from 10 days to 27 days in HIMSST analysis.

Suomi-NPP/VIIRS and NOAA-20/VIIRS SSTs and Himawari SSTs are converted into daily SST anomalies on a $0.1 \times 0.1^\circ$ grid and decomposed into small scale with cut-off wavelengths from 22 to 143 km, which are shorter than 55-143 km of MGDSST. Figure 5.2.2 shows daily HIMSSTs (left) and MGDSSTs (right) in the seas around Hokkaido, Japan. The HIMSSTs show sharper SST gradients due to the higher grid resolution and application of the shorter wavelength components of satellite-derived SSTs including Himawari-8/9.

HIMSSTs support JMA's ocean information services for the western North Pacific, and have also been used since March 2019 to provide boundary conditions for short-range NWP models such as the Meso-Scale Model (MSM) and the Local Forecast Model (LFM) (Section 3.5, Section 3.7). The analysis data are available from the NEAR-GOOS Regional Real Time Database (https://www.data.jma.go.jp/goos/data/database.html).

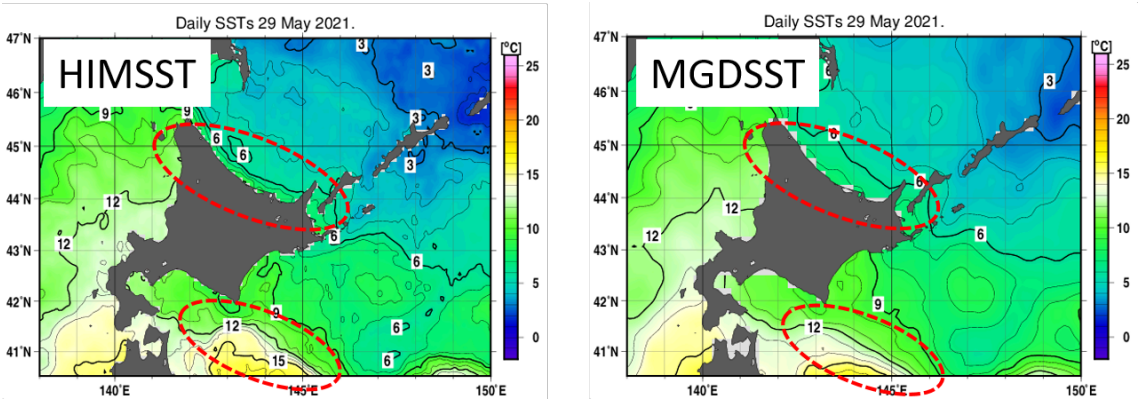


Figure 5.2.2: Comparison of HIMSST (left) and MGDSST (right) in the seas around Hokkaido.

5.2.3 Daily Sea Surface Analysis for Climate Monitoring

JMA produces historical Sea Surface Temperature (SST) datasets with analysis systems called Centennial In Situ Observation-Based Estimates of the Variability of SST and Marine Meteorological Variables (COBE; Ishii *et al.* 2005), and COBE2 (Hirahara *et al.* 2014). The COBE dataset will continue to support previous-generation monitoring for products that require it, but will eventually be superseded by COBE2.

The SST component of COBE (COBE-SST) has a resolution of 1° latitude and 1° longitude. The east-west grid points start at 0.5°E and end at 0.5°W, while the north-south points start at 89.5°S and end at 89.5°N. Analysis is based on optimum interpolation, and the deviation of the previous day's analysis from the 1950−2000 normal is multiplied by 0.95 for use as a first guess. Information on sea ice concentration is used in the estimation of SSTs in polar oceans, and analysis is performed on a daily basis with data for seven days centered on the day of interest. Observations are aggregated daily for each call sign and averaged over a $1.5^\circ \times 1.5^\circ$ grid box to form super-observations, which are later used for optimal interpolation. This procedure reduces correlations between observations and minimizes processing time.

Bias correction for past SST observation reports is based on Folland and Parker (1995). Observation data quality control is performed by checking ship tracks, dates and positions applicable to reports, and erroneous information is automatically corrected in the compilation of marine meteorological data at JMA. Based on observation data deviations from 1950−2000 normals for three months including the day of interest, the biases of data with identical ship call signs are estimated, and signs associated with large data biases are automatically blacklisted through daily analysis. Daily (final) analysis is performed with a delay of 31 days from real time to include delayed observations. Daily analysis for the latest 30-day period following the final analysis is also performed for real-time utilization. A detailed report on the characteristics of the data is available at this Japan Meteorological Agency (2006) link to the Tokyo Climate Center (TCC) website[1].

Since May 2021, COBE2 (updated from COBE) has been operated in parallel and its SST component (COBE-SST2) is utilized in century-scale monitoring of global warming and in historical oceanic analysis. The main improvements from COBE to COBE2 are as follows:

1. The historical database of in-situ observations has been updated from ICOADS 2.0 (Woodruff *et al.* 1998) to 2.5 (Woodruff *et al.* 2011).

2. A new bias correction method enables estimation for the biases of individual SST measurement types, such as insulated/uninsulated buckets and engine-room intake. The bias of observations with no measurement type information is defined as a mixture of those of specified measurement types.

3. An updated method of estimating SSTs in ice-covered regions enables clearer statistical relations between sea-ice concentration and SSTs for consideration of freezing points in sea water as a function of climatological sea surface salinity.

4. A new SST analysis method includes construction of a daily SST field as a sum of the trend, interannual variations, and daily changes.

Daily updated operational COBE-SST data are used for:

1. Monitoring of equatorial Pacific SSTs and El Niño/ La Niña evolution.

2. Input for the Japanese 55-year Reanalysis (JRA-55; Kobayashi *et al.* 2015; Section 2.11).

3. Input for the operational Ocean Data Assimilation System (MOVE/MRI.COM-G2; Toyoda *et al.* 2013).

Meanwhile, COBE-SST2 data are used for:

1. Input for the Japanese Reanalysis for Three Quarters of a Century (JRA-3Q; Kosaka *et al.* 2024; Section 2.11) covering the period from September 1947 to May 1985.

2. Input for the Japanese Reanalysis for Three Quarters of a Century with COBE-SST2 (JRA-3Q-COBE; Kosaka *et al.* 2024; Section 2.11) covering the period from June 1985 to December 1990.

3. Input for the operational Ocean Data Assimilation System (MOVE/MRI.COM-G3; Fujii *et al.* 2023; Section 5.3) and historical oceanic analysis (Section 5.3).

Monthly averaged COBE-SST data are available on the TCC website[2] and COBE-SST2 on the NEAR-GOOS Regional Real Time Database[3].

[1] https://www.data.jma.go.jp/tcc/tcc/products/elnino/cobesst_doc.html
[2] https://www.data.jma.go.jp/tcc/tcc/products/elnino/cobesst/cobe-sst.html
[3] https://www.data.jma.go.jp/goos/data/database.html

5.3 Ocean Data Assimilation and Prediction Systems

JMA has operated multiple ocean data assimilation systems based on the unified Multivariate Ocean Variational Estimation/Meteorological Research Institute Community Ocean Model (MOVE/MRI.COM) framework developed by its Meteorological Research Institute (MRI) since 2008. Currently, two such systems, MOVE-G3 and MOVE-JPN are in operation. MOVE-G3 is used to analyze the global ocean for El Niño monitoring and initialization of the oceanic part of the coupled seasonal prediction system (CPS3, Section 3.4). MOVE-JPN is used for analysis and prediction of oceanographic sub-mesoscale events and sea ice conditions over coastal areas around Japan. The MOVE/MRI.COM framework and the two operational systems are outlined below.

5.3.1 Ocean General Circulation Model and Objective Analysis Scheme: Common Framework

MOVE/MRI.COM consists of an ocean general circulation model (MRI.COM) and an objective analysis scheme (MOVE). Although MRI.COM and MOVE are continuously developed with a variety of features, this subsection outlines only features adopted in the operational systems. With the increasing number of options available in both components, operational configurations depend on resolution and system objectives (Table 5.3.1).

MRI.COM (Tsujino *et al.* 2010, 2017; Sakamoto *et al.* 2023) is a general-purpose ocean model with an Arakawa B-grid arrangement by which primitive equations are solved under hydrostatic and Boussinesq approximations. The z^* vertical coordinate system (Adcroft and Campin 2004) that allows free surface elevation and shallow (< 10 m) bottom topography is used. For nonlinear momentum advection, a generalized enstrophy-preserving scheme (Arakawa 1972) and a scheme involving the concept of diagonally upward/downward mass momentum fluxes along a sloping bottom are applied. A tracer advection scheme based on conservation of second-order moments (Prather 1986) is adopted in most systems while the more lightweight QUICK (Leonard 1979) is adopted in the analysis models. The bottom boundary layer (Nakano and Suginohara 2002) adopted in the coarse global models helps to reproduce the downslope advection of dense water. Vertical viscosity and diffusivity are determined using the Generic Length Scale scheme of Umlauf and Burchard (2003). A variety of options also exist in lateral mixing and viscosity parameterization. For coarse models that do not resolve eddies, isopycnal mixing (Redi 1982) and eddy-induced transport parameterized by Gent and McWilliams (1990) are used for tracers, and harmonic viscosity with the parameterization of Smagorinsky (1963) is used for momentum. For eddy-permitting and finer models, a biharmonic operator is used for horizontal turbulent mixing, and biharmonic friction with Smagorinsky-like viscosity (Griffies and Hallberg 2000) is used for momentum. A sea ice model with the thermodynamics of Mellor and Kantha (1989), thickness categories, and ridging/rheology following the Los Alamos sea ice model (CICE, Hunke and Lipscomb 2006) is enabled in all operational systems. Surface forcing to drive the model is based on the Japanese Reanalysis for Three Quarters of a Century (JRA-3Q, Kosaka *et al.* 2024), the Global Spectral Model (GSM, Yonehara *et al.* 2023, Sections 2.5 and 3.2), and the Global Ensemble Prediction System (GEPS, Ota *et al.* 2023, Section 3.3), depending on lead times and other considerations. River runoff climatology is commonly added to freshwater flux, and tidal forcing is enabled in coastal monitoring models.

Variational analysis with vertically coupled Temperature-Salinity (T-S) Empirical Orthogonal Function (EOF) modal decomposition (Fujii and Kamachi 2003b) is the basis for all assimilation systems. The model domain is divided into several subregions for which vertical T-S EOF modes are calculated in advance from observed T-S profiles. The amplitudes of T-S EOF modes, which are proportional to T-S increments down to 2000 m, serve as control variables. The vertical correlation of background error is naturally expressed in EOF modes, whereas horizontal correlations are expressed with different lengths prescribed for individual subregions and may be anisotropic. With the tangent linear and adjoint models of MRI.COM, MOVE now has a four-dimensional variational (4DVAR) analysis capability (Usui *et al.* 2015) in addition to the more traditional three-dimensional variational analysis with first guess at appropriate time (3DVAR-FGAT, Lorenc and Rawlins 2005). Control variables are optimized via a quasi-Newtonian approach (Fujii and Kamachi 2003a; Fujii 2005), and the analysis results are used for the model temperature and salinity fields via incremental analysis updates (IAU, Bloom *et al.* 1996). IAU is also used to adjust forecast-model T-S fields to those of the analysis

Table 5.3.1: MRI.COM and MOVE options adopted in operational ocean data assimilation systems.

Operational system	MOVE-G3		MOVE-JPN				
Subsystem/Model	G3A	G3F	Analysis		Prediction		
			GLB	NPR	GLB	NP	JPN
MRI.COM major version	4	4	4	4	5	5	5
Horizontal tracer advection							
QUICK	Y		Y	Y			
Second-order moment		Y			Y	Y	Y
Bottom boundary layer	Y		Y		Y		
Lateral diffusion/viscosity parameterization							
Isopycnal diffusion	Y		Y		Y		
Gent and McWilliams	Y		Y		Y		
Harmonic viscosity	Y		Y		Y		
Biharmonic diff./visc.		Y		Y		Y	Y
Forcing							
JRA-3Q	Y	Y	Y	Y			
Global Analysis (GA)	Y	Y					
GSM			Y	Y	Y	Y	Y
GEPS					Y	Y	
Tides					Y	Y	Y
Temperature and salinity analysis							
3DVAR-FGAT			Y				
4DVAR	Y			Y			
IAU initialization (to analysis)		Y			Y	Y	Y
Assimilated SST product							
MGDSST	Y		Y	Y			
Sea ice concentration analysis							
3DVAR	Y	Y					
Nudging				Y		Y	Y
Use of SSM/I obj. anl.	Y	Y		Y		Y	
Use of manual analysis							Y

model (which may differ in horizontal resolution or tidal forcing) in a process known as the replay method or dynamical downscaling.

MOVE/MRI.COM assimilates satellite altimeter data and in-situ observations of temperature and salinity reported from ships, profiling floats, and moored/drifting buoys via GTS and other communication systems. The satellite altimeter data are the level-3 product from the Copernicus Marine Environment Monitoring Service (CMEMS)[4]. Data from objective sea surface temperature analysis (MGDSST (Kurihara *et al.* 2006, Subsection 5.2.1), performed independently from MOVE/MRI.COM) are also assimilated. MOVE can additionally assimilate sea ice concentration data with 3DVAR (Toyoda *et al.* 2011, 2016) and nudging. Daily objective analysis of sea ice concentration from space-based SSM/I sensors (Matsumoto *et al.* 2006) and sea ice concentration subjectively estimated on the basis of data from satellites, aircraft, ships and coastal observations are also assimilated.

5.3.2 Ocean Data Assimilation System for Global Oceans (MOVE-G3)

Table 5.3.2: MOVE-G3 specifications

Horizontal grid spacing	G3A: 1° longitude × 0.3 – 0.5° latitude G3F: 0.25°
Vertical levels	60 (G3A with bottom boundary layer)
Number of grids	G3A: $364 \times 366 \times 61 = 8,126,664$ G3F: $1,444 \times 676 \times 60 = 58,568,640$
Sea surface forcing	Radiative heat flux, precipitation, surface wind, pressure, temperature, humidity Delayed analysis: from JRA-3Q Early analysis: from Global Analysis
River runoff	Climatology from JRA55-do (Tsujino *et al.* 2018)
Assimilation	G3A: 4DVAR with 10-day windows (observations for the last 5 days assimilated by giving T-S increments for the first 5 days)
Initialization	G3F: IAU for 5 days using G3A results
Sea ice assimilation	3DVAR
Observational data	In-situ temperature and salinity above -2,000 m Sea level anomaly (Jason-3, Cryosat-2, Saral, Sentinel-3A, Sentinel-3B) Sea surface temperature (MGDSST) and sea ice concentration (SSM/I, Matsumoto *et al.* 2006)

Operation of the MOVE-G3 global ocean data assimilation system began in February 2022 (Table 5.3.2, Fujii *et al.* 2023) with incorporation of lower-resolution 4DVAR analysis (G3A) and initialization of oceanic parts for coupled forecasts (G3F, Figure 5.3.1). The G3A model adopts a horizontal resolution of 0.5° latitude and 1.0° longitude except for the 15°S – 15°N band, where latitudinal grid spacing is at a minimum of 0.3° between 6°S and 6°N, has 60 vertical levels and a bottom boundary layer (Nakano and Suginohara 2002), with 23 levels above 200 m, along with realistic bottom topography and a maximum depth of 6,300 m. The computational domain is the global ocean, including the Arctic Ocean with use of tri-polar horizontal coordinates (Figure 5.3.2). A bias correction scheme (Fujii *et al.* 2012) and a global water mass correction scheme (Kuragano *et al.* 2014) are incorporated into MOVE-G3. The G3F forecast model covers global oceans using a higher resolution of 0.25 degrees, initialized by downscaling G3A temperature and salinity fields with IAU.

Sea ice concentration is also assimilated in MOVE-G3 (Toyoda *et al.* 2011, 2016). The 3DVAR scheme first combines observation data from daily objective analysis of sea ice concentration (Matsumoto *et al.* 2006) with information on background sea ice concentration from the model forecast. The analysis concentration increment is then applied to gradually adjust model sea ice with IAU. No information on sea ice is passed from G3A to G3F, and concentration is assimilated independently in both.

Data production is enhanced to provide prompt and frequent initial conditions for the coupled model. Although MOVE-G3 has a five-day data window, five staggered analysis streams are applied so that initial conditions can be determined every day (valid at 00 UTC). The operational system features early analysis on the same day for initial conditions and delayed cycle analysis with a four-day delay for more uniform

[4]Product identifier: SEALEVEL_GLO_PHY_L3_NRT_OBSERVATIONS_008_044 (https://doi.org/10.48670/moi-00147)

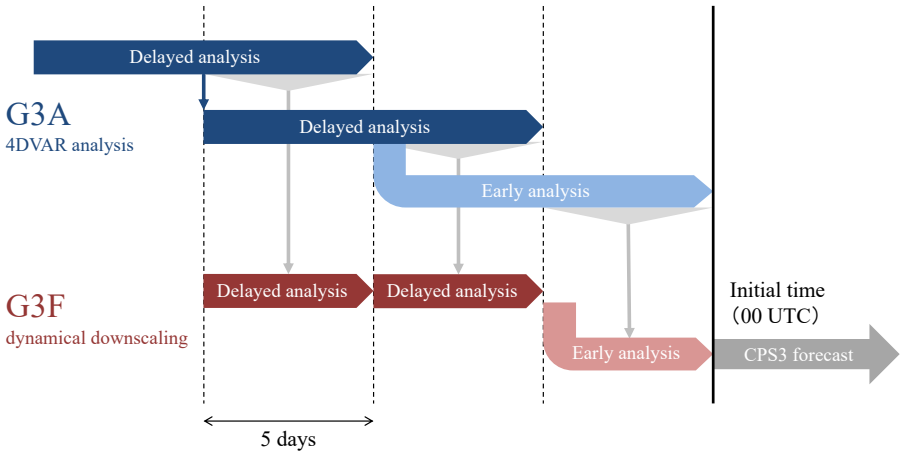


Figure 5.3.1: The MOVE-G3 system, with the abscissa representing the time of validity. Within a ten-day analysis window of the G3A subsystem, observations for the last five days are assimilated by giving T-S increments for the first five days, and restart files are created at the end of the fifth day. The G3F subsystem adjusts its five-day mean temperature and salinity fields to those of G3A using IAU (represented by downward grey triangles and arrows). Operationally, five mutually independent temporally staggered analysis streams are employed for daily initialization of the CPS3 coupled model (not shown).

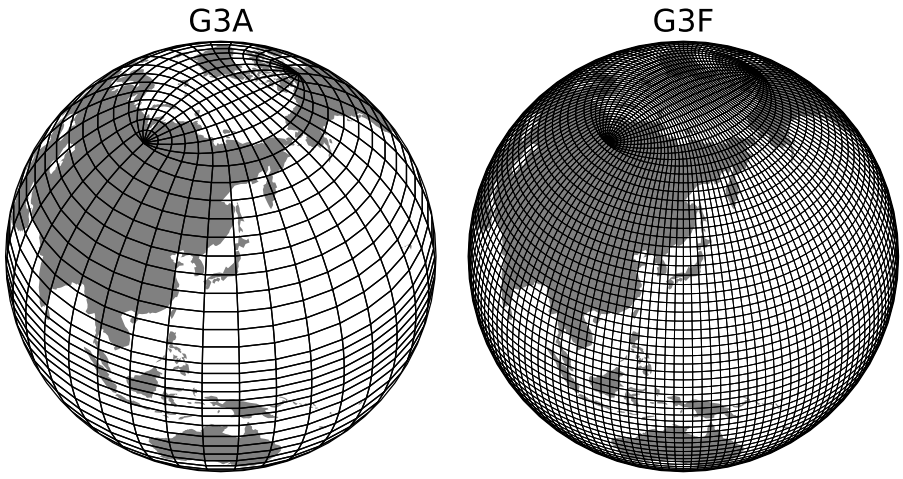


Figure 5.3.2: Tripolar coordinate adopted for the two MOVE-G3 models. Boundaries are shown for every 10 grids.

analysis. The early part involves the use of atmospheric Global Early Analysis with the GSM for surface forcing, whereas the delayed part involves the more uniform JRA-3Q. The 4DVAR system also generates ensemble perturbations approximating analysis error covariances using minimization histories (Niwa and Fujii 2020). With this capability, early analysis provides daily five-member initial ensembles for the coupled model.

MOVE-G3 output is used in various forms for JMA El Niño monitoring, and products for the equatorial Pacific region are published in Monthly Highlights on the Climate System and the El Niño Outlook. Report charts are also provided on the Tokyo Climate Center Web page[5].

5.3.3 Ocean Data Assimilation and Prediction System for the Seas Around Japan (MOVE-JPN)

MOVE-JPN consists of an analysis system that assimilates ocean observation data and simulates the latest ocean state, and a prediction system that initializes higher-resolution models based on analysis and forecasts ocean conditions.

The analysis system (Hirose *et al.* 2019, Table 5.3.3) consists of a global model (GLB) and a North Pacific model (NPR[6]), both with 60 levels, 31 of which are above 400 m. The GLB has a bottom boundary layer (Nakano and Suginohara 2002) and horizontal coordinates identical to those of G3A. The NPR model domain spans the area from 99°E to 75°W zonally and 15°S to 63°N meridionally. NPR horizontal resolution is variable, with values of 1/11° from 114°E to 163°E and 1/10° from 17°N to 56°N (≈ 10 km around Japan) and coarser elsewhere. The NPR model (one-way nested within the GLB assimilating global data based on 3DVAR) assimilates observation data based on 4DVAR (Usui *et al.* 2015).

The MOVE-JPN prediction system consists of the GLB, another North Pacific model (NP) and the JPN model (Sakamoto *et al.* 2019, Table 5.3.4). The NP model domain is identical to that of NPR, but its horizontal resolution is zonally 1/11° and meridionally 1/10° over the entire domain (Figure 5.3.3). The JPN model domain spans the area 117−160°E and 20−52°N with grid spacing of around 2 km (zonally 1/33°, meridionally 1/50°). The JPN model is one-way nested to the two-way nested set of the GLB and NP models, all of which explicitly represent tidal motion. NP and JPN are initialized using the IAU scheme with temporal and spatial filtering on temperature and salinity fields from NPR 4DVAR analysis. Using surface boundary conditions output by the GSM, the JPN model forecasts ocean states around Japan with a lead time of 11 days. The two-way nested GLB and NP models continue to forecast ocean states for a month ahead based on GEPS and JRA-55 climatology forcing.

[5] https://www.data.jma.go.jp/tcc/tcc/index.html

[6] The NPR model used for 4DVAR analysis is a reduced-resolution version of the NP model detailed later.

Table 5.3.3: MOVE-JPN analysis system specifications

Model domain	GLB: global NPR: 15°S – 63°N between 99°E and 75°W
Nesting	GLB/NPR 1-way off line
Horizontal grid spacing	GLB: 1° longitude × 0.3 – 0.5° latitude NPR: 1/11° longitude × 1/10° latitude for seas around Japan, coarser elsewhere
Vertical levels	60 (GLB with bottom boundary layer)
Number of grids	GLB: $364 \times 366 \times 61 = 8,126,664$ NPR: $1,334 \times 578 \times 60 = 46,263,120$
Sea surface forcing	Radiative heat flux, precipitation, surface wind, pressure, temperature, humidity Delayed analysis: JRA-3Q Early analysis: GSM (Section 3.2)
River runoff	Climatology from JRA55-do (Tsujino *et al.* 2018)
Assimilation	GLB: 3DVAR with 5-day windows NPR: 4DVAR with 10-day windows (observations for the last 7 days assimilated by giving T-S increments for the first 3 days)
Sea ice assimilation	NPR: nudging
Observational data	In-situ temperature and salinity above -2,000 m Sea level anomaly (Jason-3, Cryosat-2, Saral, Sentinel-3a,b, HaiYang-2b) Sea surface temperature (MGDSST) and sea ice concentration (SSM/I, Matsumoto *et al.* 2006)

Table 5.3.4: MOVE-JPN prediction system specifications

Model domain	GLB: global NP: 15°S – 63°N, 99°E – 75°W JPN: 20° – 52°N, 117° – 160°E
Nesting	GLB and NP: 2-way on line JPN: 1-way off line to NP
Horizontal grid spacing	GLB: 1° longitude × 0.3 – 0.5° latitude NP: 1/11° longitude × 1/10° latitude JPN: 1/33° longitude × 1/50° latitude
Vertical levels	60 (GLB with bottom boundary layer)
Number of grids	GLB: $364 \times 366 \times 61 = 8,126,664$ NP: $2,049 \times 784 \times 60 = 96,384,960$ JPN: $1,423 \times 1,604 \times 60 = 136,949,520$
Tidal forcing	8 main tidal constituents
Sea surface forcing	Radiative heat flux, precipitation, surface wind, pressure, temperature, humidity Up to 11th day: GSM From 12th day: GEPS and JRA-55 climatology
River runoff	Climatology from JRA55-do (Tsujino *et al.* 2018)
Initialization	IAU for 3 days using GLB-3DVAR and NPR-4DVAR results
Sea ice initialization	NP: nudging to SSM/I JPN: nudging to JMA manual sea ice analysis
Prediction period	GLB and NP: 31 days JPN: 11 days

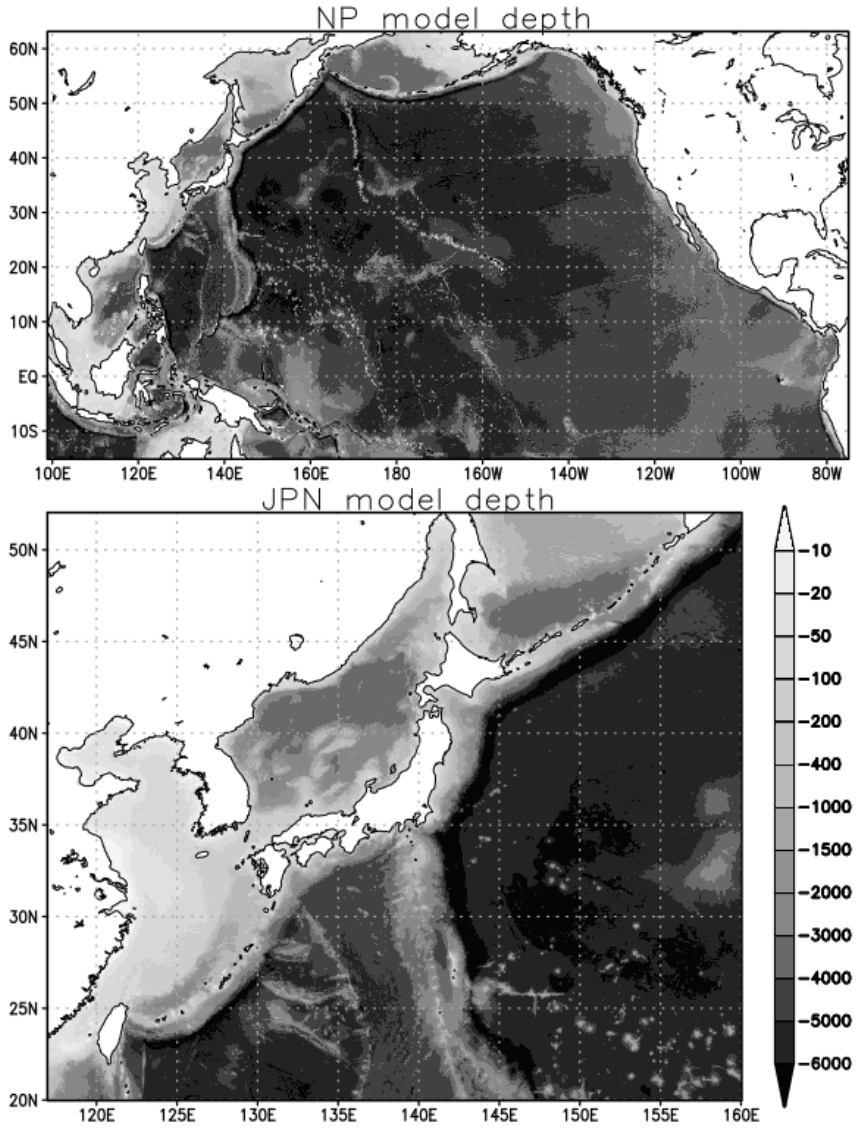


Figure 5.3.3: Bottom topography of the NP model (top) and the JPN model (bottom) in m. The figures show the domain of each model.

Table 5.4.1: Specifications of the Global Wave Model, the Coastal Wave Model and the Wave Ensemble System.

Model name	Global Wave Model	Coastal Wave Model	Wave Ensemble System
Type of wave model	spectral model (third-generation wave model, MRI-III)		
Area	global 75°N – 75°S 180°W – 0° – 180°E	coastal sea of Japan 50°N – 20°N 120°E – 150°E	global 75°N – 75°S 180°W – 0° – 180°E
Grid size	0.25° × 0.25° (1440 × 601)	0.05° × 0.05° (601 × 601)	0.5° × 0.5° (720 × 301)
Time step (advection term) (source term)	 5 minutes 15 minutes	 1 minute 3 minutes	 10 minutes 30 minutes
Forecast range (initial time)	132 hours (06 and 18 UTC) 264 hours (00 and 12 UTC)	132 hours (00, 06, 12 and 18 UTC)	264 hours (00 and 12 UTC)
Spectral component	900 components 25 frequencies from 0.0375 to 0.3 Hz (logarithmically partitioned) 36 directions		
Initial condition	Analysis by Optimal Interpolation		
Boundary condition		Global Wave Model	
Wind field	Global Spectral Model (GSM)		Global Ensemble Prediction System (GEPS)
	Fujita's empirical formula and a corresponding gradient wind for a typhoon		
Shallow-water effects	Refraction and bottom friction		

5.4 Ocean Wave Models

5.4.1 Introduction

The Japan Meteorological Agency (JMA) operates the Global Wave Model (GWM), the Coastal Wave Model (CWM), and the Wave Ensemble System (WENS). The Shallow-water Wave Model (SWM) is also operated in selected regions.

The GWM, CWM and WENS are based on MRI-III developed by JMA's Meteorological Research Institute (Ueno and Kohno 2004). The current versions of the GWM and the CWM, which incorporate shallow water effects, have been used for short-range forecasting since May 2017. The WENS, which is a prediction system incorporating probability information for mid-range forecasts, has been in operation since June 2016. The specifications of the three models are given in Table 5.4.1, and their domains are shown in Figure 5.4.1.

The SWM is based on the WAM (The WAMDI Group 1988) with modification by the National Institute for Land and Infrastructure Management of the Ministry of Land, Infrastructure, Transport, and Tourism (MLIT), with quasi-operation under a cooperative framework with MLIT's Water and Disaster Management Bureau. It has a high resolution of 1 minute (see Table 5.4.2), and is operated over 22 areas. SWM products are used exclusively within JMA and MLIT's Regional Development Bureaus.

5.4.2 Ocean Wave Model Structure

The ocean wave models forecast the wave energy density (spectrum) of each frequency and direction (i.e., the two-dimensional (directional) wave spectrum). The basic equation is the energy balance expression:

$$\frac{\partial F}{\partial t} + \nabla \cdot (\boldsymbol{C}_g F) + \frac{\partial}{\partial \theta}(\Omega F) = S_{net} = S_{in} + S_{nl} + S_{ds} + S_{btm} \tag{5.4.1}$$

where

$$\Omega = \frac{C_g}{C_p}\left(-\frac{\partial C_p}{\partial x}\cos\theta + \frac{\partial C_p}{\partial y}\sin\theta\right)$$

which represents refraction in shallow water. $F(f, \theta, \boldsymbol{x}, t)$ is a two-dimensional spectrum, where f is the frequency and θ is the wave direction, $\boldsymbol{C}_g(f, \theta, \boldsymbol{x})$ is the group velocity, and $C_p(f, \boldsymbol{x})$ is the phase speed. The

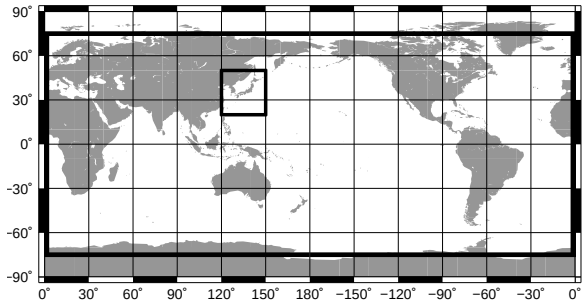


Figure 5.4.1: Calculation areas for the Global Wave Model and the Wave Ensemble System (outer thick lines), and the Coastal Wave Model (inner thick lines).

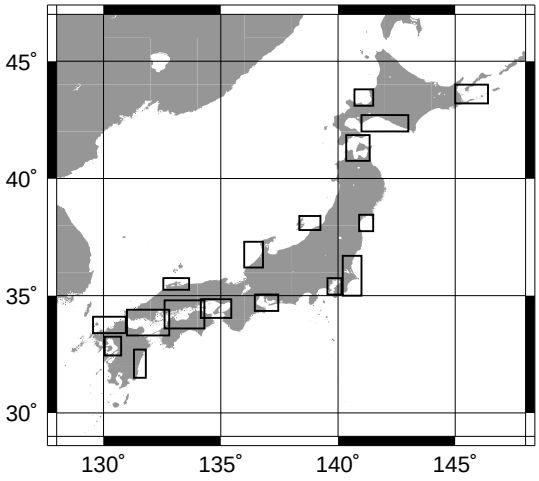


Figure 5.4.2: Calculation areas for the Shallow-water Wave Model.

Table 5.4.2: Specifications of the Shallow-water Wave Model.

Type of wave model	spectral model (third-generation wave model, WAM)					
Areas	Domain name	Grid size	Integration domain	Domain name	Grid size	Integration domain
	Tokyo Bay	37×43	35.75°N – 35.05°N 139.55°E – 140.15°E	Off Shimane	67×31	35.75°N – 35.25°N 132.55°E – 133.65°E
	Ise Bay	61×43	35.05°N – 34.35°N 136.45°E – 137.45°E	Ishikari Bay	49×43	43.80°N – 43.10°N 140.70°E – 141.50°E
	Harima-Nada Osaka Bay	79×49	34.85°N – 34.05°N 134.15°E – 135.45°E	Off Ishikawa	49×67	37.30°N – 36.20°N 136.00°E – 136.80°E
	Ariake Sea Shiranui Sea	43×49	33.25°N – 32.45°N 130.05°E – 130.75°E	Off Nemuro	85×49	44.00°N – 43.20°N 145.00°E – 146.40°E
	Off Niigata	55×37	38.40°N – 37.80°N 138.35°E – 139.25°E	Off Miyazaki	31×73	32.70°N – 31.50°N 131.30°E – 131.80°E
	Sendai Bay	37×43	38.45°N – 37.75°N 140.90°E – 141.50°E	Tsugaru Strait	61×67	41.85°N – 40.75°N 140.35°E – 141.35°E
	Off Tomakomai	121×43	42.70°N – 42.00°N 141.00°E – 143.00°E	Off Ibaraki Off Boso	49×103	36.70°N – 35.00°N 140.20°E – 141.00°E
	Suo-Nada Iyo-Nada Aki-Nada	109×67	34.40°N – 33.30°N 131.00°E – 132.80°E	Genkai-Nada	83×43	34.10°N – 33.40°N 129.55°E – 130.95°E
	Hiuchi-Nada	103×73	34.80°N – 33.60°N 132.60°E – 134.30°E			
Grid resolution	$1' \times 1'$					
Time step (advection term) (source term)	 1 minute 1 minute					
Forecast range	39 hours					
Spectral component	1260 components 35 frequencies from 0.0418 to 1.1 Hz (logarithmically partitioned) 36 directions					
Initial condition	Coastal Wave Model					
Boundary condition	Coastal Wave Model					
Wind field	Meso-Scale Model (MSM)					
	Fujita's empirical formula and a corresponding gradient wind for a typhoon					

magnitude of group velocity is simply $C_g(f)$ for deep-water waves, and depends on water depth for shallow-water waves. S_{net} is a net source function consisting of S_{in}, S_{nl}, S_{ds}, and S_{btm}, which are briefly outlined below. Only the model numerics of the MRI-III are described here, as those of the WAM are already extensively referenced elsewhere (e.g. Janssen 2004).

1. S_{in}: energy input from wind. This value generally takes the form $S_{in} = A + BF$, where A is linear wave growth and BF is exponential growth. In the MRI-III, the formula of Cavaleri and Rizzoli (1981) is used for linear growth

$$A = 1.5 \times 10^{-3} \left(\frac{u_*^4}{2\pi g^2}\right) \exp\left[-\left(\frac{f_{PM}}{f}\right)^4\right] (\max\{0, \cos(\theta - \theta_W)\})^4 \tag{5.4.2}$$

where u_* is the friction velocity of wind, θ_W is the wind direction, and g is the gravitational acceleration. In general, the linear term has little influence on wave growth except in the very early stages. Meanwhile, the exponential term BF has a key role in wave growth. In the MRI-III, The B is expressed as

$$B(f, u_*, \theta_W - \theta) = c_{in} \left(\frac{u_*}{C_{pd}}\right)^2 f \frac{\cos^3(\theta_W - \theta)}{|\cos(\theta_W - \theta)|} \tag{5.4.3}$$

where C_{pd} is the phase speed of deep water waves, i.e., $C_{pd} = \dfrac{g}{\omega} = \dfrac{g}{2\pi f}$.

This expression is based on Mitsuyasu and Honda (1982) and Plant (1982).

2. S_{nl}: nonlinear energy transfer associated with resonant interaction. Since rigorous calculation is highly time-consuming, a practical scheme known as discrete interaction approximation (DIA) (Hasselmann *et al.* 1985) is commonly used in operational wave models. This approach involves the use of only one parameter for the set of four resonant waves:

$$\left.\begin{aligned} &f_1 = f_2 = f, \\ &f_3 = f(1+\lambda) = f_+, \\ &f_4 = f(1-\lambda) = f_-, \\ &\theta_1 = \theta_2 = \theta, \\ &\theta_3 - \theta = \pm \cos^{-1}\left\{\frac{1 + 2\lambda + 2\lambda^3}{(1+\lambda)^2}\right\}, \\ &\theta_4 - \theta = \mp \cos^{-1}\left\{\frac{1 - 2\lambda - 2\lambda^3}{(1-\lambda)^2}\right\} \end{aligned}\right\} \tag{5.4.4}$$

$$\begin{Bmatrix} \delta S_{nl} \\ \delta S_{nl}^+ \\ \delta S_{nl}^- \end{Bmatrix} = \begin{Bmatrix} -2\dfrac{\Delta f \Delta\theta}{\Delta f \Delta\theta} \\ \dfrac{(1+\lambda)(\Delta f \Delta\theta)}{\Delta f_+ \Delta\theta} \\ \dfrac{(1-\lambda)(\Delta f \Delta\theta)}{\Delta f_- \Delta\theta} \end{Bmatrix}$$

$$\times C f^{11} g^{-4} \left[F^2 \left\{\frac{F_+}{(1+\lambda)^4} + \frac{F_-}{(1-\lambda)^4}\right\} - 2\frac{F F_+ F_-}{(1-\lambda^2)^4}\right] \tag{5.4.5}$$

$$S_{nl} = \delta S_{nl} + \delta S_{nl}^+ + \delta S_{nl}^- \tag{5.4.6}$$

where $F \equiv F(f, \theta)$, $F_+ \equiv F(f_+, \theta_3)$, $F_- \equiv F(f_-, \theta_4)$. The coefficient C is determined to fit exact calculation for the JONSWAP spectrum. Hasselmann *et al.* (1985) defined the related parameters as $\lambda = 0.25$, corresponding to $\theta_3 - \theta = \pm 11.5^\circ$, $\theta_4 - \theta = \mp 33.6^\circ$ and $C = 3 \times 10^7$. DIA calculation is found to support highly accurate estimation if multiple parameters are combined. δS_{nl} is the rate of change in nonlinear energy transfer for the parameter λ. δS_{nl}, δS_{nl}^+ and δS_{nl}^- are calculated with three configurations in the MRI-III. The parameters used are $\lambda_1 = 0.19\,(C_1 = 1.191 \times 10^7)$, $\lambda_2 = 0.23\,(C_2 = 6.835 \times 10^6)$, and $\lambda_3 = 0.33\,(C_3 = 1.632 \times 10^6)$.

3. S_{ds}: energy dissipation associated with wave breaking and other effects. In the MRI-III, dissipation terms are expressed as local energy dissipation as proposed by Ueno (1998).

$$S_{ds} = -c_b \frac{u_*}{g^3} f^7 \left(F(f,\theta)\right)^2 \tag{5.4.7}$$

where c_b is a coefficient determined to fit wave generation. In the MRI-III, a slightly artificial swell decay process is included.

$$S_{sds} = -2.96 \times 10^{-6} \tanh\left[4\frac{f_s - f}{f_p}\right] F(f,\theta) \tag{5.4.8}$$

where $f_p = 0.156g/U_{10N}$ represents the peak frequency of the Pierson-Moskowitz (PM) spectrum from the 10m height wind speed U_{10N}. This decay function is applied to the spectrum of frequencies lower than $f_s = 1.8f_p$ when the significant wave height exceeds 1.5m.

4. S_{btm}: the energy loss associated with bottom friction. This effect needs to be considered for shallow water. In the MRI-III, this term is taken from Hasselmann *et al.* (1973).

$$S_{btm} = -\frac{0.038}{g^2} \cdot \frac{(2\pi f)^2}{\sinh^2(kd)} \cdot F(f,\theta) \tag{5.4.9}$$

where k is the wave number and d is depth. Energy loss increases for shallow water and long-period waves.

5.4.3 Wind Field

Wind fields for the GWM and the CWM are given by the Global Spectral Model (GSM), while the SWM uses Meso-Scale Model (MSM) winds. For the WENS, 51-member wind fields of the Global Ensemble Prediction System (GEPS) are employed.

Wind fields around typhoons are empirically modified in the GWM, CWM and SWM. As typhoons generate extremely high waves in the western North Pacific, accurate wave forecasting is crucial to maritime and coastal safety. As NWP models may not adequately predict typhoon intensity or tracks, wind fields based on operational typhoon analysis and forecasting are imposed onto NWP winds (known as bogus winds) when a typhoon is present over the western North Pacific. Wave conditions may differ significantly with changes in typhoon track, especially in the limited region covered by the SWM. Accordingly, wave fields are also predicted with the assumption of five different paths for each approaching typhoon (center, faster, slower, and right/left of forecast circle) within forecast error circles. See Subsection 5.5.1.3 for a description of forecast circles.

To create bogus wind data, sea level pressure distribution near a typhoon is assumed to have a profile expressed by Fujita's empirical formula (Fujita 1952)

$$P(r) = P_\infty - \frac{P_\infty - P_0}{\sqrt{1 + (r/r_0)^2}} \tag{5.4.10}$$

where P_∞, P_0 and r_0 denote the ambient pressure, the central pressure of the typhoon, and the scaling factor of the radial distribution of the pressure, respectively. Surface winds near the typhoon are estimated from the pressure field by assuming the gradient wind balance with modifications based on the typhoon movement and surface friction effects.

5.4.4 Wave Analysis

An assimilation scheme (Kohno *et al.* 2012) for the GWM and the CWM was introduced in October 2012. In these models, initial conditions (wave spectra) are modified based on significant wave heights under the Objective Wave Analysis System (OWAS) (Kohno *et al.* 2009), which is used for objective analysis of wave heights using optimal interpolation (OI) with observations from radar altimeters on board satellites, buoys and coastal wave recorders, and ship reports (Table 5.4.3). Data assimilation improves the prediction of ocean wave fields, especially in terms of shorter forecast times and swell propagation.

Table 5.4.3: JMA Objective Wave Analyses System specifications.

Analysis scheme	Optimal interpolation
Data cut-off time	6 hours and 25 minutes for early run analysis 12hours for delayed analysis
First guess	6-hour forecast by the GWM and the CWM
Analysis variables	Significant wave height
Grid size	0.25°×0.25° grid for the GWM, 0.05°×0.05° grid for the CWM
Domain	Global oceans for the GWM, Coastal sea of Japan for the CWM
Observational data	0.25°×0.25° grid super-observational data from BUOY, SHIP, Nowphas, GPS wave meter, Jason-3, SARAL, and Sentinel-3A/B
Assimilation window	6 hours

5.4.5 Products

JMA provides various ocean wave products based on model data via its website[7].

Such output based on GWM information comes from RSMC activities for numerical ocean wave prediction.

Charts showing analysis and 24-hour ocean wave forecasts are provided twice a day via the JMH radio facsimile broadcast service and the JMA website for the Western North Pacific and seas around Japan (Figure 5.4.3). The charts indicate significant wave heights, peak wave directions and peak wave periods. In addition, information on rough sea areas that may hinder maritime navigation was incorporated for forecast charts in 2017. The horizontal hatching in Figure 5.4.3 (a) indicates areas of crossing waves that may give rise to unexpectedly high-sea conditions. The vertical hatching in Figure 5.4.3 (b) indicates areas of increased wave height and steepness due to the effects of opposing ocean currents.

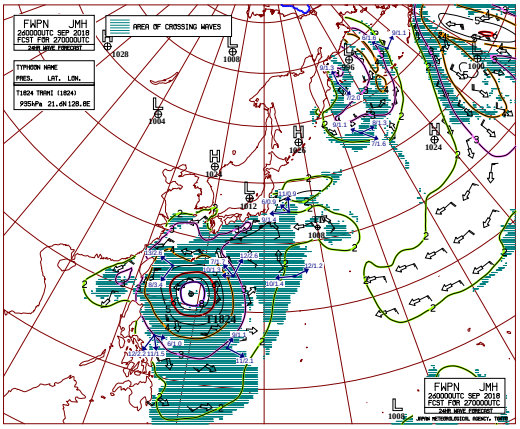
The following probabilistic medium-range forecast products highlighting significant wave heights and periods are statistically produced from WENS predictions:

1. Distribution maps for ensemble mean, maximum, 3rd quartile, probability of threshold breaches and ensemble spread of elements for ocean wave forecasts

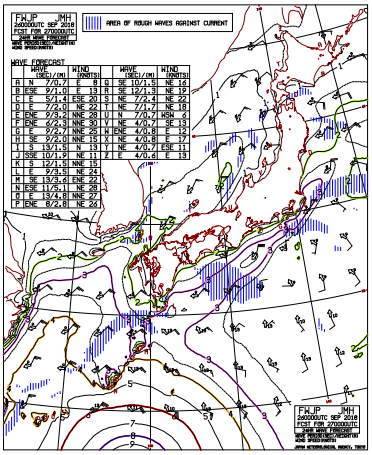
2. Time-series representations with box plots and probability of threshold breaches

These products are provided on JMA's website for the WMO Severe Weather Forecasting Programme (SWFP[8]) and on the Numerical Typhoon Prediction (NTP) website for Typhoon Committee members. Figure 5.4.4 and Figure 5.4.5 show examples for TY Surigae (2102).

[7] https://www.jma.go.jp/jma/en/NMHS/JMA_RSMC.html
[8] https://www.wis-jma.go.jp/swfdp/

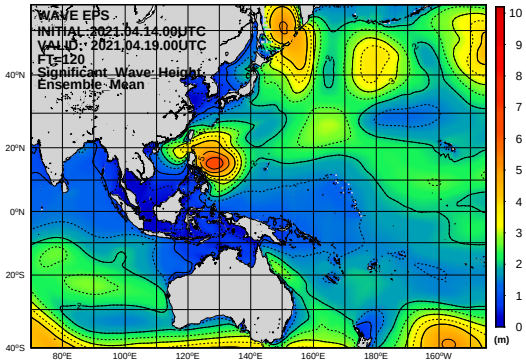


(a) Wave forecast chart for the Western North Pacific (FWPN) created from the GWM.

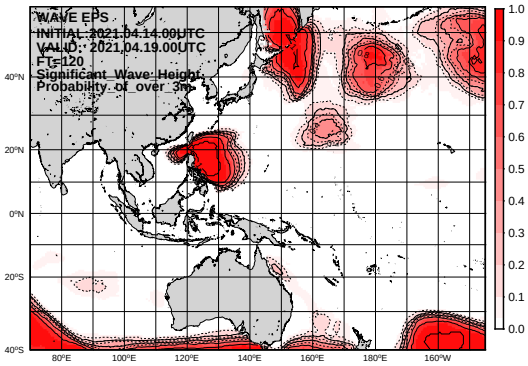


(b) Wave forecast chart for coastal regions of Japan (FWJP) created from the CWM.

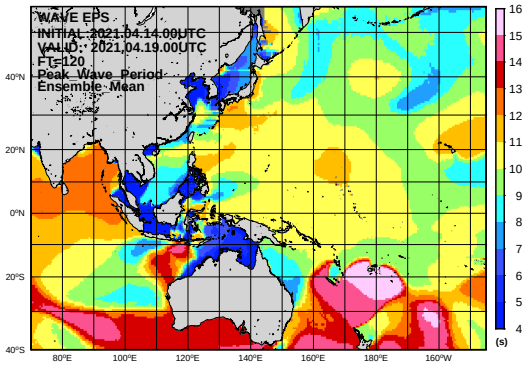
Figure 5.4.3: Wave forecast charts based on 24-hour model predictions with an initial time of 00 UTC on 26 September 2018, including significant wave height(contours), wave direction(arrows), sea surface wind speed and direction(barbs).



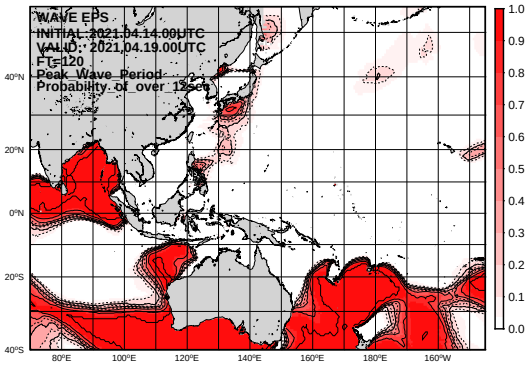
(a) Ensemble means of significant wave heights.



(b) Probability of wave heights over 3 m.

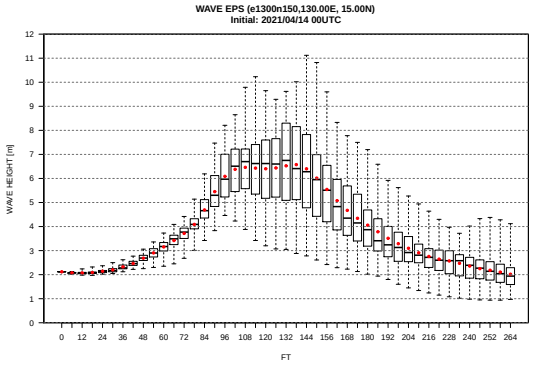


(c) Ensemble means of wave periods.

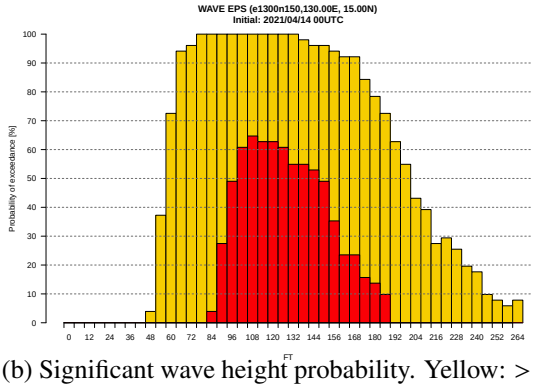


(d) Probability of wave periods over 12 sec.

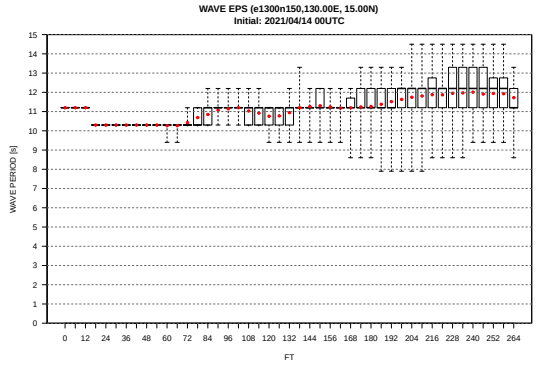
Figure 5.4.4: Distribution maps from 120-hour WENS predictions with an initial time of 00 UTC 14 on April 2021.



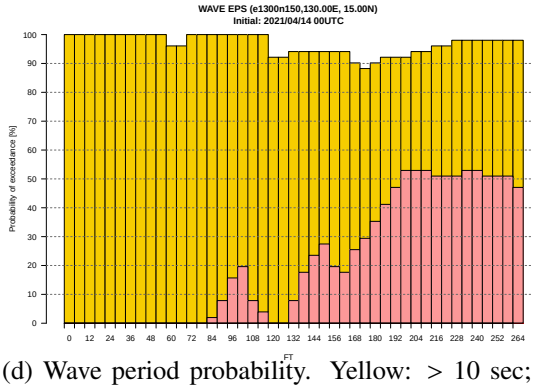
(a) Box plots for significant wave heights.



(b) Significant wave height probability. Yellow: > 3 m; red: > 6 m.



(c) Box plots for wave periods.



(d) Wave period probability. Yellow: > 10 sec; pink: > 13 sec; red: > 15 sec.

Figure 5.4.5: Time-series representations for areas off the east coast of the Philippines from WENS prediction with an initial time of 00 UTC 14 on April 2021.

5.4.6 Improvement and Development

The main improvements made to JMA ocean wave models since 2017 are as follows:

1. The shallow water effect was introduced into the GWM and the CWM in 2017, and into the WENS in 2020.
2. WENS horizontal resolution was upgraded from 1.25° to 0.5° in 2020.
3. The number of WENS members was increased from 27 to 51 in 2021.
4. GWM horizontal resolution was enhanced from 0.5° to 0.25° in January 2023.

Figure 5.4.6 shows monthly root mean square errors (RMSEs) for the significant wave height of the GWM against all kinds of observations measured by buoys, ships and satellites from 2012 to 2022. The remarkable improvement of T+0 RMSE in late 2012 corresponds to the implementation of data assimilation. The gradual decrease in RMSEs for the all lead times suggests the improvement of the performance of the JMA's weather and wave prediction systems.

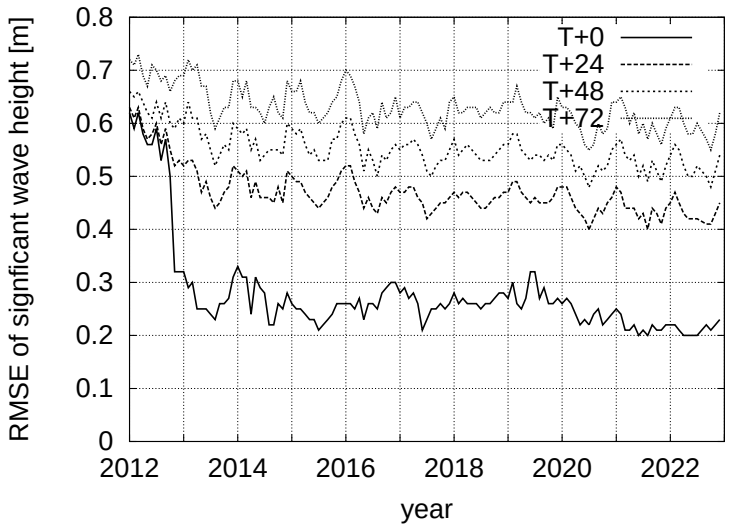


Figure 5.4.6: Time-series representation of analysis and forecast scores for the GWM.

Wave forecast errors are occasionally observed in relation to typhoons. For example:

1. Wave distribution resembling ball pairs sometimes appears due to the bogus wind embedding method.
2. As the model does not resolve the details of typhoon wind fields, especially in the extra-tropical stage, data on initial winds with bogus winds need to be improved.

5.5 Storm Surge Model

5.5.1 Japan Area Storm Surge Model

5.5.1.1 Introduction

The Japanese Archipelago is vulnerable to storm surges because of its topography characterized by gulfs open to the south and shallow coasts, making accurate and timely forecasts/warnings crucial in mitigating related threats to life and property.

The Japan Meteorological Agency (JMA) is responsible for issuing storm surge warnings/advisories for the area, and has operated a numerical storm surge model since July 1998 to provide basic related information. The model was initially run four times a day whenever a typhoon was present in the vicinity, and has undergone continuous enhancements including model domain enlargement, prediction of extratropical cyclone conditions, extension of forecast times and addition of advection terms. Since May 2010, a new storm surge model with higher resolution (approx. 1-km mesh near coastal areas) and gridded astronomical tide analysis have been implemented in storm tide calculation to produce even more practical information and warnings.

JMA also began providing information on the probability of warnings in 2017. This gave rise to the need for probabilistic prediction modeling specific to Japan for estimation of storm surge risk with a 3+ day lead time. To meet this demand, a new prediction system named Japan-area storm surge Probabilistic Forecast System (PFS; Subsection 5.5.1.6) was incorporated in August 2022.

5.5.1.2 Dynamics

Storm surges are mainly caused by the effects of wind setup due to strong onshore sea surface winds and inverse barometer effects associated with pressure drops in low-pressure systems. The effects of wind setup are proportional to the square of wind speed and inversely proportional to water depth, and are related to coastal topography, meaning that they are amplified in open bays against the wind.

The JMA storm surge model is similar to the one described in Higaki *et al.* (2009), and is based on two-dimensional shallow water equations driven by meteorological fields. These equations are composed of vertically integrated momentum equations in two horizontal directions:

$$\frac{\partial U}{\partial t} + u\frac{\partial U}{\partial x} + v\frac{\partial U}{\partial y} - fV = -g(D+\eta)\frac{\partial(\eta-\eta_0)}{\partial x} + \frac{\tau_{sx}}{\rho_w} - \frac{\tau_{bx}}{\rho_w} \tag{5.5.1a}$$

$$\frac{\partial V}{\partial t} + u\frac{\partial V}{\partial x} + v\frac{\partial V}{\partial y} + fU = -g(D+\eta)\frac{\partial(\eta-\eta_0)}{\partial y} + \frac{\tau_{sy}}{\rho_w} - \frac{\tau_{by}}{\rho_w} \tag{5.5.1b}$$

and the continuity equation:

$$\frac{\partial \eta}{\partial t} + \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0 \tag{5.5.2}$$

where U and V are volume fluxes in the x- and y-directions, defined as:

$$U \equiv \int_{-D}^{\eta} u\, dz \tag{5.5.3a}$$

$$V \equiv \int_{-D}^{\eta} v\, dz \tag{5.5.3b}$$

Other notations are as follows. f is the Coriolis parameter; g is gravity acceleration; D is the water depth below mean sea level; η is surface elevation; η_0 is the inverse barometer effect converted into the equivalent water column height; ρ_w is the density of water; τ_{sx} and τ_{sy} are the x- and y-components of wind stress on the sea surface, respectively; and τ_{bx} and τ_{by} are the x- and y-components of the stress of bottom friction, respectively. As for the drag coefficient, adjusted parameterization is adopted in reference to the results of Smith and Banke (1975) and Frank (1984):

$$c_{\mathrm{d}} = \begin{cases} (0.63 + 0.1185W) \times 10^{-3} \ (W < 20 \text{ m/s}) \\ \{3.00 + 0.0120(W - 25)\} \times 10^{-3} \ (W \geq 20 \text{ m/s}) \end{cases} \tag{5.5.4}$$

where W is wind speed.

The equations are solved via numerical integration using the explicit finite difference method. A staggered (Arakawa-C) grid (Arakawa and Lamb 1977) is adopted for the grid system.

5.5.1.3 Meteorological Forcing

Dynamically predicted meteorological fields (surface wind and sea-level pressure) from the Meso-Scale Model (MSM) are used for external forcing to drive the storm surge model. Simple parametric TC model values (referred to here as typhoon bogus) are also used for forcing when a tropical cyclone (TC) intensifying into a typhoon is expected in the area around Japan.

The main purpose of typhoon bogus is to predict storm surges based on official JMA typhoon forecasts, which are the most reliable at the time but not necessarily equivalent to those of the MSM. Single results remain inadequate for risk analysis because storm surges greatly depend on TC tracks. To take into account the influence of TC track uncertainty on storm surges, five runs of the storm surge model are conducted with potential TC tracks prescribed at the center of and at four points in the circle within which the TC is forecast to be with a probability of 70% (Figure 5.5.1). The five runs are (1) center track, (2) fastest track, (3) rightward-biased track, (4) slowest track, and (5) leftward-biased track. These data are used to create meteorological fields with the parametric TC model.

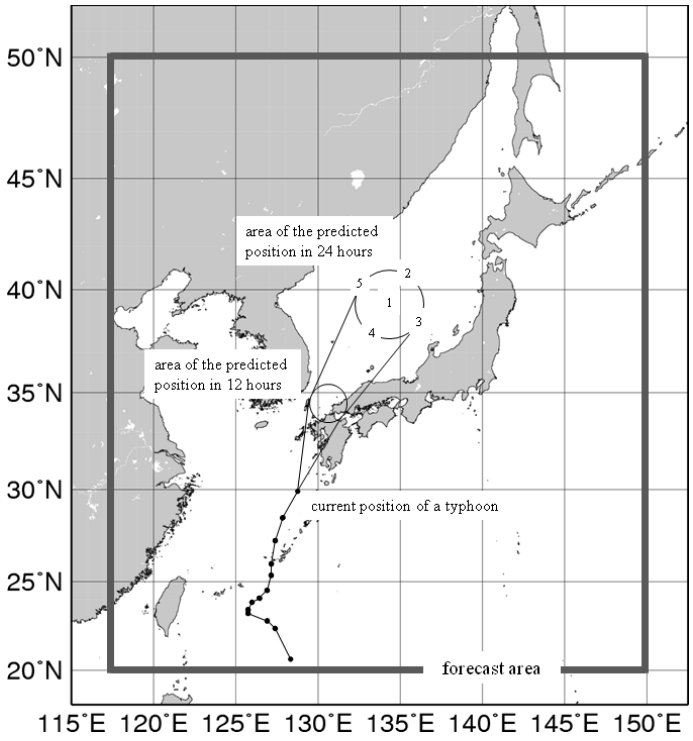


Figure 5.5.1: Bogus TC tracks and the domain of the Japan-area storm surge model

The typhoon bogus utilizes Fujita's formula (Fujita 1952), which represents radial pressure distribution in a TC:

$$P = P_{\infty} - \frac{P_{\infty} - P_{\mathrm{c}}}{\sqrt{1 + (r/r_0)^2}} \tag{5.5.5}$$

and the gradient wind relation:

$$-\frac{v_g^2}{r} - f v_g = -\frac{1}{\rho_a}\frac{\partial P}{\partial r} \tag{5.5.6}$$

In Eq. (5.5.5), P is atmospheric pressure at distance r from the center of the TC, P_∞ is environmental pressure, P_c is the central pressure of the TC and r_0 is a scaling factor for radial distribution of pressure. In Eq. (5.5.6), ρ_a is the density of air and v_g is the gradient wind speed.

To represent the asymmetry of the wind field $\boldsymbol{w}$ in a TC, the moving velocity vector of the TC multiplied by a weight that decays exponentially with distance from the TC center is added to the gradient wind:

$$\boldsymbol{w} = C_1 \left\{ \boldsymbol{v_g} + \boldsymbol{C} \cdot \exp\left(-\pi \frac{r}{r_e}\right) \right\} \tag{5.5.7}$$

$\boldsymbol{C}$ is the TC velocity vector, and r_e is the decay coefficient.

Moreover, an upwind directional land roughness parameterization of Westerink *et al.* (2008) is adopted to represent the wind speed reduction in coastal areas.

TC analysis and forecast information, such as the center position, central pressure and maximum wind, is applied to these formulas to synthesize the wind and pressure fields (Konishi 1995).

5.5.1.4 Model Specifications

Table 5.5.1 gives the specifications of the storm surge model, whose domain covers the whole of Japan (Figure 5.5.1).

Table 5.5.1: Japan-area storm surge model specification

Model	two-dimensional model
Grid	Lat-Lon Arakawa-C grid
Region	20°N - 50°N, 117.4°E - 150°E
Resolution	approximately 1, 2, 4, 8, 16 km (Adaptive mesh)
Boundary conditions	Modified radiation condition at open boundaries and zero normal flows at coastal boundaries
Time step	4 seconds
Initial time	00, 03, 06, 09, 12, 15, 18, 21 (UTC)
Forecast time	78 hours (00, 12 UTC) and 39 hours (03,06,09,15,18,21 UTC)
Member	TC case: 6 members (MSM+5 typhoon bogus) no TC case: 1 member (MSM)

Since storm surge is essentially a long wave, its phase speed is proportional to the square root of water depth. It is thus inefficient to set the same resolution for all grids in consideration of computer resources. Accordingly, the Adaptive Mesh Refinement (Berger and Oliger 1984), in which the mesh is fine over shallow water and coarse over deep water, is adopted. The resolution is varied over five levels (1, 2, 4, 8 and 16 km) with water depth (Figure 5.5.2). This method makes storm surge calculation efficient, compared to the normal lat-lon grid system.

The storm surge model runs every three hours for forecasts up to 78 or 39 hours ahead. Initial conditions such as surface elevation (η) and volume fluxes (U and V) are created from the previous calculation using the latest MSM prediction as forcing (hindcasts). As these conditions are less important than those in atmospheric models, observation data are not assimilated.

The model computes only storm surges, i.e. anomalies from the level of astronomical tides. However, storm tides (storm surge plus astronomical tide) are required to issue storm surge warnings. Astronomical tides are predicted using harmonic analysis of sea level observations based on JMA's gridded astronomical tide method, in which astronomical tide calculation is performed even for no-observation grid areas (Subsection

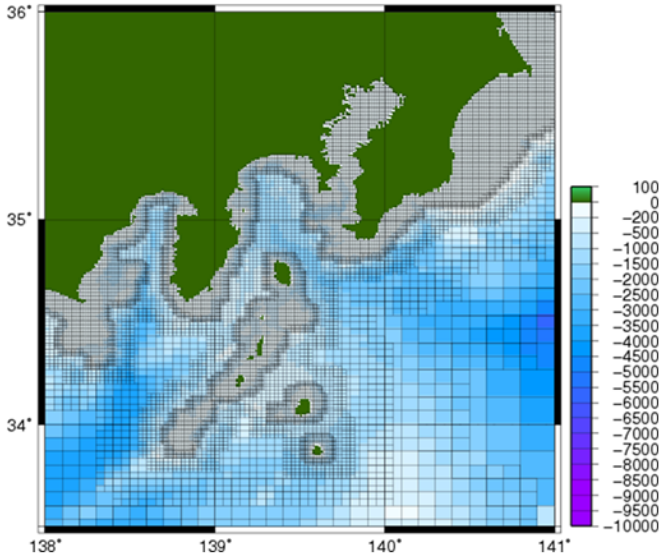


Figure 5.5.2: Storm surge model horizontal grid system and water depth (around the Kanto region)

5.5.3). After storm surge model computation, the astronomical tide level for the coastal area is linearly added to the predicted storm surge.

5.5.1.5 Verification

The accuracy of storm surge prediction depends on that of the storm surge model itself and atmospheric model conditions. To help eliminate forecast errors in atmospheric model input, the accuracy of the model was evaluated using storm surge predictions driven by atmospheric analysis data.

Figure 5.5.3 shows a scatter diagram of storm surge hindcasts against observation values from 210 tide stations for statistical periods of TC presence from 2019 to 2022. These stations are managed by organizations including JMA, the Ports and Harbours Bureau, the Japan Coast Guard and the Geospatial Information Authority of Japan. The figure shows prediction errors in the range of ± 50 cm, although large errors (exceeding observation by more than 50 cm) are also observed. These errors may be attributable to factors not considered in the storm surge model, such as wave setup, ocean currents and sea water stratification.

Storm surge associated with Typhoon Jebi (T1821) is presented here as an example of related prediction. Figure 5.5.4 shows the track of the typhoon, which passed over central Japan and caused storm surge conditions in Osaka Bay and elsewhere on 4 September 2018. Figure 5.5.5 (a) shows storm surge distribution around the bay as of 06 UTC on 4 September as predicted by the storm surge model, which reasonably forecasted the large surge behaviors associated with wind setup in the bay's inner part. Figure 5.5.5 (b) shows a time-series chart of storm surge at the port of Osaka. The peak was slightly underestimated and the forecast of its timing was delayed by about an hour, but in general the extreme storm surge conditions were fairly predicted.

5.5.1.6 Probabilistic Prediction

When a typhoon is forecast, the Japan-area storm surge model calculates predictions based on a potential ensemble system with six possible scenarios. As this number is insufficient for consideration of typhoon forecast uncertainties up to five days ahead, JMA developed a new Japan-area storm surge PFS that provides perturbations to TC tracks with reference to the Probabilistic storm Surge (P-Surge) model (Taylor and Glahn 2008; Gonzalez and Taylor 2018) operated by the National Hurricane Center (NOAA/NHC).

Most system specifications are common to those of deterministic modeling, with 21 members and no

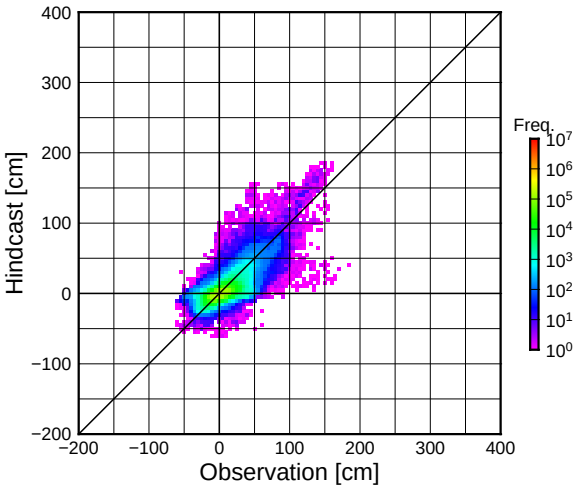


Figure 5.5.3: Scatter diagram of storm surge hindcast values against observed values

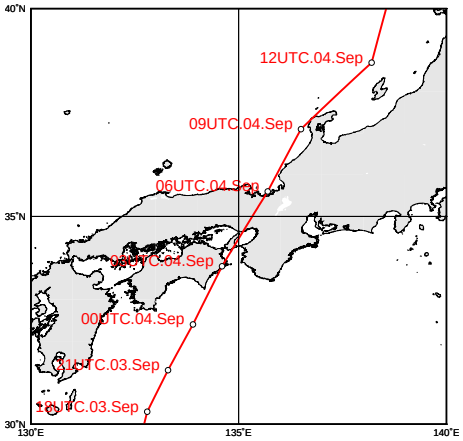
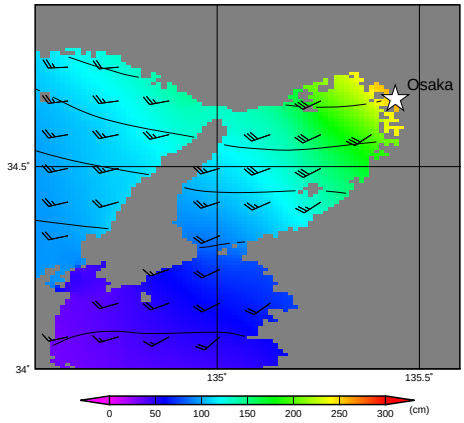
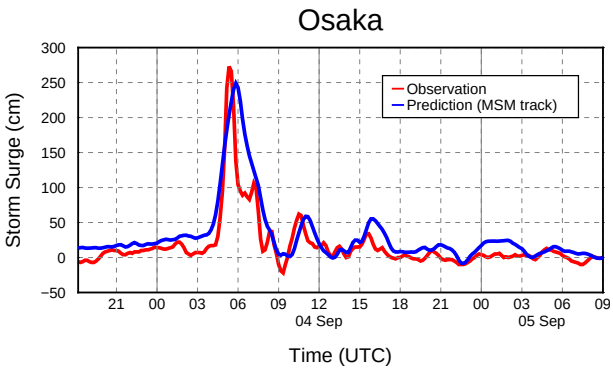


Figure 5.5.4: The track of Typhoon Jebi (T1821)



(a) Distribution of storm surge at 06 UTC on 4 September around Osaka Bay



(b) Time-series chart of storm surge at the port of Osaka

Figure 5.5.5: Storm surge distribution and time-series chart for the port of Osaka as predicted by the storm surge model with an initial time of 18 UTC on 3 September 2018

execution during non-TC periods. The PFS is driven only by typhoon bogus as a meteorological forcing factor.

Figure 5.5.6 shows PFS track placement, with each member set with horizontal shifting of the official TC track at equal intervals. The overall tracks of cross-track members cover a 70% probability circle and an outer margin region. The PFS implants typhoon bogus into 21 tracks and operates storm surge predictions up to 132 hours (5.5 days) ahead. In consideration of time-direction errors, prediction post-processing is implemented to create along-track members by shifting each PFS cross-track member with time interpolation in 21 ways to create 441 predictions. Figure 5.5.7 shows an example of such calculation for storm tide data at Osaka for Typhoon Jebi. In such cases, warning criteria breaches can be predicted for some members.

Probabilistic prediction requires assignment of a probability density to each member. In the system, two-dimensional normal distribution is assumed for probability density distribution that TC center locations may assume with respect to typhoon forecast circles, with members closer the circle center being assigned higher probabilities. Based on probability densities of members overall, predictions of variables such as storm surge potential and storm tide percentiles are calculated, and operational products are created accordingly (see Figure 5.5.8).

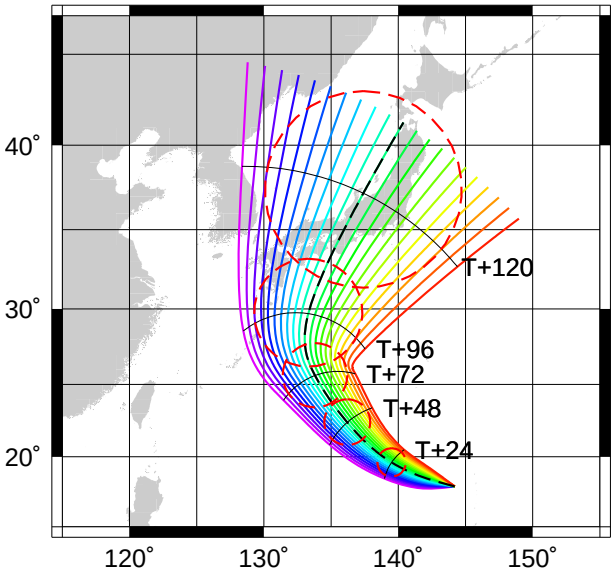


Figure 5.5.6: PFS member tracks

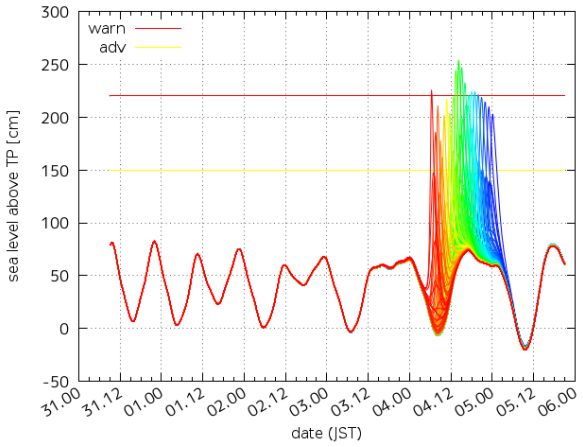


Figure 5.5.7: Time-series representation of storm tides at the port of Osaka as predicted by the PFS with an initial time of 00 UTC on 31 August 2018

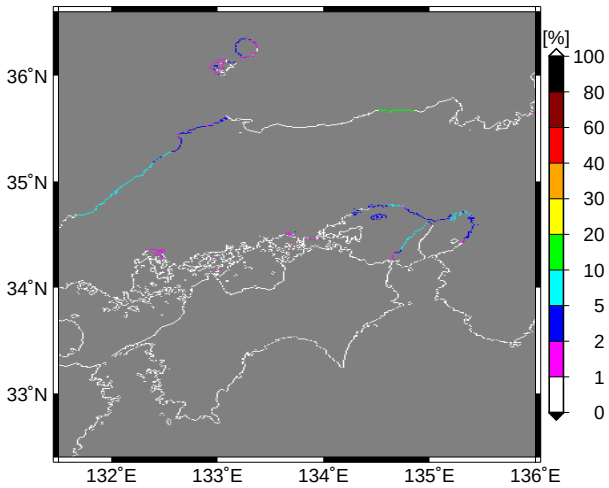


Figure 5.5.8: Probability of storm tide warning criteria breach on the fourth day based on PFS products with an initial time of 00 UTC on 2 September 2022

5.5.2 Asia Area Storm Surge Model

5.5.2.1 Introduction

The late 2000s saw severe storm surge disasters worldwide, including on the coast of the Gulf of Mexico (caused by Hurricane Katrina) in 2005, on the coast of Bangladesh (caused by Cyclone Sidr) in 2007, and on the coast of Myanmar (caused by Cyclone Nargis) in 2008.

In response to a request by the WMO Executive Council (60th session, June 2008), WMO initiated the development of the regional Storm Surge Watch Scheme (SSWS) for areas affected by tropical cyclones. In relation to the western North Pacific and the South China Sea, the ESCAP/WMO Typhoon Committee (41st session, January 2009) endorsed a commitment by the RSMC Tokyo - Typhoon Center to produce storm surge forecasts with the aim of strengthening the storm surge warning capabilities of National Meteorological and Hydrological Services (NMHSs) in the region. JMA began development of a storm surge model for the Asia region in 2010 in collaboration with Typhoon Committee Members providing sea level observation and sea bathymetry data. Horizontal distribution maps of predicted storm surges and time-series charts are published on JMA's Numerical Typhoon Prediction website (Hasegawa *et al.* 2017).

In the last decade, Hurricane Sandy made landfall on the eastern coast of USA in 2012, causing serious damages including the traffic paralysis, massive blackouts and cessation of economic activity in New York. Typhoon Haiyan also caused more than 6,000 fatalities in the Philippines in 2013. Against such a background, storm surge and inundation countermeasures play pivotal roles in efforts to prevent typhoon-related disaster condition.

5.5.2.2 Dynamics

The basic equations for the Asia-area storm surge model are essentially the same as those for the Japan-area model (Subsection 5.5.1), with the exception of definition in a spherical coordinate system. The expressions incorporate vertically integrated momentum fluxes associated with the influence of the earth's rotation with gravity acceleration:

$$\frac{\partial U}{\partial t} + \frac{1}{R\cos\varphi}\left(\frac{\partial u^2 H}{\partial \lambda} + \frac{\partial uvH\cos\varphi}{\partial \varphi}\right) - fvH - \frac{uvH}{R}\tan\varphi = -\frac{gH}{R\cos\varphi}\frac{\partial(\eta-\eta_0)}{\partial\lambda} + \frac{\tau_{s\lambda}}{\rho_w} - \frac{\tau_{b\lambda}}{\rho_w} \tag{5.5.8a}$$

$$\frac{\partial V}{\partial t} + \frac{1}{R\cos\varphi}\left(\frac{\partial uvH}{\partial \lambda} + \frac{\partial v^2 H\cos\varphi}{\partial \varphi}\right) + fuH + \frac{u^2 H}{R}\tan\varphi = -\frac{gH}{R}\frac{\partial(\eta-\eta_0)}{\partial\varphi} + \frac{\tau_{s\varphi}}{\rho_w} - \frac{\tau_{b\varphi}}{\rho_w} \tag{5.5.8b}$$

and the continuity equation:

$$\frac{\partial \eta}{\partial t} + \frac{1}{R\cos\varphi}\left(\frac{\partial uH}{\partial \lambda} + \frac{\partial vH\cos\varphi}{\partial \varphi}\right) = 0 \tag{5.5.9}$$

Here, H is total depth, R is the earth's radius, and (λ, φ) are the zonal and meridional direction coordinates. Other values are as per those of the Japan-area storm surge model. Wind stress is expressed as:

$$\tau_{s\lambda} = c_d\,\rho_a W u_w \tag{5.5.10a}$$

$$\tau_{s\varphi} = c_d\,\rho_a W v_w \tag{5.5.10b}$$

where c_d is the drag coefficient, ρ_a is the density of air, $W \equiv \sqrt{u_w^2 + v_w^2}$ is the wind speed and (u_w, v_w) is the wind velocity. The drag coefficient is set based on the results of Peng and Li (2015):

$$c_d = \begin{cases} (3.146 - 0.00188 \times (W-33)^2) \times 10^{-3} & (W < 60\ \text{m/s}) \\ (1.5 + 0.2755 \times e^{-0.3685(W-60)}) \times 10^{-3} & (W \geq 60\ \text{m/s}) \end{cases} \tag{5.5.11}$$

5.5.2.3 Data

Bathymetry data for the Asia-area storm surge model come from 15-second-interval grid datasets of the General Bathymetric Chart of the Oceans (GEBCO 2021) (Figure 5.5.9).

Astronomical tides are determined via harmonic analysis with past tide observation data provided by Typhoon Committee Members and results from FES2014[9] (Lyard *et al.* 2021), allowing calculation of astronomical tides for stations where observation data are not available.

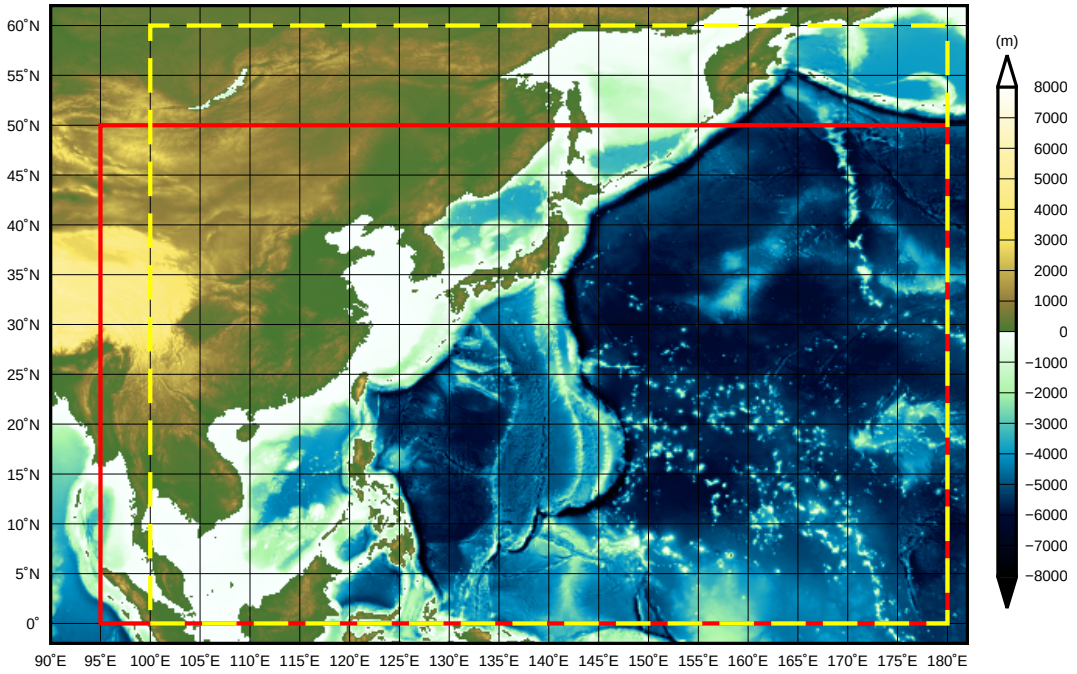


Figure 5.5.9: GEBCO 2021 topography. The red line indicates the domain of the Asia-area storm surge model, and the dashed yellow line denotes the RSMC Tokyo - Typhoon Center's area of responsibility.

5.5.2.4 Meteorological Forcing

Operation of the Asia-area storm surge model involves the blending of simple parametric TC model data and output from JMA's operational Global Spectral Model (GSM) as meteorological forcing fields. The simple parametric TC model in this resource is as per that of the Japan area storm surge model (typhoon bogus in Subsection 5.5.1). Related calculation requires an atmospheric model covering the Asian region, but the resolution of the GSM (13 km) is insufficient for resolving adequate TC intensity. Accordingly, meteorological forcing is generated by implanting typhoon bogus information into atmospheric model gridded data.

5.5.2.5 Ensemble Prediction

The Asia-area storm surge model previously adopted multi-scenario predictions (Hasegawa *et al.* 2017), and six scenarios were considered. However, this was insufficient for risk management because the occurrence and intensity of storm surges strongly depend on TC tracks, and probabilistic forecast products could not be provided in this way. Against such a background, JMA introduced an ensemble prediction system with 52 members and began providing ensemble forecast products (Hasegawa *et al.* 2023) using the GSM with typhoon bogus and the Global Ensemble Prediction System (GEPS) (Kyouda and Higaki 2015), which has 51 members.

[9]FES2014 was produced by NOVELTIS, LEGOS, CLS Space Oceanography Division and CNES. It is distributed by AVISO, with support from CNES (http://www.aviso.altimetry.fr/).

As the horizontal resolution of the GEPS (27 km) is considered too coarse for adequate TC prediction, typhoon bogus is implemented for TC tracks of all GEPS members as in the GSM.

5.5.2.6 Model Specifications

Table 5.5.2 outlines the specifications of the Asia-area storm surge model. Here, an unstructured Arakawa-B approach is adopted with a grid system, and a finite volume approach is used for equation resolution with reference to the Finite Volume Community Ocean Model (FVCOM) (Chen *et al.* 2003) and the Finite-volumE Sea ice-Ocean Model (FESOM2) (Danilov *et al.* 2017). The resolution of the unstructured grid ranges from 1.5 to 50 km, becoming fine in coastal areas and coarse in open seas (Figure 5.5.10). The model covers most of the RSMC Tokyo - Typhoon Center's area of responsibility (Figure 5.5.9), running every six hours and calculating storm surge predictions up to 132 hours ahead. When no TCs are present, a single calculation is conducted with GSM prediction. If one or more TCs are present or expected in the model domain, ensemble predictions for 51 members are additionally carried out based on the GEPS with wind and pressure fields modified by typhoon bogus.

Three-hourly distribution maps of the whole domain and enlarged versions showing only areas around the TC are available up to 132 hours ahead. The time-series charts provided include data on predicted/astronomical tides, storm surge, sea level pressure and surface wind. Time-series charts for 78 locations are currently provided to Typhoon Committee Members.

Table 5.5.2: Specifications of the Asia area storm surge model

Model	two-dimensional non-linear model
Grid	unstructured Arakawa-B grid
Region	0° - 50°N, 95°E - 180°E
Resolution	1.5 - 50 km
Time step	4 seconds
Initial time	00, 06, 12, 18 (UTC)
Forecast period	132 hours
Member	TC case: 52 members (GSM + GEPS 51 members with typhoon bogus) no TC case: 1 member (GSM)

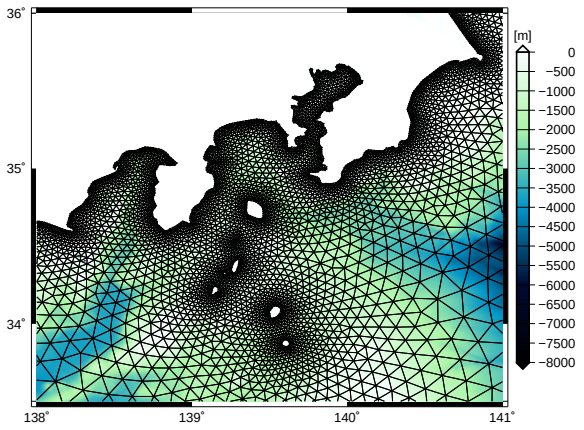


Figure 5.5.10: Unstructured grid around the Kanto region in the Asia-area storm surge model. JIGSAW (Engwirda 2017) is used as a mesh-generating tool.

5.5.2.7 Verification

To evaluate the performance of the Asia-area storm surge model, accuracy was verified by comparing predicted and observed values of hourly storm surge for numerous tide stations in the Japan and Asia areas over the statistical period from June 2018 to March 2021. Figure 5.5.11 shows a scatter diagram of storm surge hindcasts against observation values for 214 tide stations in the Japan area (left) and eight tide stations in the Asia area (right). Note that the GSM resolution is 20 km in this statistical period. Most storm surge prediction errors lie within the range of ± 50 cm in both panels, although some underestimation is seen for reasons similar to those detailed in Subsection 5.5.1.5 for the Japan area. For the Asia area, underestimations were seen at Quarry Bay (Hong Kong) for the intense Typhoon Mangkhut (T1822), and the atmospheric model (GSM) did not adequately resolve sea level pressure and surface wind. Inaccuracy in topographical expression and calculation in the Asia-area storm surge model also caused underestimation.

Yearly verification details are provided in the Annual Report on the Activities of the RSMC Tokyo - Typhoon Center [10].

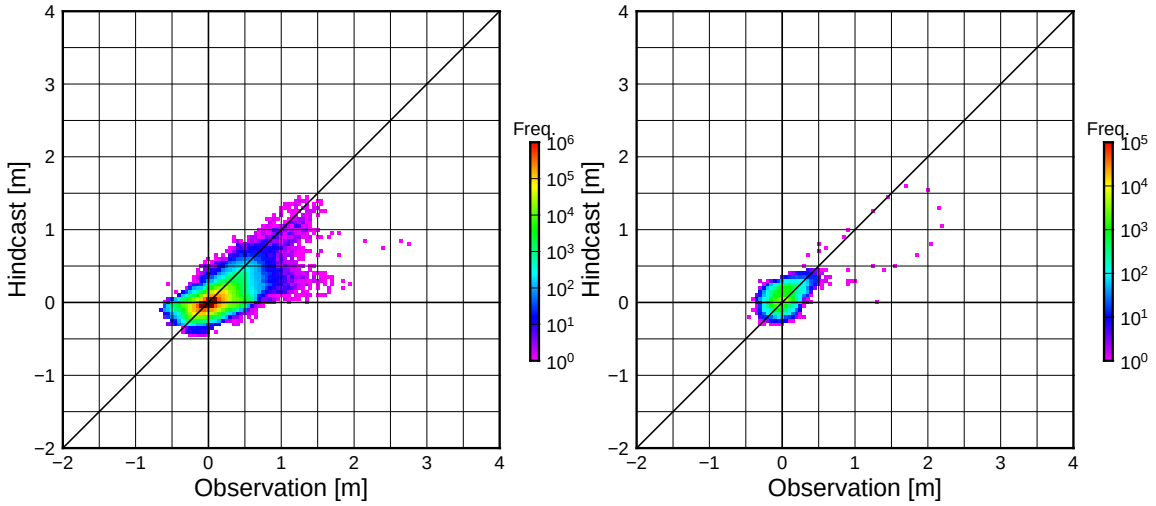


Figure 5.5.11: Scatter diagram of storm surge hindcast values against observations. Left: results for 214 tide stations in the Japan area; right: results for eight tide stations in the Asia area.

Figures 5.5.12 and 5.5.13 show distribution mapping for the probability of storm surges exceeding 1 meter in height during the forecast period and time-series charts for Quarry Bay (Hong Kong) for Typhoon Ma-on (T2209), respectively. The typhoon passed northwest over the South China Sea and generated high storm surges around Hong Kong. The results show that ensemble predictions captured the potential for high storm surges and peak storm surge/tide conditions well.

5.5.3 Astronomical Tide Analysis

5.5.3.1 Introduction

The model described in Subsection 5.5.1 calculates only storm surges, defined as anomalies from the astronomical tide level. However, prediction of storm tides (i.e., storm surge plus astronomical tides) is needed for storm surge warning issuance. In 2010, JMA changed its storm surge warning criteria to cover issuance for all coastal areas of Japan in consideration of inundation risk at all points. Appropriate issuance of warnings requires calculation to determine astronomical tides in all coastal areas.

[10] http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/annualreport.html

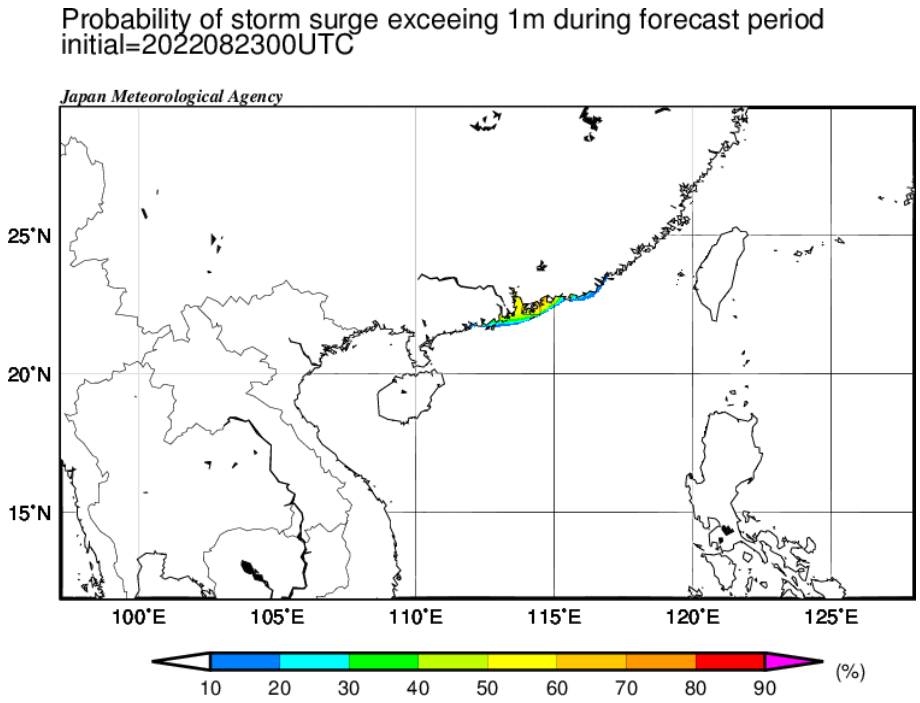


Figure 5.5.12: Distribution for the probability of storm surges exceeding 1 meter in height during the forecast period as predicted by the Asia-area storm surge ensemble prediction system with an initial time of 00 UTC on 23 August.

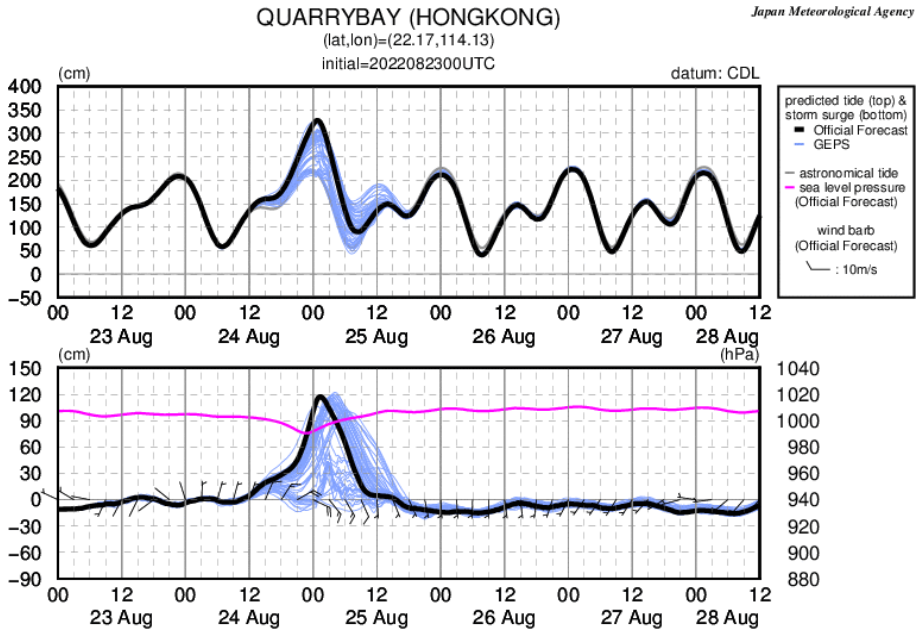


Figure 5.5.13: Time-series charts of storm/astronomical tides (top), storm surge, sea level pressure and surface wind (bottom) for Quarry Bay (Hong Kong) as per the Asia-area storm surge ensemble prediction system with an initial time of 00 UTC on 23 August. The black lines show observations.

5.5.3.2 Analysis Method

Tidal variations are expressed as a composite of periodic oscillations with various frequencies, as observed with semi-diurnal, diurnal and annual tides. Semi-diurnal and diurnal tides are caused by lunar and solar gravitational forces, while annual tides are brought by seasonal variations in seawater temperature and sea surface pressure. Harmonic constants are sets of amplitudes and phases for individual tidal constituents. Harmonic constants at tide station points can be derived by analyzing hourly tidal observation data, but cannot be determined for arbitrary coastal points where such data are unavailable using this method.

To enable analysis of astronomical tides for the whole of Japan's coast, both short-period (semi-diurnal and diurnal) and long-period tides (annual) tides are considered as shown in Figure 5.5.14. The constituents used in this method (Takasa *et al.* (2011)) are shown in Table 5.5.3.

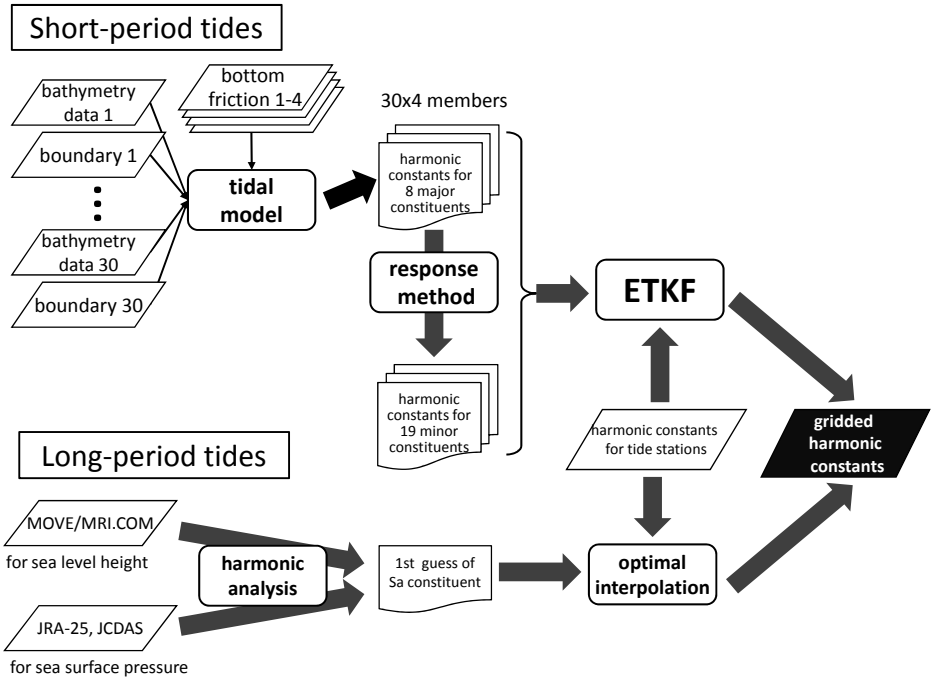


Figure 5.5.14: Flow of astronomical tide analysis

5.5.3.3 Short-period Tides

Eight major constituents with relatively large amplitudes (K1, O1, P1, Q1, K2, M2, N2 and S2; Table 5.5.3) are estimated using the Oregon State University Tidal Inversion Software (OTIS) tidal model package (Egbert and Erofeeva 2002), which involves data preparation, ocean dynamics and data assimilation. Only the linearized version of ocean dynamics (a tidal model), in which Fourier transform is applied to eliminate time variation, is used:

$$i\omega U - fV + gH\frac{\partial \zeta}{\partial x} + \kappa U = F_U \tag{5.5.12}$$

$$i\omega V + fU + gH\frac{\partial \zeta}{\partial y} + \kappa V = F_V \tag{5.5.13}$$

$$\left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right) + i\omega\zeta = 0 \tag{5.5.14}$$

Table 5.5.3: Constituents used in astronomical tide analysis.

Name	Type	Major/Minor	Name	Type	Major/Minor
S_a	annual	-	θ_1	diurnal	minor
$2Q_1$	diurnal	minor	J_1	diurnal	minor
σ_1	diurnal	minor	OO_1	diurnal	minor
Q_1	diurnal	major	$2N_2$	semi-diurnal	minor
ρ_1	diurnal	minor	μ_2	semi-diurnal	minor
O_1	diurnal	major	N_2	semi-diurnal	major
MP_1	diurnal	minor	ν_2	semi-diurnal	minor
M_1	diurnal	minor	M_2	semi-diurnal	major
χ_1	diurnal	minor	λ_2	semi-diurnal	minor
π_1	diurnal	minor	L_2	semi-diurnal	minor
P_1	diurnal	major	T_2	semi-diurnal	minor
K_1	diurnal	major	S_2	semi-diurnal	major
ψ_1	diurnal	minor	R_2	semi-diurnal	minor
ϕ_1	diurnal	minor	K_2	semi-diurnal	major

where ω is the tidal constituent frequency, U and V are the x and y components of current integrated from the sea surface to the bottom, respectively, f is the Coriolis parameter, g is gravity acceleration, H is depth, ζ is the anomaly from mean sea level, κ is the dissipation coefficient of bottom friction, and F is the tide-generating force.

The Ensemble Transform Kalman Filter (ETKF) (Bishop *et al.* 2001) is used to assimilate harmonic constants at tide stations. As the model results contain uncertainty due to a lack of resolution and accuracy in bathymetry data and lateral boundary conditions, perturbations are added to these conditions to create an ensemble. There are 30 sets of bathymetry data (incorporating random errors) and boundary condition data generated by blending results from four tidal models (NAO.99Jb (Matsumoto *et al.* 2000), FES2004 (Lyard *et al.* 2006), GOT00.2 (an update to Ray (1999)) and TPXO (Egbert and Erofeeva 2002)). Four sets of bottom friction data are also used in consideration of the influence of such friction on tidal amplitude (Yano *et al.* (2010) and An (1977)). The 120 ensemble members are associated with a combination of the 30 sets of bathymetry and boundary condition data and 4 sets of bottom friction data.

A total of 19 minor constituents (see Table 5.5.3) are estimated from major constituents of similar frequency using the response method (Munk and Cartwright 1966).

5.5.3.4 Long-period Tides

The first guess of annual constituents (S_a) is derived from the results of harmonic analysis of reanalyzed sea level height from MOVE-WNP (see Section 5.3 and Usui *et al.* (2006)) corrected with sea surface pressure from the Japanese 25-year Reanalysis (JRA-25) and the JMA Climate Data Assimilation System (JCDAS) (Onogi *et al.* (2007)) assuming hydrostatic balance. This is modified using harmonic constants for tide stations with the optimal interpolation (OI) method.

5.5.3.5 Verification

To verify astronomical tide analysis based on the method described, the outcomes are compared with those from harmonic constants at tide stations. Figure 5.5.15 shows a comparison of root mean square errors (RMSEs) for the respective distributions. For most stations, the error is less than 3 cm, although larger values are observed for some stations, especially in bays and inland sea areas.

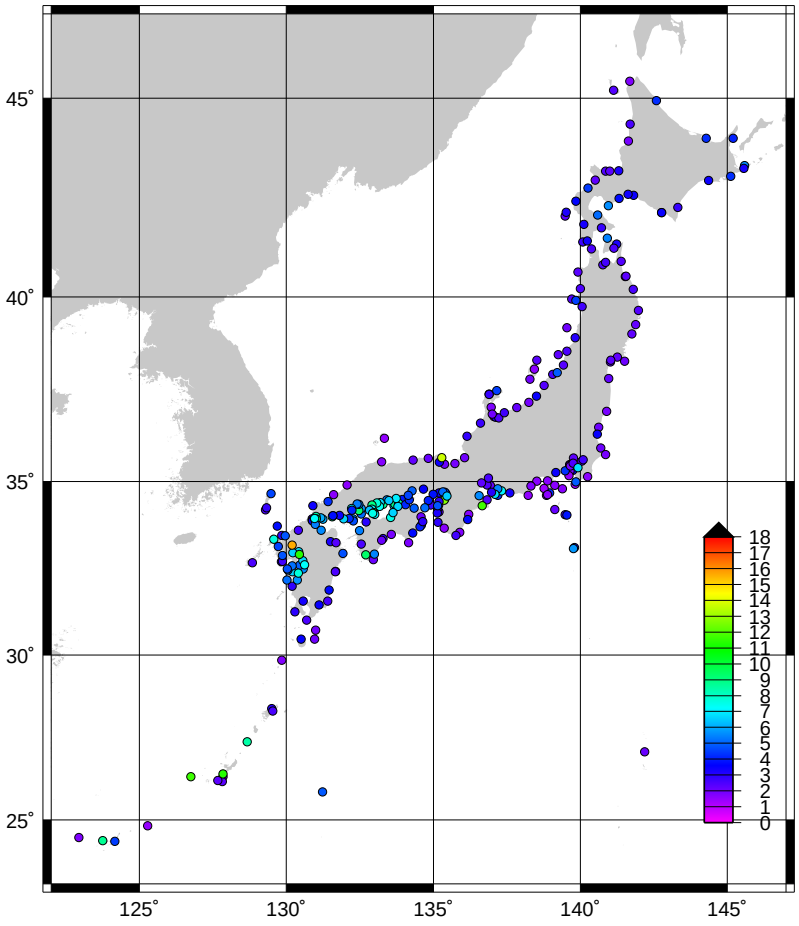


Figure 5.5.15: Distribution of RMSEs from astronomical tide analysis. The unit is cm.

5.6 Oil Spill Prediction Model

5.6.1 Introduction

In the 1990s, large-scale oil spills (such as those involving the vessels ABT Summer, MB Braer and Sea Empress) frequently occurred around the world. In Japan, the wrecking of the Russian tanker Nakhodka resulted in a serious oil spill in the Sea of Japan in January 1997, causing major environmental damage along Japan's western coast. Following the Japanese Government's subsequent consideration of countermeasures for large-scale oil spills, JMA has operated its Oil Spill Prediction Model since October 1999.

The model predicts the large-scale behavior of oil spilled in offshore seas. Based on accident information from the Japan Coast Guard, JMA operates the model to produce forecasts with lead time up to 192 hours. The results are provided to personnel responsible for emergency response operations.

The model is applicable to the entire western North Pacific. The domain of calculation is selected from seven settings from $0.8^\circ \times 0.8^\circ$ to $12^\circ \times 12^\circ$ in latitude and longitude based on consideration of incident conditions.

5.6.2 Basic Equation

The oil spill prediction model is generally described by the following equation including terms of advection and diffusion,

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \mathbf{V} \cdot \nabla C = \nabla \cdot (K \nabla C) + S \tag{5.6.1}$$

where C is pollutant concentration, t is time, $\mathbf{V}$ is advection velocity, K is the turbulent diffusion coefficient, and S (referred to as the source term) represents the process involved in changing the total amount of oil spilled through changes in oil properties.

Equation (5.6.1) in general can be numerically solved either by calculating C directly using the finite difference method or by simulating the behaviors of a large number of particles representing oil components. The latter approach is used in JMA's Oil Spill Prediction Model. Spilled oil is expressed as numerous particles $C_n (n = 1, 2, \cdots)$ using:

$$C_n \{\mathbf{x}(t + \delta t), s(t + \delta t); t + \delta t\} = \Phi[C_n (\mathbf{x}(t), s(t); t), \delta t] \tag{5.6.2}$$

where $\mathbf{x} = (x, y, z)$ indicates the position of each particle and s is the chemical status of the oil. Φ is a general function describing oil property changes over time.

In the advection term, the effects of surface winds, ocean waves, and ocean currents are taken into account as potentially major factors. Ekman drift current generated by sea surface winds is an example of such as an influence. In the JMA model, surface flows are determined as 2.5% of the wind speed with an angle of 15° clockwise with respect to the wind direction. As another example, Stokes drift involves forward movement of particles at the sea surface in the wave direction as a result of wave motion back and forth in each wave cycle. This effect is more significant when high waves are present, and is independent of wind when swell is predominant. Accordingly, Stokes drift is included explicitly and calculated from predictions by JMA ocean wave models. Ocean currents are provided by the JMA Ocean Data Assimilation and Prediction System for the seas around Japan (MOVE-JPN; Subsection 5.3.3).

The three-dimensional diffusion of oil is basically calculated via the shear diffusion treatment proposed by Elliott (1986). Surface flow is assumed to have a logarithmic profile in the vertical direction, and spilled oil is assumed to be carried at a particular horizontal speed in each water level. The shear mechanism is also associated with vertical diffusion. Spilled oil is divided into a large number of droplets with varying levels of buoyancy in line with their size. Consideration of this buoyancy and the present depth of oil drops allows determination of oil motion in the vertical direction and clarification of whether the oil floats on the surface.

In addition to the above shear diffusion process, isotropic diffusion may also be generated by small scale eddies and similar influences as estimated using the constant diffusion coefficient $K_h = 95.0\ \mathrm{m}^2/\mathrm{s}$. As such diffusion may be greater in conditions of strong wind or high waves, the influence is parameterized with additional diffusion coefficients:

Table 5.6.1: Oil Spill Prediction Model specifications

Applicable area		15°S – 63°N, 99°E – 80°W
Domain of calculation		7 options ($0.8° \times 0.8° - 12° \times 12°$)
Grid spacing		7 options (2–30km), according to the domain of calculation
Number of grids		41×41
Prediction period		192 hours
Physical and chemical process	Advection	Ekman drift (estimated from wind field of Global Spectrum Model) Stokes drift (estimated from wave field of Global/Coastal Wave Models) Ocean current (MOVE-JPN)
	Diffusion	Elliott (1986) etc.
	Evaporation	Fingas (2010)
	Emulsification	Reed (1989)

$$\begin{cases} \text{waves: } K_{wv} & = 500.0 H_w^2 / T_w \\ \text{winds: } K_{wnd} & = 5.0 W^3 / g \end{cases} \tag{5.6.3}$$

where H_w and T_w are the wave height and period, W is wind speed, and g is gravitational acceleration. The coefficients are empirically determined on the basis of actual cases.

Additional diffusion of oil parcels is estimated from the total value of the diffusion coefficients (K_h, K_{wv}, and K_{wnd}). Specific values are calculated using the random walk method with such diffusion assumed to be horizontal.

Due to the complex behavior of spilled oil, consideration of all related chemical processes is largely impractical. Accordingly, only evaporation and emulsification are considered as major processes. Evaporation is estimated using empirical formulae (Fingas 2010), in which the evaporation rate E_v (%) of most oils can be expressed by the form of either the logarithmic or the root profile over time.

$$E_v = \begin{cases} (a + b \cdot T) \ln t \\ (a + b \cdot T) \sqrt{t} \end{cases} \tag{5.6.4}$$

The constant coefficients a and b are based on experimental results and are listed in the Environment Canada oil data catalogue. T represents oil temperature, and is assumed to be equivalent to sea surface temperature (SST). t is the number of minutes elapsed since the spill.

Emulsification is calculated using the formula of Reed (1989), by which the water content F_{wc} is estimated as:

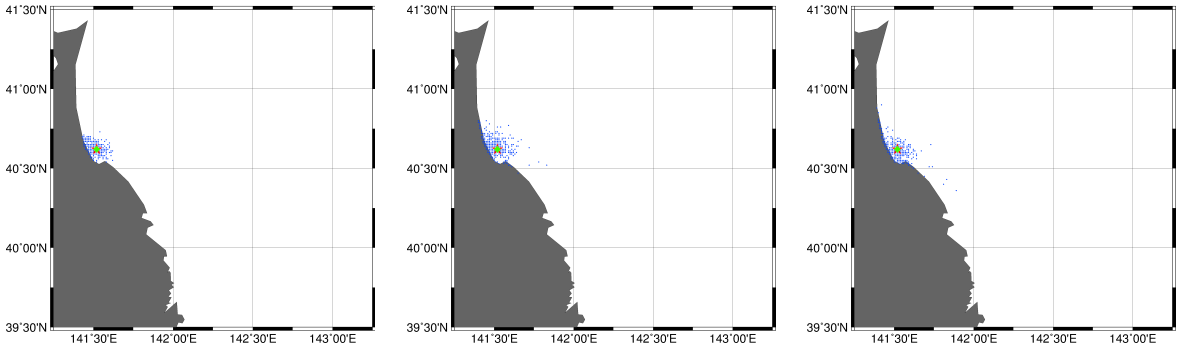
$$\frac{dF_{wc}}{dt} = 2.0 \times 10^{-6} (W + 1)^2 \cdot \left(1 - \frac{F_{wc}}{C_3}\right) \tag{5.6.5}$$

where W (m/s) is wind speed. C_3 is a constant parameter for the upper limit of water content, and differs among oil types. Oil density is also calculated in consideration of water content, which can change the behavior of oil.

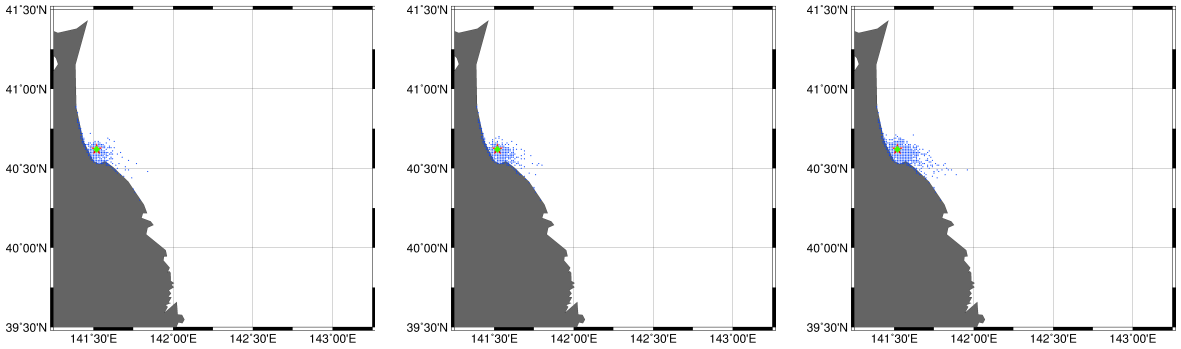
The specifications of the Oil Spill Prediction Model and related processes are summarized in Table 5.6.1.

5.6.3 Products

The model is operated in the event of a large-scale oil spill in offshore deep-water seas, where short-term tidal currents can be negligible. The results of oil spill prediction are provided to the Japanese Government and/or the Japan Coast Guard along with various marine meteorological charts. An example of prediction is shown in Figure 5.6.1.



(a)Forecast period: 24 h (b)Forecast period: 48 h (c)Forecast period: 72 h



(d)Forecast period: 96 h (e)Forecast period: 120 h (f)Forecast period: 144 h

Figure 5.6.1: A simulation for the sea east of Japan supposing an accident at 19UTC on 11 Aug 2021 at 40.37 °N 141.31 °E. The stars show the accident location, and the rhombuses show the source of the spillage (assuming the source is carried by ocean currents). The area and amount of spilled oil are indicated by dot distribution.

Appendix A

Verification Indices

This appendix highlights a number of verification indices referenced in this document. The indices are also used in international verification via the Global Data-processing and Forecasting System (GDPFS) of the World Meteorological Organization (WMO 2019).

A.1 Basic Verification Indices

A.1.1 Mean Error

Mean Error (ME), also called Bias, represents the mean value of deviations between forecasts and verification values, and is defined by

$$\mathrm{ME} \equiv \left(\sum_{i=1}^{n} w_i D_i \right) \Big/ \sum_{i=1}^{n} w_i, \tag{A.1.1a}$$

$$D_i = F_i - A_i, \tag{A.1.1b}$$

$$w_i = \frac{1}{n} \ (\text{or } \cos\phi_i, \text{ and so on}), \tag{A.1.1c}$$

where F_i, A_i, and D_i represent the forecast, the verification value, and the deviation between the forecast and the verification value, respectively. w_i represents the weighting coefficient, n is the number of samples, and ϕ_i is latitude. In general, observational values, initial values or objective analysis values are often used as verification values. When the forecast is fully correct, called *a perfect forecast*, ME is equal to zero.

Calculation of an average over an extensive region such as the Northern Hemisphere requires evaluation with weighting coefficients in consideration of latitude-related differences among areas. By way of example, to evaluate objective analysis in an equirectangular projection, the weighting coefficient "$w_i = 1/n$" is often replaced with the cosine of latitude "$\cos\phi_i$". The other indices in Section A.1 are handled in the same manner.

A.1.2 Root Mean Square Error

Root Mean Square Error (RMSE) is often used to represent forecast accuracy, and is defined by

$$\mathrm{RMSE} \equiv \sqrt{\sum_{i=1}^{n} w_i D_i^2} \Big/ \sqrt{\sum_{i=1}^{n} w_i}, \tag{A.1.2}$$

where D_i represents deviation between the forecast and the verification values in Eq. (A.1.1b), w_i represents the weighting coefficient in Eq. (A.1.1c), and n is the number of samples. Proximity of the RMSE to zero indicates that forecast values are closer to verification values. For a perfect forecast, RMSE is equal to zero. With the components of ME and random error separated, RMSE is expressed as follows:

$$\mathrm{RMSE}^2 = \mathrm{ME}^2 + \sigma_e^2, \tag{A.1.3}$$

where σ_e represents standard deviation (SD) for the deviation D_i, and is given by

$$\sigma_e^2 = \left(\sum_{i=1}^{n} w_i (D_i - \mathrm{ME})^2 \right) \Big/ \sum_{i=1}^{n} w_i. \tag{A.1.4}$$

A.1.3 Anomaly Correlation Coefficient

The anomaly correlation coefficient (ACC) is one of the most widely used measures in the verification of spatial fields (Jolliffe and Stephenson 2003), and represents the correlation between anomalies of forecasts and those of verification values with reference values such as climatological data. ACC is defined as follows:

$$\mathrm{ACC} \equiv \frac{\sum_{i=1}^{n} w_i \left(f_i - \overline{f}\right)(a_i - \overline{a})}{\sqrt{\sum_{i=1}^{n} w_i \left(f_i - \overline{f}\right)^2 \sum_{i=1}^{n} w_i \left(a_i - \overline{a}\right)^2}}, \quad (-1 \le \mathrm{ACC} \le 1), \tag{A.1.5}$$

where n is the number of samples, and f_i, $\overline{f}$, a_i and $\overline{a}$ are given by the following equations:

$$f_i = F_i - C_i, \quad \overline{f} = \left(\sum_{i=1}^{n} w_i f_i \right) \Big/ \sum_{i=1}^{n} w_i, \tag{A.1.6a}$$

$$a_i = A_i - C_i, \quad \overline{a} = \left(\sum_{i=1}^{n} w_i a_i \right) \Big/ \sum_{i=1}^{n} w_i, \tag{A.1.6b}$$

where F_i, A_i, and C_i represent the forecast value, the verification value and a reference such as a climatological value, respectively. $\overline{f}$ is the mean of f_i, $\overline{a}$ is the mean of a_i, and w_i represents the weighting coefficient in Eq. (A.1.1c). If the variation pattern of forecast anomalies is perfectly coincident with that of verification anomalies, the ACC will take the maximum value of 1. Conversely, if the variation pattern is completely reversed, it will take the minimum value of -1.

A.1.4 Ensemble Spread

Ensemble Spread is a familiar measure representing the degree of forecast uncertainty in the ensemble forecast. It is the standard deviation of the ensembles as defined by

$$\text{Ensemble Spread} \equiv \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{M} \sum_{m=1}^{M} (F_{m,i} - \bar{F}_i)^2 \right)}, \tag{A.1.7}$$

where M is the number of ensemble members, N is the number of samples, $F_{m,i}$ represents the forecast of the mth member, and $\bar{F}_i$ is the ensemble mean, defined by

$$\bar{F}_i \equiv \frac{1}{M} \sum_{m=1}^{M} F_{m,i}. \tag{A.1.8}$$

Table A.2.1: Schematic contingency table for categorical forecasts of a binary event. The numbers of outcomes in each category are indicated by FO, FX, XO and XX, and N is the total number of events.

	Observed	Not Observed	Total
Forecasted	FO (hits)	FX (false alarms)	$FO + FX$
Not Forecasted	XO (misses)	XX (correct rejections)	$XO + XX$
Total	M	X	N

A.1.5 S1 Score

The S1 Score is often used to measure the degree of error in the depiction of forecast pressure fields, and is defined by

$$\mathrm{S1} \equiv 100 \times \frac{\sum_{i=1}^{n} w_i \left\{ \left|\partial_x D_i\right| + \left|\partial_y D_i\right| \right\}}{\sum_{i=1}^{n} w_i \left[\max\left(\left|\partial_x F_i\right|, \left|\partial_x A_i\right|\right) + \max\left(\left|\partial_y F_i\right|, \left|\partial_y A_i\right|\right) \right]}, \tag{A.1.9}$$

where F_i and A_i represent forecast and verification values, respectively. D_i is the deviation between the forecast and verification values in Eq. (A.1.1b), w_i is the weighting coefficient in Eq. (A.1.1c), and the subscripts x and y denote the differential with respect to x and y, as expressed by

$$\partial_x X = \frac{\partial X}{\partial x}, \quad \partial_y X = \frac{\partial X}{\partial y}. \tag{A.1.10}$$

Lower S1 Scores indicate superior forecasts.

A.2 Verification Indices for Categorical Forecasts

Many meteorological phenomena can be regarded as simple binary events, and related forecasts or warnings are often issued as unqualified statements indicating that such events will or will not occur (Jolliffe and Stephenson 2003). In the verification of forecasts for binary events, outcomes for the targeted phenomenon are distinguished in terms of correspondence between forecasts and observations using a 2×2 contingency table as shown in Table A.2.1.

A.2.1 Contingency Table

In the contingency table, categorical forecasts for a binary event are divided into hits, false alarms, misses and correct rejections (or correct negatives) with numbers expressed as FO, FX, XO and XX, respectively. The total number of events is the sum of numbers for all outcomes, given by $N = FO + FX + XO + XX$. The numbers of observed events and non-observed events are $M = FO + XO$ and $X = FX + XX$, respectively.

A.2.2 Proportion Correct

Proportion Correct (PC) is the ratio of the number of correct events $FO + XX$ to the total number of events N, and is defined by

$$\text{PC} \equiv \frac{FO + XX}{N}, \qquad (0 \leq \text{PC} \leq 1). \tag{A.2.1}$$

Higher PC values indicate higher forecast accuracy.

A.2.3 False Alarm Ratio

The false alarm ratio (FAR) is the ratio of the number of false alarm events FX to the number of forecast events $FO + FX$, and is defined by

$$\text{FAR} \equiv \frac{FX}{FO + FX}, \qquad (0 \leq \text{FAR} \leq 1). \tag{A.2.2}$$

Lower FAR values indicate a lower number of false alarm events. In some cases, the total number N is used as the denominator in Eq. (A.2.2) instead of $FO + FX$.

A.2.4 Undetected Error Rate

The undetected error rate (Ur) is the ratio of the number of miss events XO to the number of observed events M, and is defined by

$$\text{Ur} \equiv \frac{XO}{M}, \qquad (0 \leq \text{Ur} \leq 1). \tag{A.2.3}$$

Lower Ur values indicate a lower number of miss events. In some cases, the total number N is used as the denominator in Eq. (A.2.3) instead of M.

A.2.5 Hit Rate

The hit rate (Hr) is the ratio of the number of hit events FO to the number of observed events M, and is defined by

$$\text{Hr} \equiv \frac{FO}{M}, \qquad (0 \leq \text{Hr} \leq 1). \tag{A.2.4}$$

Higher Hr values indicate a lower number of miss events. The hit rate is used to plot the ROC curve described in Subsection A.3.5.

A.2.6 False Alarm Rate

The false alarm rate (Fr) is the ratio of the number of false alarm events FX to the number of non-observed events X, and is defined by

$$\text{Fr} \equiv \frac{FX}{X}, \qquad (0 \leq \text{Fr} \leq 1). \tag{A.2.5}$$

Lower Fr values indicate a lower number of false alarm events and higher forecast accuracy. The denominator of the false alarm rate differs from that of the false alarm ratio (see Subsection A.2.3). The false alarm rate is also used to plot the ROC curve described in Subsection A.3.5.

A.2.7 Bias Score

The bias score (BI) is the ratio of the number of forecasted events $FO + FX$ to the number of observed events M, and is defined by

$$\mathrm{BI} \equiv \frac{FO + FX}{M}, \qquad (0 \leq \mathrm{BI}). \tag{A.2.6}$$

If the number of forecasted events $FO + FX$ is equal to the number of observed events M, BI will be unity. If BI is larger than unity, the frequency of events is overestimated. Conversely, if BI is smaller than unity, the frequency of events is underestimated.

A.2.8 Climatological Relative Frequency

Climatological relative frequency (P_c) is the probability of occurrence of events estimated from samples, and is defined by

$$\mathrm{P}_c \equiv \frac{M}{N}, \tag{A.2.7}$$

where M is the number of observed events occurring, and N is the total number of events. P_c is derived from the number of observed events, and is independent of forecast accuracy.

A.2.9 Threat Score

The threat score (TS) is an index value focused on hit events. It represents the ratio of the number of hit events FO to the number of events other than correct rejections $FO + FX + XO$, and is defined by

$$\mathrm{TS} \equiv \frac{FO}{FO + FX + XO}, \qquad (0 \leq \mathrm{TS} \leq 1). \tag{A.2.8}$$

If the number of observed events is extremely small (i.e. $N \gg M$, and $XX \gg FO$, FX, or XO), the proportion correct (PC) value will be close to unity due to the the major contribution from the number of non-observed events. The TS is applicable to validation of forecasts accuracy without contribution from correct rejection events. Forecast accuracy rises as the TS value approaches the maximum value of unity. As TS values are often affected by climatological relative frequency, they are not applicable to comparison regarding the accuracy of forecasts validated under different conditions. To address this issue, equitable threat scores are often used for validation.

A.2.10 Equitable Threat Score

The equitable threat score (ETS) is similar to the threat score, but with the removal of contribution from hits by chance in *random forecasts*, and is defined by (Schaefer 1990)

$$\mathrm{ETS} \equiv \frac{FO - S_f}{FO + FX + XO - S_f}, \qquad (-\frac{1}{3} \leq \mathrm{ETS} \leq 1), \tag{A.2.9}$$

and

$$S_f = P_c(FO + FX), \quad P_c = \frac{M}{N}, \tag{A.2.10}$$

where P_c is the climatological relative frequency and S_f is the number of hit events being forecast randomly $FO + FX$ times. Proximity to the maximum value of unity indicates higher forecast accuracy. For random forecasts, the ETS is zero. This metric has a minimum value of $-1/3$ if $FO = XX = 0$ and $FX = XO = N/2$.

A.2.11 Heidke Skill Score

The Heidke skill score (HSS) is used to remove the effects of issues in individual forecasts in consideration of the number of correct events in a random forecast estimated from climatological probability, and is defined by

$$\mathrm{HSS} \equiv \frac{FO + XX - S}{N - S}, \qquad (-1 \leq \mathrm{HSS} \leq 1), \tag{A.2.11}$$

where

$$S = P_c(FO + FX) + Px_c(XO + XX), \tag{A.2.12}$$

and

$$P_c = \frac{M}{N}, \quad Px_c = \frac{X}{N} = 1 - P_c, \tag{A.2.13}$$

where P_c and Px_c are the climatological relative frequencies of observed and non-observed events in random forecasting, respectively. Proximity to the maximum value of unity indicates higher forecast accuracy. The Heidke skill score is zero in random forecasts and unity in perfect forecasts. The index has a minimum value of -1 if $FO = XX = 0$ and $FX = XO = N/2$.

A.2.12 Fractions Skill Score

The fractions skill score (FSS) is an index of how forecast skill varies with spatial scale. In other words, it is a measure to verify forecasted fractional event frequencies. The verification method (Roberts and Lean 2008) is described here.

First, all model and observation data are projected onto the same verification grid. Suitable thresholds (q) are chosen and used to convert the observed (O) and forecast (F) fields into binary fields I_O and I_F. All grid squares exceeding the threshold have a value of 1 and all others a value of 0,

$$I_o = \begin{cases} 1 & (O \geq q) \\ 0 & (O < q) \end{cases} \quad and \quad I_F = \begin{cases} 1 & (F \geq q) \\ 0 & (F < q) \end{cases} \tag{A.2.14}$$

Second, for every grid point in the binary fields obtained from Eq. (A.2.14), computation is performed to determine the fraction of surrounding points within a given square of length n that have a value of 1. These are described by

$$O(n)_{i,j} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_o \left[i + k - 1 - \frac{(n-1)}{2}, j + l - 1 - \frac{(n-1)}{2} \right], \tag{A.2.15}$$

$$F(n)_{i,j} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_F \left[i + k - 1 - \frac{(n-1)}{2}, j + l - 1 - \frac{(n-1)}{2} \right]. \tag{A.2.16}$$

Third, the mean square error (MSE) for the observed and forecast fractions from the neighborhood of length n is computed using

$$MSE(n) = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [O(n)_{i,j} - F(n)_{i,j}]^2. \tag{A.2.17}$$

Here i goes from 1 to N_x, where N_x is the number of columns in the domain, and j goes from 1 to N_y, where N_y is the number of rows. $O(n)_{i,j}$ is the resultant field of observed fractions for the square of length n and $F(n)_{i,j}$ is the resultant field of model forecast fractions. However, the MSE is not in itself very useful because it is highly dependent on the frequency of the event itself. The fractions skill score is defined by

$$FSS(n) = \frac{MSE(n) - MSE(n)_{ref}}{MSE(n)_{perfect} - MSE(n)_{ref}} = 1 - \frac{MSE(n)}{MSE(n)_{ref}} \tag{A.2.18}$$

where $MSE(n)_{perfect}$ is the MSE of a perfect forecast for neighborhood length n and $MSE(n)_{ref}$ is the MSE of the reference.

A.3 Verification Indices for Probability Forecasts

A.3.1 Brier Score

The Brier score (BS) is a basic verification index for probability forecasts, and is defined by

$$\mathrm{BS} \equiv \frac{1}{N}\sum_{i=1}^{N}(p_i - a_i)^2, \quad (0 \le \mathrm{BS} \le 1), \tag{A.3.1}$$

where p_i is the forecast probability of occurrence of an event ranging from 0 to 1 in probability forecasts, a_i indicates observation with binary values (1 for observed and 0 for not observed), and N is the number of samples. Smaller BS values indicate higher forecast accuracy. In a perfect forecast, the BS has a minimum value of 0.

The Brier score for *climatological forecasts* (BS_c), in which the climatological relative frequency $P_c = M/N$ is always used as the forecast probability p_i ,is defined by

$$\mathrm{BS}_c \equiv P_c(1 - P_c), \tag{A.3.2}$$

Since the Brier score is influenced by the climatological frequency of events in the verification sample, it is not applicable to comparison of accuracy for forecasts with different sets of samples and/or different phenomena. For example, BS_c may differ with differing values of P_c even under the same forecast method (e.g., the climatological approach) because of its dependence on P_c (Stanski and Burrows 1989). To reduce this effect, the Brier skill score is often used for verification with improvement from the climatological forecast (see Subsection A.3.2).

A.3.2 Brier Skill Score

The Brier skill score (BSS) is an index based on the Brier score. It indicates the degree of forecast improvement in reference to climatological forecasts, and is defined by

$$\mathrm{BSS} \equiv \frac{\mathrm{BS}_c - \mathrm{BS}}{\mathrm{BS}_c}, \quad (\mathrm{BSS} \le 1), \tag{A.3.3}$$

where BS is the Brier score and BS_c is the Brier score for the climatological forecast. BSS is unity for a perfect forecast and zero for the climatological forecast. Its value is negative if the forecast error exceeds that of the climatological forecast.

A.3.3 Murphy's Decompositions

To provide deeper insight into the relationship between the Brier score (BS) and the properties of probability forecasts, Murphy (1973) decomposed the score into reliability, resolution and uncertainty terms (Eq. A.3.4a), refered to as Murphy's Decompositions.

Consider the probability of forecasts classified to L intervals. Let the sample number in the lth interval be N_l, and let the number of observed events in N_l be M_l. It follows that $N = \sum_{l=1}^{L} N_l$ and $M = \sum_{l=1}^{L} M_l$. The BS value can therefore be represented with Murphy's decompositions as follows:

$$\mathrm{BS} = \mathrm{Reliability} - \mathrm{Resolution} + \mathrm{Uncertainty}, \tag{A.3.4a}$$

$$\mathrm{Reliability} = \sum_{l=1}^{L}\left(p_l - \frac{M_l}{N_l}\right)^2 \frac{N_l}{N}, \tag{A.3.4b}$$

$$\mathrm{Resolution} = \sum_{l=1}^{L}\left(\frac{M}{N} - \frac{M_l}{N_l}\right)^2 \frac{N_l}{N}, \tag{A.3.4c}$$

$$\mathrm{Uncertainty} = \frac{M}{N}\left(1 - \frac{M}{N}\right), \tag{A.3.4d}$$

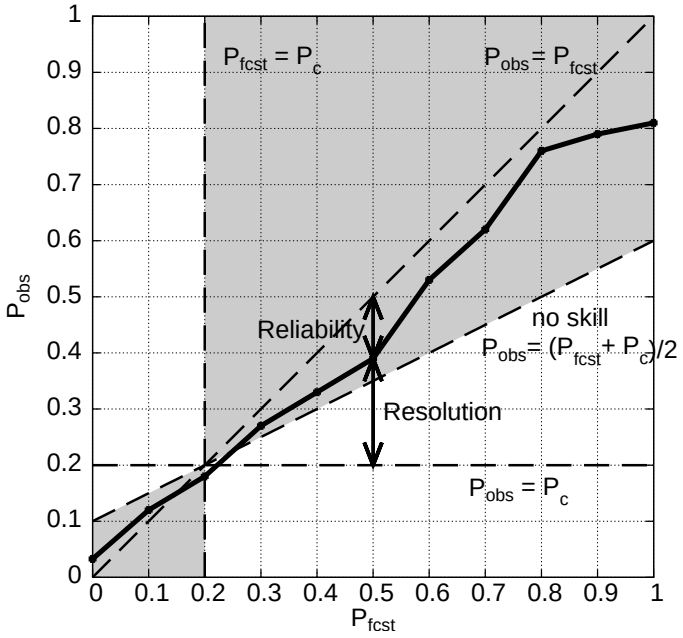


Figure A.3.1: Reliability diagram. The ordinate represents the relative frequencies of observed events P_{obs}, the abscissa is the probability of forecast event occurrence P_{fcst}, and the solid line is the reliability curve. Grey shading indicates positive contribution to the BSS.

where p_l is the representative value in the lth interval of predicted probability. Reliability is the minimum value of zero when p_l is equal to the relative frequency of the observed events M_l/N_l. If the distance between $M/N\,(=P_c)$ and M_l/N_l is longer, resolution will have a large value. Uncertainty depends on observed events regardless of forecast methods. When $P_c = 0.5$, Uncertainty will have the maximum value of 0.25. Uncertainty is equal to the Brier score for climatological forecasts (BS_c). In this regard, the Brier skill score (BSS) can be expressed as

$$\text{BSS} = \frac{\text{Resolution} - \text{Reliability}}{\text{Uncertainty}}. \tag{A.3.5}$$

A.3.4 Reliability Diagram

Probability forecast performance is often evaluated using a reliability diagram, also called an attributes diagram. This is a chart detailing the relative frequencies of observed events P_{obs} as the ordinate and the probability of forecast event occurrence P_{fcst} as abscissa as shown in Figure A.3.1. The plots are generally displayed in the form of a reliability curve.

The properties of the curve can be related to the reliability and resolution terms of Murphy's decompositions. Contribution to reliability (or resolution) for each value of P_{fcst} is associated with the squared distance from a point on the reliability curve to the line $P_{\text{obs}} = P_{\text{fcst}}$ (or $P_{\text{obs}} = P_c$), and is derived from its weighted mean using the number of samples as weights. The contributions are the same for both reliability and resolution on the line $P_{\text{obs}} = (P_{\text{fcst}} + P_c)/2$, called the no-skill line, and contribution to the Brier score is zero on this line. The shading enclosed by the no-skill line, the line $P_{\text{fcst}} = P_c$ and the axes in Figure A.3.1 indicate the area of positive contribution to the BSS, since the contribution to reliability is larger than that to resolution. For further details of reliability diagrams, see Wilks (2006).

In climatological forecasting (see Subsection A.3.1) as a special case, the reliability curve corresponds to a point $(P_{\text{fcst}}, P_{\text{obs}}) = (P_c, P_c)$. Probability forecasts with the following properties will have higher accuracy.

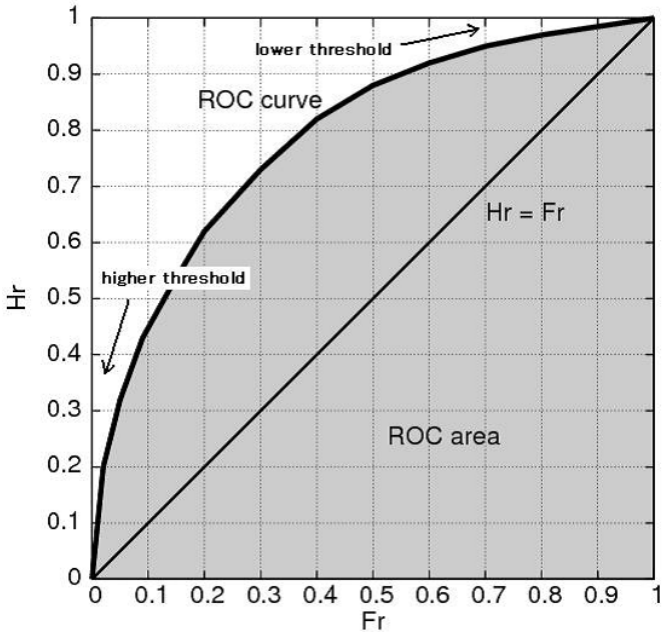


Figure A.3.2: Schematic diagram of an ROC curve. The ordinate is Hr and the abscissa is Fr. Gray shading indicates the ROC area.

- The reliability curve is close to the line $P_{\text{obs}} = P_{\text{fcst}}$ (reliability close to zero).
- Points with a large number of samples on the reliability curve are distributed away from the point of the climatological forecast $(P_{\text{fcst}}, P_{\text{obs}}) = (P_c, P_c)$ (around the lower left or the upper right of the reliability diagram) with higher resolution.

A.3.5 ROC Area Skill Score

If two alternatives in a decision problem, whether the event occurs or not, must be chosen on the basis of a probability forecast for a dichotomous variable, the choice will depend on the probability threshold. A relative operating characteristic (ROC) curve is often used to evaluate such decision problems. This involves the use of a schematic diagram whose ordinate and abscissa represent the hit rate (Hr) and the false alarm rate (Fr), respectively, and is made from contingency tables with variations of threshold values as shown in Figure A.3.2.

The threshold value is lower around the upper right of the diagram and higher around the lower left. Probability forecasting is more accurate when the curve is more convex to the top because the hit rate is higher than the false alarm rate; that is, $Hr > Fr$ around the upper left. The gray shaded area below the ROC curve, called the ROC area (ROCA), will be larger with higher values of information in probability forecasts. For further details of ROC curves, see Wilks (2006).

The ROC area skill score (ROCASS) is a validation index in reference to probability forecasts with no information values (i.e. $Hr = Fr$), and is defined by

$$\text{ROCASS} \equiv 2(\text{ROCA} - 0.5), \quad (-1 \leq \text{ROCASS} \leq 1). \tag{A.3.6}$$

ROCASS is unity for a perfect forecast and zero for a forecast with no information values, such as one with a uniform probability as randomly sampled from the range [1, 0].

References

Adcroft, A. and J.-M. Campin, 2004: Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modell.*, **7**, 269–284.

Agresti, A., 2002: *Categorical Data Analysis, 2nd ed.* New York: Wiley, 734pp.

An, H. S., 1977: A Numerical Experiment of the M2 Tide in the Yellow Sea. *J. Oceanogr.*, **33**, 103–110.

An, Soon-Il, 2008: Interannual variations of the tropical ocean instability wave and ENSO. *J. Climate*, **21**, 3680–3686.

Anderson, J. L., 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.

Andersson, E. and H. Järvinen, 1999: Variational Quality Control. *Quart. J. Roy. Meteor. Soc.*, **125**, 697–722.

Aoki, Te., Ta. Aoki, M. Fukabori, and T. Takao, 2002: Characteristics of UV-B Irradiance at Syowa Station, Antarctica: Analyses of the Measurements and Comparison with Numerical Simulations. *J. Meteor. Soc. Japan*, **80**, 161–170.

Arakawa, A., 1972: Design of the UCLA general circulation model. Numerical simulation weather and climate, Tech. Rep. 7, Dept. of Meteorology, University of California, Los Angeles. 116pp.

Arakawa, A. and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics*, **17**, 174–265, Academic Press.

Arakawa, A. and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.*, **31**, 674–701.

Aranami, K. and J. Ishida, 2004: Implementation of two dimensional decomposition for JMA non-hydrostatic model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **34**, 03.01–03.02.

Aranami, K., T. Hara, Y. Ikuta, K. Kawano, K. Matsubayashi, H. Kusabiraki, T. Ito, T. Egawa, K. Yamashita, Y. Ota, Y. Ishikawa, T. Fujita, and J. Ishida, 2015: A new operational regional model for convection-permitting numerical weather prediction at JMA. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **45**, 05.05–05.06.

Baines, P. G. and T. N. Palmer, 1990: Rationale for a new physically-based parametrization of subgrid-scale orographic effects. 11.

Barnes, R. T. H., R. Hide, A. A. White, and C. A. Wilson, 1983: Atmospheric angular momentum fluctuations, length-of-day change and polar motion. *Proc. R. Soc.*, **A387**, 31–73.

Bartholomé, E. and A. S. Belward, 2005: GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, **26**, 1959–1977.

Bechtold, P., M. Kohler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **134**, 1337–1351.

Beljaars, A. C. M., 1995: The parameterization of surface fluxes in large-scale models under free convection. *Quart. J. Roy. Meteor. Soc.*, **121**, 255–270.

Beljaars, A. C. M., A. R. Brown, and N. Wood, 2004a: A new parametrization of turbulent orographic form drag. *Quart. J. Roy. Meteor. Soc.*, **130**, 1327–1347.

Beljaars, A. C. M., A. R. Brown, and N. Wood, 2004b: A new parametrization of turbulent orographic form drag. **130**, 1327–1347.

Beljaars, A. C. M. and A. A. M. Holtslag, 1991: Flux parameterization over land surfaces for atmospheric models. *J. Appl. Meteor.*, **30**, 327–341.

Berger, M. J. and J. Oliger, 1984: Adaptive mesh refinement for hyperbolic partial differential equations. *J. Comp. Phys.*, **53**, 484–512.

Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A Spectral Stochastic Kinetic Energy Backscatter Scheme and Its Impact on Flow-Dependent Predictability in the ECMWF Ensemble Prediction System. *J. Atmos. Sci.*, **66**, 603–626.

Best, M. J., A. C. M. Beljaars, J. Polcher, and P. Viterbo, 2004: A proposed structure for coupling tiled surfaces with the planetary boundary layer. *J. Hydr. Meteorol.*, **5**, 1271–1278.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform

Kalman Filter. Part I : Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Blackadar, A. K., 1962: The vertical distribution of wind and turbulent exchange in a neutral atmosphere. *J. Geophys. Res.*, **67**, 3095–3102.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.

Bourke, W., 1974: A multi-level spectral model. I. Formulation and hemispheric integrations. *Mon. Wea. Rev.*, **102**, 687–701.

Brasnett, B., 1999: A global analysis of snow depth for numerical weather prediction. *J. Appl. Meteor.*, **38**, 726–740.

Braud, I., J. Noilhan, P. Bessemoulin, P. Mascart, R. Haverkamp, and M. Vauclin, 1993: Bare-ground surface heat and water exchanges under dry conditions: Observations and parameterization. *Bound.-Layer Meteor.*, **66**, 173–200.

Briegleb, B. P., P. Minnis, V. Ramanathan, and E. Harrison, 1986: Comparison of Regional Clear-Sky Albedos Inferred from Satellite Observations and Model Computations. *Journal of Climate and Applied Meteorology*, **25**, 214–226.

Bringi, V. N. and V. Chandrasekar, 2001: *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press.

Brown, P. R. and P. N. Francis, 1995: Improved Measurements of the Ice Water Content in Cirrus Using a Total-Water Probe. *J. Atmos. and Oceanic Technol.*, **12**, 410–414.

Broxton, P. D., X. Zeng, W. Scheftic, and P. A. Troch, 2014a: A MODIS-Based 1 km Maximum Green Vegetation Fraction Dataset. *J. Appl. Meteor. Climat.*, **53**, 1996–2004.

Broxton, P. D., X. Zeng, W. Scheftic, and P. A. Troch, 2014b: A MODIS-Based 1 km Maximum Green Vegetation Fraction Dataset. *J. Appl. Meteor. Climat.*, **53**, 1996–2004.

Buizza, R., 1998: Impact of Horizontal Diffusion on T21, T42, and T63 Singular Vectors. *J. Atmos. Sci.*, **55**, 1069–1083.

Buizza, R., M. Milleer, and T. N. Palmer, 1999a: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.

Buizza, R., M. Miller, and T. N. Palmer, 1999b: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.

Buizza, R. and T. N. Palmer, 1995: The Singular-Vector Structure of the Atmospheric Global Circulation. *J. Atmos. Sci.*, **52**, 1434–1456.

Cameron, J. and W. Bell, 2018: The testing and implementation of variational bias correction (VarBC) in the Met Office global NWP system. *Weather Science Technical Report No: 631,Met Office*, 1–22.

Caron, J. F., 2013: Mismatching perturbations at the lateral boundaries in limited-area ensemble forecasting: A case study. *Mon. Wea. Rev.*, **141**, 356–374.

Carter, W.P.L., 2000: Documentation of the SAPRC-99 Chemical Mechanism for VOC reactivity assesment. Final report to california air resources board.

Cats, G., Van T. Vu, and L. Wolters, 2008: Overlapping communications with calculations. *Hirlam newsletter*, **54**, 189–201.

Caumont, O., V. Ducrocq, É. Wattrelot, G. Jaubert, and S. Pradier-vabre, 2010: 1D+3DVar assimilation of radar reflectivity data: a proof of concept. *Tellus*, **62A**, 173–187.

Cavaleri, L. and P. M. Rizzoli, 1981: Wind wave prediction in shallow water: Theory and applications. *J. Geophys. Res.*, **86**, 10 961–10 973.

Charnock, H., 1955: Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc.*, **81**, 639–640.

Chen, C., R. C. Beardsley, and G. Cowles, 2003: An unstructured grid, finite-volume community ocean model: FVCOM User Manual. SMAST/UMASSD Technical Report-13-0701, pp404.

Chiba, J. and H. Kawai, 2021: Improved SST-shortwave radiation feedback using an updated stratocumulus parameterization. *WGNE blue book, Res. Activ. Earth Sys. Modell*, **51**, 4–03.

Chikamoto, Y., H. Mukougawa, T. Kubota, H. Sato, A. Ito, and S. Maeda, 2007: Evidence of growing bred vector associated with the tropical intraseasonal oscillation. *Geophys. Res. Lett.*, **34**, L04 806.

Chou, M.-D., M. J. Suarez, X.-Z. Liang, and M. M.-H. Yan, 2001: A thermal infrared radiation parameterization for atmospheric studies. *Technical report series on global modeling and data assimilation, Vol. 19, NASA Goddard Space Flight Center*, 56pp.

Coakley, J. A., Jr., R. D. Cess, and F. B. Yurevich, 1983: The effect of tropospheric aerosols on the Earth's radiation budget: A parameterization for climate models. *J. Atmos. Sci.*, **40**, 116–138.

Collins, W. D., 2001: Parameterization of generalized cloud overlap for radiative calculations in general circulation models. *J. Atmos. Sci.*, **58**, 3224–3242.

Collins, W. D., J. M. Lee-Taylor, D. P. Edwards, and G. L. Francis, 2006: Effects of increased near-infrared absorption by water vapor on the climate system. *J. Geophys. Res.*, **111**, D18 109.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387.

Culverwell, I. D., H. W. Lewis, D. Offiler, C. Marquardt, and C. P. Burrows, 2015: The Radio Occultation Processing Package, ROPP. *Atmos. Meas. Tech.*, **8**, 1887–1899.

Danilov, S., D. Sidorenko, Q. Wang, and T. Jung, 2017: The finite-volume sea ice-ocean model (FESOM2). *Geosci. Model Dev.*, **10**, 765–789.

de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, 2013: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **139**, 1199–1213.

Dee, D. P., 2004: Variational bias correction of radiance data in the ECMWF system. *Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP*, Reading, UK, 28 June - 1 July 2004, 97–112.

DeFries, R. S., M. C. Hansen, J. R. G. Townshend, A. C. Janetos, and T. R. Loveland, 2000: A new global 1km data set of percent tree cover derived from remote sensing. *Global Chg. Biol.*, **6**, 247–254.

Derber, J. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221.

Derber, J. C. and W.-S Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131**, 3385–3396.

Deushi, M. and K. Shibata, 2011: Development of a Meteorological Research Institute chemistry-climate model version 2 for the study of tropospheric and stratospheric chemistry. *Pap. Meteor. Geophys*, **62**, 1–46.

Dobbie, J. S., J. Li, and P. Chýlek, 1999: Two- and four-stream optical properties for water clouds and solar wavelengths. *J. Geophys. Res.*, **104**, 2067–2079.

Drusch, M. and P. Viterbo, 2007: Assimilation of screen-level variables in ECMWF's integrated forecast system: A study on the impact on the forecast quality and analyzed soil moisture. *Mon. Wea. Rev.*, **135**, 300–314.

Ebert, E. E. and J. A. Curry, 1992: A parameterization of ice cloud optical properties for climate models. *J. Geophys. Res.*, **97**, 3831–3836.

ECMWF, 2014: Part IV: Physical Process, Chapter 7 Clouds and large-scale precipitation. *IFS Documentation—Cy40r1*, 91–108.

Egbert, G. D. and S. Y. Erofeeva, 2002: Efficient Inverse Modeling of Barotropic Ocean Tides. *J. Atmos. Oceanic Technol.*, **19**, 183–204.

Ehrendorfer, M., R. M. Errico, and K. D. Raeder, 1999: Singular vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.*, **56**, 1627–1648.

Elliott, A. J., 1986: Shear diffusion and the spread of oil in the surface layers of the North Sea. *Dt. Hydrogr. Zeit.*, **39**, 113–137.

Engwirda, D., 2017: JIGSAW-GEO (1.0): locally orthogonal staggered unstructured grid generation for general circulation modelling on the sphere. *Geosci. Model Dev.*, **10**, 2117–2140.

Epifanio, C. C. and T. Qian, 2008: Wave–turbulence interactions in a breaking mountain wave. *J. Atmos. Sci.*, **65**, 3139–3158.

Essery, R., J. Pomeroy, J. Parviainen, and P. Storck, 2003: Sublimation of Snow from Coniferous Forests in a

Climate Model. *J. Climate*, **16**, 1855–1864.

Eyre, J. R. and W. P. Menzel, 1989: Retrieval of Cloud Parameters from Satellite Sounder Data: A Simulation Study. *J. Appl. Meteor.*, **28**, 267–275.

FAO, IIASA, ISRIC, ISSCAS, and JRC, 2012: Harmonized World Soil Database (version 1.2). 42pp., URL http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HWSD_Documentation.pdf.

Fećan, F., B. Marticorena, and G. Bergametti, 1998: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas. *Ann. Geophy.*, **17**, 149–157.

Field, P. R., A. J. Heymsfield, and A. Bansemer, 2007: Snow Size Distribution Parameterization for Midlatitude and Tropical Ice Clouds. *J. Atmos. Sci.*, **64**, 4346–4365.

Fingas, M. F., 2010: OIL SPILL SCIENCE and TECHNOLOGY. *Gulf Professional Publishing*, 201–242.

Folland, C. K. and D. E. Parker, 1995: Correction of instrumental biases in historical sea surface temperature data. *Quart. J. Roy. Meteor. Soc.*, **121**, 319–367.

Frank, W. M., 1977: The structure and energetics of the tropical cyclone I. Storm structure. *Mon. Wea. Rev.*, **105**, 1119–1135.

Frank, W. M., 1984: A composite analysis of the core of a mature hurricane. *Mon. Wea. Rev.*, **112**, 2401–2420.

Freidenreich, S. M. and V. Ramaswamy, 1999: A new multiple-band solar radiative parameterization for general circulation models. *J. Geophys. Res.*, **104**, 31 389–31 409.

Fu, Q. and K. N. Liou, 1992: On the correlated *k*-distribution method for radiative transfer in nonhomogeneous atmospheres. *J. Atmos. Sci.*, **49**, 2139–2156.

Fujii, Y., 2005: Preconditioned Optimizing Utility for Large-dimensional analyses (POpULar). *J. Oceanogr.*, **61**, 655–662.

Fujii, Y. and M. Kamachi, 2003a: A nonlinear preconditioned quasi-Newton method without inversion of a first-guess covariance matrix in variational analyses. *Tellus A*, **55**, 450–454.

Fujii, Y. and M. Kamachi, 2003b: A reconstruction of observed profiles in the sea east of Japan using vertical coupled temperature-salinity EOF modes. *J. Oceanogr.*, **59**, 173–186.

Fujii, Y., M. Kamachi, S. Matsumoto, and S. Ishizaki, 2012: Barrier layer and relevant variability of the salinity field in the equatorial Pacific estimated in an ocean reanalysis experiment. *Pure Appl. Geophys.*, **169**, 579–594.

Fujii, Y., T. Yoshida, H. Sugimoto, I. Ishikawa, and S. Urakawa, 2023: Evaluation of a global ocean reanalysis generated by a global ocean data assimilation system based on a Four-Dimensional Variational (4DVAR) method. *Front. Clim.*, **4**, 1–20, doi:10.3389/fclim.2022.1019673.

Fujita, T., 1952: Pressure Distribution within Typhoon. *Geophys. Mag.*, **23**, 437–451.

Garratt, J. R. and R. J. Francey, 1978: Bulk characteristics of heat transfer in the unstable, baroclinic atmospheric boundary layer. *Bound.-Layer Meteor.*, **15**, 399–421.

Gauthier, P. and J.-N. Thépaut, 2001: Impact of digital filter as a weak constraint in the preoparational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.*, **129**, 2089–2102.

Geer, A. J., F. Baordo, N. Bormann, and S. J. English, 2014: All-sky assimilation of microwave humidity sounders. *ECMWF Tech. Memo.*, **741**.

Geer, A. J. and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037.

Geleyn, J.-F. and A. Hollingsworth, 1979: An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. *Beitr. Phys. Atmos.*, **52**, 1–16.

Gent, P. R. and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155.

Gifford, F. A., 1982: Horizontal diffusion in the atmosphere: a Lagrangian-dynamical theory. *Atmos. Env.*, **16**, 505–512.

Giglio, L., J. T. Randerson, van der G. R. Werf, P. S. Kasibhatla, G. J. Collatz, D. C. Morton, and R. S. DeFries, 2010: Assessing variability and long-term trends in burned area by merging multiple satellite fire products. *Biogeosciences*, **7**, 1171–1186.

Godfrey, J. S. and A. C. M. Beljaars, 1991: On the turbulent fluxes of buoyancy, heat and moisture at the air-sea interface at low wind speeds. *J. Geophys. Res.-Oceans*, **96**, 22 043–22 048.

Godske, C. L. et al., 1957: *Dynamic Meteorology and Weather Forecasting*, chap. 18. Amer. Met. Soc.

Gonzalez, T. and A. Taylor, 2018: Development of the NWS' Probabilistic Tropical Storm Surge Model. *Proceedings of the 33rd Conference on Hurricanes and Tropical Meteorology*, **11**.

Griffies, S. M. and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Mon. Wea. Rev.*, **128**, 2935–2946.

Gryanik, V. M., C. Lüpkes, A. Grachev, and D. Sidorenko, 2020: New Modified and Extended Stability Functions for the Stable Boundary Layer based on SHEBA and Parametrizations of Bulk Transfer Coefficients for Climate Models. *J. Atmos. Sci.*, **77**, 2687–2716.

Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P. I. Palmer, and C. Geron, 2006: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys.*, **6**, 3181–3210.

Haimberger, L., C. Tavolato, and S. Sperka, 2008: Toward elimination of the warm bias in historic radiosonde temperature records–Some new results from a comprehensive intercomparison of upper-air data. *J. Climate*, **21**, 4587–4606.

Haimberger, L., C. Tavolato, and S. Sperka, 2012: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *J. Climate*, **25**, 8108–8131.

Hamrud, M., M. Bonavita, and L. Isaksen, 2015: EnKF and hybrid gain ensemble data assimilation part I: EnKF implementation. *Mon. Wea. Rev.*, **143**, 4847–4864.

Han, J. and H.-L. Pan, 2011: Revision of Convection and Vertical Diffusion Schemes in the NCEP Global Forecast System. *Weather and Forecasting*, **26**, 520–533.

Hara, T., 2015: Necessity of parameterizations for convective initiation in high resolution cloud-permitting models. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **45**, 04.06–04.07.

Hara, T., K. Kawano, K. Aranami, Y. Kitamura, M. Sakamoto, H. Kusabiraki, C. Muroi, and J. Ishida, 2012: Development of Physics Library and its application to ASUCA. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **42**, 05.05–05.06.

Hasegawa, H., N. Kohno, M. Higaki, and M. Itoh, 2017: Upgrade of JMA's Storm Surge Prediction for the WMO Storm Surge Watch Scheme (SSWS). *RSMC Tokyo-Typhoon Center Technical Review*, **19**, 26–34.

Hasegawa, H., J. Sugano, T. Fukuura, and M. Higaki, 2023: Upgrade of JMA's Strom Surge Prediction for the WMO Storm Surge Watch Scheme (SSWS) in 2022. *RSMC Tokyo-Typhoon Center Technical Review*, **25**, 1–14.

Hasegawa, Y., A. Sugai, Yo. Hayashi, Yu. Hayashi, S. Saito, and T. Shimbori, 2015: Improvements of volcanic ash fall forecasts issued by the Japan Meteorological Agency. *J. Appl. Volcanol.*, **4**, 2.

Hasselmann, K., T. P. Barnett, E. Bouws, H. Carlson, D. E. Cartwright, K. Enke, J. A. Ewing, H. Gienapp, D. E. Hasselmann, P. Kruseman, A. Meerburg, P. Muller, D. J. Olbers, K. Richter, W. Sell, and H. Walden, 1973: Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Dtsch. Hydrogr. Z.*

Hasselmann, S., K. Hasselmann, J. H. Allender, and T. P. Barnett, 1985: Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part II: Parameterizations of the nonlinear energy transfer for application in wave models. *J. Phys. Oceanogr.*, **15**, 1378–1391.

Hess, R., 2001: Assimilation of screen-level observations by variational soil moisture analysis. *Meteor. Atmos. Phys.*, **77**, 145–154.

Higaki, M., H. Hayashibara, and F. Nozaki, 2009: Outline of the storm surge prediction model at the Japan Meteorological Agency. *RSMC Tokyo-Typhoon Center Technical Review*, **11**, 25–38.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75.

Hirahara, S., Y. Kubo, T. Yoshida, T. Komori, J. Chiba, T. Takakura, T. Kanehama, R. Sekiguchi, K. Ochi, H. Sugimoto, Y. Adachi, I. Ishikawa, and Y. Fujii, 2023: Japan Meteorological Agency/Meteorological Research Institute Coupled Prediction System version 3 (JMA/MRI-CPS3). *J. Meteor. Soc. Japan*, **101**,

149–169.

Hirai, M., K. Miyaoka, H. Sato, H. Sugimoto, A. Minami, and C. Matsukawa, 2014: March 2014 upgrade of JMA's One-month Ensemble Prediction System. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **44**, 6.09–6.10.

Hirose, N., N. Usui, K. Sakamoto, H. Tsujino, G. Yamanaka, H. Nakano, S. Urakawa, T. Toyoda, Y. Fujii, and N. Kohno, 2019: Development of a new operational system for monitoring and forecasting coastal and open-ocean states around Japan. *Ocean Dyn*, **69**, 1333–1357.

Hoffman, R. N. and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus A: Dynamic Meteorology and Oceanography*, **35**, 100–118.

Hogan, R. J. and A. Bozzo, 2015: Mitigating errors in surface temperature forecasts using approximate radiation updates. *J. Adv. Model. Earth Syst*, **7**, 836–853.

Hogan, R. J. and S. Hirahara, 2016: Effect of solar zenith angle specification in models on mean shortwave fluxes and stratospheric temperatures. *Geophys. Res. Lett.*, **43**, 482–488.

Holmlund, K., 1998: The utilization of statistical properties of satellite-derived atmospheric motion vectors to derive quality indicators. *Wea. Forecasting*, **13**, 1093–1104.

Holtslag, A. A. M. and B. A. Boville, 1993: Local versus nonlocal boundary-layer diffusion in a global climate model. *J. Climate*, **6**, 1825–1842.

Honda, Y., M. Nishijima, K. Koizumi, Y. Ohta, K. Tamiya, T. Kawabata, and T. Tsuyuki, 2005: A pre-operational variational data assimilation system for a non-hydrostatic model at the Japan Meteorological Agency: Formulation and preliminary results. *Quart. J. Roy. Meteor. Soc.*, **131**, 3465–3475.

Hortal, M., 2002: The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. *Quart. J. Roy. Meteor. Soc.*, **128**, 1671–1687.

Hoskins, B. J. and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.*, **101**, 637–655.

Hotta, D. and Y. Ota, 2019: Statistical generation of SST perturbations with spatio-temporally coherent growing patterns. *Quart. J. Roy. Meteor. Soc.*, **145**, 1660–1673.

Hu, Y. X. and K. Stamnes, 1993: An accurate parameterization of the radiative properties of water clouds suitable for use in climate models. *J. Climate*, **6**, 728–742.

Hunke, E. C. and W. H. Lipscomb, 2006: CICE. *The Los Alamos Sea Ice Model Documentation and Software User's Manual*, accessed 6 August 2021, https://github.com/CICE-Consortium/CICE/wiki.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

ICAO, 2018: *Annex 3 - Meteorological Service for International Air Navigation*. 20th ed., International Civil Aviation Organization, 224 pp.

Ikawa, M. and K. Saito, 1991: Description of a nonhydrostatic model developed at the Forecast Research Department of the MRI. *Tech. Rep. MRI*, **28**, 238pp.

Ikuta, Y. and Y. Honda, 2011: Development of 1D+4DVAR data assimilation of radar reflectivity in JNoVA. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **41**, 01.09–01.10.

Ikuta, Y., H. Kusabiraki, K. Kawano, T. Anzai, M. Sawada, M. Ujiie S. Nishimoto, Y. Ota, and M .Narita, 2020: A New Data assimilation System and Upgrading of Phaysical Processes in JMA's Meso-scale NWP System. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **50**, 01.07–01.08.

Ikuta, Y., M. Sato, M.Sawada, H. Kusabiraki, and T Kubota, 2021: Improvement of the Cloud Microphysics Scheme of the Mesoscale Model at the Japan Meteorological Agency Using Spaceborne Radar and Microwave Imager of the Global Precipitation Mesurement as Reference. *Mon. Wea. Rev.*, **149**, 3803–3819.

Imai, T. and R. Yoshida, 2016: Algorithm Theoretical Basis for Himawari-8 Cloud Mask Product. *Meteorological Satellite Center Technical Note*, **61**, 1–17.

Ishida, J., 2007: Development of a hybrid terrain-following vertical coordinate for JMA Non-hydrostatic Model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **37**, 03.09–03.10.

Ishida, J., K. Aranami, K. Kawano, K. Matsubayashi, Y. Kitamura, and C. Muroi, 2022: ASUCA: the JMA operational non-hydrostatic model. *J. Meteor. Soc. Japan*, **100**, 825 – 846.

Ishida, J., C. Muroi, and Y. Aikawa, 2009: Development of a new dynamical core for the nonhydrostatic model.

CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell., **39**, 05.09–05.10.

Ishida, J., C. Muroi, K. Kawano, and Y. Kitamura, 2010: Development of a new nonhydrostatic model "ASUCA" at JMA. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **40**, 05.11–05.12.

Ishihara, M., Y. Kato, T. Abo, K. Kobayashi, and Y. Izumikawa, 2006: Characteristics and Performance of the Operational Wind Profiler Network of the Japan Meteorological Agency. *J. Meteor. Soc. Japan*, **84**, 1085–1096.

Ishii, K., T. Shimbori, R. Kai, Y. Hasegawa, Y. Hayashi, and H. Tsuchiyama, 2021: Improvement of volcanic ash cloud prediction in the Tokyo Volcanic Ash Advisory Center. *WGNE Res. Activ. Earth Sys. Modell.*, **51**, 05.03–05.04.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865–879.

Ishikawa, Y., 2010: Data assimilation of GPS precipitable water vapor into the JMA mesoscale numerical weather prediction model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **40**, 1–13.

Ishikawa, Y. and K. Koizumi, 2002: Meso-scale analysis. *Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System*, Tokyo, Japan, Japan Meteorological Agency, 26–31.

Iwamura, K. and H. Kitagawa, 2008: An upgrade of the JMA operational global NWP model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **38**, 06.3–06.4.

Iwasaki, T., T. Maki, and K. Katayama, 1998: Tracer transport model at Japan Meteorological Agency and its application to the ETEX data. *Atmos. Env.*, **32**, 4285–4295.

Iwasaki, T., S. Yamada, and K. Tada, 1989: A parameterization scheme of orographic gravity wave drag with two different vertical partitionings, Part I: Impacts on medium-range forecasts. *J. Meteor. Soc. Japan*, **67**, 11–27.

Jakob, C. and A. P. Siebesma, 2003: A new subcloud model for mass-flux convection schemes: Influence on triggering, updraft properties, and model climate. *Mon. Wea. Rev.*, **131**, 2765–2778.

Janssen, P., 2004: *The Interaction of Ocean Waves and Wind*. Cambridge Univ. Press, 308pp.

Japan Meteorological Agency, 2006: Characteristics of Global Sea Surface Temperature Analysis Data (COBE-SST) for Climate Use. *Monthly Report on Climate System Separated Volume*, **12**, 116pp.

JMA, 1997: Objective Interpretation of NWP Products. *Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System*, Japan Meteorological Agency, Tokyo, Japan, 123–124.

JMA, 2002: *Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System*. Japan Meteorological Agency, Tokyo, Japan.

JMA, 2007: *Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System and Numerical Weather Prediction Research*. Japan Meteorological Agency, Tokyo, Japan.

JMA, 2013: *Outline of the operational numerical weather prediction at the Japan Meteorological Agency, Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research*. Japan Meteorological Agency, Tokyo, Japan.

JMA, 2017: Joint WMO Technical Progress Report on the Global Data Processing and Forecasting System and Numerical Weather Prediction Research Activities for 2016. 31–32.

JMA, 2018: *Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System*. Japan Meteorological Agency, Tokyo, Japan.

JMA, 2019: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System and Numerical Weather Prediction.

Jolliffe, I. T. and D. B. Stephenson, 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley, 254pp.

Joseph, J. H., W. J. Wiscombe, and J. A. Weinman, 1976: The delta-Eddington approximation for radiative flux transfer. *J. Atmos. Sci.*, **33**, 2452–2459.

Juang, H.-M. H., 2004: A Reduced Spectral Transform for the NCEP Seasonal Forecast Global Spectral Atmospheric Model. *Mon. Wea. Rev.*, **132**, 1019–1035.

Kadowaki, T, 2005: A 4-dimensional variational assimilation system for the JMA Global Spectrum Model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **35**, 01.17–01.18.

Kadowaki, T., Y. Ota, and S. Yokota, 2020: Introduction of a new hybrid data assimilation system for the JMA Global Spectral Model. *WGNE Res. Activ. Earth System Modell.*, **50**, 01.09–01.10.

Kadowaki, T. and K. Yoshimoto, 2012: A new inner model with a higher horizontal resolution (TL319) in JMA's Global 4D-Var data assimilation system. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, 01.09–01.10.

Kain, J. S., 2004: The Kain-Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181.

Kain, J. S. and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802.

Kajino, M., A. Kamada, N. Tanji, M. Kuramochi, M. Deushi, and T. Maki, 2022: Quantitative influences of interannual variations in meteorological factors on surface ozone concentration in the hot summer of 2018 in Japan. *Atmos. Environ.: X*, **16**, 100191 pp.

Kajino, M., M. Deushi, T. T. Sekiyama, N. Oshima, K. Yumimoto, T. Y. Tanaka, J. Ching, A. Hashimoto, T. Yamamoto, M. Ikegami, A. Kamada, M. Miyashita, Y. Inomata, S. Shima, A. Takami, A. Shimizu, S. Hatakeyama, Y. Sadanaga, H. Irie, K. Adachi, Y. Zaizen, Y. Igarashi, H. Ueda, T. Maki, and M. Mikami, 2019: NHM-Chem, the Japan Meteorological Agency's Regional Meteorology — Chemistry Model: Model Evaluations toward the Consistent Predictions of the Chemical, Physical, and Optical Properties of Aerosols. *J. Meteor. Soc. Japan*, **97**, 337–374.

Kakehata, T., M. Kunii, K. Kawano, and H. Kawada, 2021: Upgrade of JMA's Mesoscale Ensemble Prediction System. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **51**, 5.05–5.06.

Kanda, M., T. Kawai, M. Kanega, R. Moriwaki, K. Narita, and A. Hagishima, 2005: A Simple Energy Balance Model for Regular Building Arrays. *Bound.-Layer Meteor.*, **116**, 423–443.

Kanehama, T., H. Yonehara, and M. Ujiie, 2023a: The impact of a high-accuracy high-resolution digital elevation model on numerical weather predictions. *Res. Activ. Earth. Sys. Modell,*, 6.5–6.6.

Kanehama, T., H. Yonehara, and M. Ujiie, 2023b: The impact of a high-accuracy high-resolution digital elevation model on numerical weather predictions. *Res. Activ. Earth. Sys. Modell*, 6.5–6.6.

Kanehama, T., H. Yonehara, and M. Ujiie, 2023c: The impact of a high-accuracy high-resolution digital elevation model on numerical weather predictions. *Res. Activ. Earth. Sys. Modell*, 6.5–6.6.

Kannari, A., Y. Tonooka, T. Baba, and K. Murano, 2007: Development of multiple-species 1km × 1km resolution hourly basis emissions inventory for Japan. *Atmos. Environ.*, **41**, 3428–3439.

Kawada, H, T. Kakehata, and K. Kawano, 2023: Implementation of the SPPT scheme in JMA's Mesoscale Ensemble Prediction System. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **53**, 5.11–5.12.

Kawai, H., 2002: Forecast of sulfur dioxide flow from Miyake volcano with a high resolution regional transport model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **32**, 05.24–05.25.

Kawai, H., 2005: Improvement of a cloud ice fall scheme in GCM. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **35**, 04.11–04.12.

Kawai, H. and T. Inoue, 2006: A simple parameterization scheme for subtropical marine stratocumulus. *SOLA*, **2**, 17–20.

Kawai, H., T. Koshiro, and M. J. Webb, 2017: Interpretation of factors controlling low cloud cover and low cloud feedback using a unified predictive index. *J. Climate*, **30**, 9119–9131.

Kawai, H., H. Yonehara, and M. Ujiie, 2013: Vertical Layer Placement in Eta Coordinate for Models with a High Model Top. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **43**, 03.3–03.4.

Kessler, E., 1969: *On the distribution and continuity of water substance in atmospheric circulation*. Meteorol. Monogr., American Meteorol. Soc., Boston, MA, 84pp.

Kikuchi, M., H. Murakami, K. Suzuki, T. M. Nagao, and A. Higurashi, 2018: Improved Hourly Estimates of Aerosol Optical Thickness Using Spatiotemporal Variability Derived From Himawari-8 Geostationary

Satellite. *IEEE Trans. Geosci. Remote Sens.*, **56**, 3442–3455.

Kim, H., 2017: Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set]. *Data Integration and Analysis System (DIAS)*, URL https://doi.org/10.20783/DIAS.501.

Kitada, T., 1994: Modelling of transport, reaction and deposition of acid rain. *Kishou Kenkyu Note*, **182**, 95–117, (in Japanese).

Kitada, T., G. R. Carmichael, and L. K. Peters, 1986: Effects of dry deposition on the concentration-distributions of atmospheric pollutants within land- and sea-breeze circulations. *Atmos. Env.*, **20**, 1999–2010.

Kitamura, Y. and J. Ito, 2016: Revisiting the bulk relation for heat flux in the free convection limit. *Bound.-Layer Meteor.*, **158**, 93–103.

Klemp, J. B., W. C. Skamarock, and J. Dudhia, 2007: Conservative split-explicit time integration methods for the compressible nonhydrostatic equations. *Mon. Wea. Rev.*, **135**, 2897–2913.

Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48.

Kohno, N., D. Miura, and K. Yoshita, 2012: The development of JMA wave data assimilation system. *Proceedings of the 12th International Workshop on Wave Hindcasting and Forecasting and Coastal Hazards Symposium*, Kohala Coast, Hawaii, 30 October - 4 November 2011, H2, (http://www.waveworkshop.org).

Kohno, N., K. Murotani, H. Minematsu, and D. Miura, 2009: The improvement of JMA operational wave forecasting system. *Proceedings of the 11th International Workshop on Wave Hindcasting and Forecasting and Coastal Hazards Symposium*, Halifax, Canada, 18 - 23 October 2009, R3, (http://www.waveworkshop.org).

Koizumi, K., Y. Ishikawa, and T. Tsuyuki, 2005: Assimilation of precipitation data to the JMA mesoscale model with a four-dimensional variational method and its impact on precipitation forecasts. *SOLA*, **1**, 45–48.

Komori, T., S. Hirahara, and R. Sekiguchi, 2020: Improved representation of convective moistening in JMA's next-generation coupled seasonal prediction system. *WGNE blue book, Res. Activ. Earth Sys. Modell*, **50**, 4–05.

Konishi, T., 1995: An experimental storm surge prediction for the western part of the Inland Sea with application to Typhoon 9119. *Pap. Meteor. Geophys.*, **46**, 9–17.

Koren, B., 1993: A robust upwind discretization method for advection, diffusion and source terms. *CWI Technical Report NM-R 9308*, 1–22, URL http://oai.cwi.nl/oai/asset/5293/05293D.pdf.

Kosaka, Y., S. Kobayashi, Y. Harada, C. Kobayashi, H. Naoe, K. Yoshimoto, M. Harada, N. Goto, J. Chiba, K. Miyaoka, R. Sekiguchi, M. Deushi, H. Kamahori, T. Nakaegawa, T. Y. Tanaka, T. Tokuhiro, Y. Sato, Y. Matsushita, and K. Onogi, 2024: The JRA-3Q reanalysis. *J. Meteor. Soc. Japan*, **102**, doi:10.2151/jmsj.2024-004.

Koyaguchi, T., 2008: *Modeling of Volcanic Phenomena*. Univ. Tokyo Press, 637 pp.

Kudo, A, 2011: Development of JMA's new turbulence index. *15th Conference on Aviation, Range, and Aerospace Meteorology*, Amer. Met. Soc., Los Angeles, CA, 1 - 4 August 2011.

Kuragano, T., Y. Fujii, T. Toyoda, N. Usui, K. Ogawa, and M. Kamachi, 2014: Seasonal barotropic sea surface height fluctuation in relation to regional ocean mass variation. *J. Oceanogr.*, **70**, 45–62.

Kuragano, T. and M. Kamachi, 2000: Global statistical space-time scales of oceanic variability estimated from the TOPEX/POSEIDON altimeter data. *J. Geophys. Res.*, **105**, 955–974.

Kurihara, Y., Murakami H., and M. Kachi, 2016: Sea surface temperature from the new Japanese geostationary meteorological Himawari-8 satellite. *Geophys. Res. Lett.*, **43**, 1234–1240.

Kurihara, Y., T. Sakurai, and T. Kuragano, 2006: Global daily sea surface temperature analysis using data from satellite microwave radiometer, satellite infrared radiometer and in-situ observations. *Weather Bulletin, JMA*, **73**, s1–s18.

Kyouda, M. and M. Higaki, 2015: Upgrade of JMAs Typhoon Ensemble Prediction System. *RSMC Tokyo-Typhoon Center Technical Review*, **17**, 1–13.

Lalaurette, F., 2003: Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Q. J. R. Meteorol. Soc.*, **129**, 3037–3057.

Lee, T. J. and R. A. Pielke, 1992: Estimating the soil surface specific humidity. *J. Appl. Meteor.*, **31**, 480–484.

Leonard, B. P., 1979: A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Computer Methods in Applied Mechanics and Engineering*, **19**, 59–98, URL https://www.sciencedirect.com/science/article/pii/0045782579900343.

Li, J., 2002: Accounting for unresolved clouds in a 1D infrared radiative transfer model. Part I: Solution for radiative transfer, including cloud scattering and overlap. *J. Atmos. Sci.*, **59**, 3302–3320.

Li, J. and H. W. Barker, 2005: A radiation algorithm with correlated-k distribution. Part I: Local thermal equilibrium. *J. Atmos. Sci.*, **62**, 286–309.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.

Lindner, T. H. and J. Li, 2000: Parameterization of the optical properties for water clouds in the infrared. *J. Climate*, **13**, 1797–1805.

Liu, D. C. and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization. *Math. Programming*, **45**, 503–528.

Liu, H., K. C. Jezek, B. Li, and Z. Zha, 2015: Radarsat Antarctic Mapping Project Digital Elevation Model, Version 2. *NASA National Snow and Ice Data Center Distributed Active Archive Center*.

Locarnini, R. A., A. V. Mishonov, O. K. Baranova, T. P. Boyer, M. M. Zweng, H. E. Garcia, J. R. Reagan, D. Seidov, K. Weathers, C. R. Paver, and I. Smolyar, 2018: *World Ocean Atlas 2018, Volume 1: Temperature*. NOAA Atlas NESDIS 81, 52 pp.

Lord, S. J. and A. Arakawa, 1980: Interaction of a cumulus cloud ensemble with the large-scale environment. Part II. *J. Atmos. Sci.*, **37**, 2677–2692.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP: a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.

Lorenc, A. C. and F. Rawlins, 2005: Why does 4D-Var beat 3D-Var? *Quart. J. Roy. Meteor. Soc.*, **131**, 3247–3257.

Lott, F. and M. J. Miller, 1997: A New Subgrid Orographic Drag Parameterization : Its Formulatin and Testing. *Quart. J. Roy. Meteor. Soc.*, **123**, 101–127.

Louis, J.-F., M. Tiedtke, and J.-F. Geleyn, 1982: A short history of the operational PBL-parameterization at ECMWF. *Workshop on Planetary Boundary Layer Parameterization*, Reading, United Kingdom, 25 - 27 November 1981, ECMWF, 59–79.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Youing, and J. W. Merchant, 2000: Development of a global land cover characteristics database and IGBP DISCover from 1km AVHRR data. *Int. J. Remote Sensing*, 1303–1330.

Lyard, F., F. Lefèvre, T. Letellier, and O. Francis, 2006: Modelling the global ocean tides: Modern insights from FES2004. *Ocean Dyn.*, **56**, 394–415.

Lyard, F. H., D. J. Allain, M. Cancet, L. Carrère, and N Picot, 2021: FES2014 global ocean tide atlas: design and performance. *Ocean Science*, **17**, 615–649.

Lynch, P., 1997: The Dolph-Chebyshev Window: A Simple Optimal Filter. *Mon. Wea. Rev.*, **125**, 655–660.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialization. *Contrib. Atmos. Phys.*, **50**, 253–271.

Mahfouf, J.-F., 1999: Influence of physical process on the tangent-linear approximation. *Tellus*, **51A**, 147–166.

Makihara, Y., 2000: Algorithms for precipitation nowcasting focused on detailed analysis using radar and raingauge data. *Technical Reports of the Meteorological Research Institute*, **39**, 63–111.

Martin, G. M., D. W. Johnson, and A. Spice, 1994: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *J. Atmos. Sci.*, **51**, 1823–1842.

Matsumoto, K., T. Takanezawa, and M. Ooe, 2000: Ocean Tide Models Developed by Assimilating TOPEX/POSEIDON Altimeter Data into Hydrodynamical Model: A Global Model and a Regional Model around Japan. *J. Oceanogr.*, **56**, 567–581.

Matsumoto, T., M. Ishii, Y. Fukuda, and S. Hirahara, 2006: Sea Ice Data Derived from Microwave Radiometer for Climate Monitoring. *14th Conference on Satellite Meteorology and Oceanography, Amer. Meteor. Soc.*, P2.21.

McLandress, C. and J. F. Scinocca, 2005: The GCM response to current parameterizations of nonorographic gravity wave drag. *J. Atmos. Sci.*, **62**, 2394–2413.

Meador, W. E. and W. R. Weaver, 1980: Two-stream approximations to radiative transfer in planetary atmospheres: A unified description of existing methods and a new improvement. *J. Atmos. Sci.*, **37**, 630–643.

Mellor, G. L., 1977: The Gaussian cloud model relations. *J. Atmos. Sci.*, **34**, 356–358.

Mellor, G. L. and L. Kantha, 1989: An ice-ocean coupled model. *J. Geophys. Res.*, **94**, 10 937–10 954.

Mellor, G. L. and T. Yamada, 1974: A hierarchy of turbulence closure models for planetary boundary layers. *J. Atmos. Sci.*, **31**, 1791–1806.

Mellor, G. L. and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875.

Mitsuyasu, H. and T. Honda, 1982: Wind-induced growth of water waves. *J. Fluid Mech.*, **123**, 425–442.

Miyamoto, K., 2006: Introduction of the Reduced Gaussian Grid into the Operational Global NWP Model at JMA. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **36**, 06.09–06.10.

Mizuta, R., H. Yoshimura, H. Murakami, M. Matsueda, H. Endo, T. Ose, K. Kamiguchi, M. Hosaka, M. Sugi, S. Yukimoto, S. Kusunoki, and A. Kitoh, 2012: Climate simulations using MRI-AGCM3.2 with 20-km grid. *J. Meteor. Soc. Japan*, **90A**, 233–258.

Moorthi, S. and M. J. Suarez, 1992: Relaxed Arakawa-Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.*, **120**, 978–1002.

Munk, W. H. and D. E. Cartwright, 1966: Tidal Spectroscopy and Prediction. *Phil. Trans. Roy. Soc. A*, **259**, 533–581.

Murphy, A. H., 1973: A New Vector Partition of the Probability Score. *J. Appl. Meteor.*, **12**, 595–600.

Murray, Ross J, 1996: Explicit generation of orthogonal grids for ocean models. *Journal of Computational Physics*, **126**, 251–273.

Myneni, Knyazikhin Y. Park T., R., 2015: MOD15A2H MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC, Accessed 2022-07-20 from https://doi.org/10.5067/MODIS/MOD15A2H.006.

Myneni, R. B., S. Hoffman, Y. Knyazikhin, J. L. Privette, J. Glassy, Y. Tian, Y. Wang, X. Song, Y. Zhang, G. R. Smith, A. Lotsch, M. Friedl, J. T. Morisette, P. Votava, R. R. Nemani, and S. W. Running, 2002: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.*, **83**, 214–231.

Nagasawa, R., 2012: The problem of cloud overlap in the radiation process of JMA's global NWP model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **42**, 4.15–4.16.

Naito, I. and N. Kikuchi, 1992: Atmospheric contributions to non-seasonal variations in the length of day. *Geophys. Res. Lett.*, **19**, 1843–1846.

Nakagawa, M., 2009: Outline of the high resolution global model at the Japan Meteorological Agency. *RSMC Tokyo-Typhoon Center Technical Review*, **11**, 1–13.

Nakanishi, M., 2001: Improvement of the Mellor-Yamada turbulence closure model based on large-eddy simulation data. *Bound.-Layer Meteor.*, **99**, 349–378.

Nakanishi, M. and H. Niino, 2004: An improved Mellor-Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.*, **112**, 1–31.

Nakanishi, M. and H. Niino, 2006: An improved Mellor-Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407.

Nakanishi, M. and H. Niino, 2009: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *J. Meteor. Soc. Japan*, **87**, 895–912.

Nakano, H. and N. Suginohara, 2002: Effects of bottom boundary layer parameterization on reproducing deep and bottom waters in a world ocean model. *J. Phys. Oceanogr.*, **32**, 1209–1227.

Narui, A, 2006: Changing the Resolution of the Inner Loop of Global 4D-Var at JMA. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **36**, 01.23–01.24.

Nishizawa, S. and Y. Kitamura, 2018: A Surface Flux Scheme Based on the Monin-Obukhov Similarity for Finite Volume Models. *J. Adv. Model. Earth Syst.*, **10**, 3159–3175.

Niwa, Y. and Y. Fujii, 2020: A conjugate BFGS method for accurate estimation of a posterior error covariance matrix in a linear inverse problem. *Quart. J. Roy. Meteor. Soc.*, **146**, 3118–3143.

Nocedal, J. and S. J. Wright, 2006: *Numerical Optimization*. 2d ed., Springer, 664pp.

Noilhan, J. and P. Lacarrére, 1995: GCM Grid-Scale Evaporation from Mesoscale Modeling. *J. Climate*, **8**, 206–223.

Noilhan, J. and J.-F. Mahfouf, 1996: The ISBA land surface parameterisation scheme. *Glob. Planet. Chang.*, **13**, 145–159.

Noilhan, J. and S. Planton, 1989: A Simple Parameterization of Land Surface Processes for Meteorological Models. *Mon. Wea. Rev.*, **117**, 536–549.

Ohara, T., H. Akimoto, J. Kurokawa, N. Horii, K. Yamaji, X. Yan, and T. Hayasaka, 2007: An Asian emission inventory of anthropogenic emission sources for the period 1980-2020. *Atmos. Chem. Phys.*, **7**, 4419–4444.

Oleson, K. W., D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, P. J. Lawrence, S. Levis, S. C. Swenson, P. E. Thornton, A. Dai, M. Decker, R. Dickinson, J. Feddema, C. L. Heald, F. Hoffman, J.-F. Lamarque, N. Mahowald, G.-Y. Niu, T. Qian, J. Randerson, S. Running, K. Sakaguchi, A. Slater, R. Stöckli, A. Wang, Z.-L. Yang, Xiaodong Zeng, and Xubin Zeng, 2010: Technical Description of version 4.0 of the Community Land Model (CLM). NCAR Technical Note 478, NCAR. 257pp.

Olson, J. B., J. S. Kenyon, W. M. Angevine, J. M. Brown, M. Pagowski, and K. Sušelj, 2019: A description of the MYNN-EDMF scheme And the Coupling to Other Components in WRF-ARW. *NOAA Tech. Memo. OAR GSD*, **61**, 37pp.

Ono, K., M. Kunii, and Y. Honda, 2021: The regional model-based Mesoscale Ensemble Prediction System, MEPS, at the Japan Meteorological Agency. *Quart. J. Roy. Meteor. Soc.*, **147**, 465–484.

Onogi, K., 1998: A Data Quality Control Method Using Forecasted Horizontal Gradient and Tendency in a NWP System: Dynamic QC. *J. Meteor. Soc. Japan*, **76**, 497–516.

Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, S. Kadokura, K. Wada, K. Kato, R. Oyama, T. Ose, N. Mannoji, and R. Taira, 2007: The JRA-25 reanalysis. *J. Meteor. Soc. Japan*, **85**, 369–432.

Ota, Y., J. Chiba, Y. Ichikawa, H. Oashi, T. Takakura, and H. Yamaguchi, 2023: Upgrade of JMA's Global Ensemble Prediction System. *WGNE Res. Activ. Earth Sys. Modell.*, **53**, 6.11–6.12.

Ota, Y., M. Ikegami, and H. Yamaguchi, 2019: Upgrade of initial perturbations made using the Local Ensemble Transform Kalman Filter in JMA's Global EPS. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **49**, 6.11–6.12.

Ou, S. and K.-N. Liou, 1995: Ice microphysics and climatic temperature feedback. *Atmos. Res.*, **35**, 127–138.

Oyama, R., 2010: Upgrade of atmospheric motion vector derivation algorithms at JMA/MSC. Meteorological Satellite Center Technical Note 54, Japan Meteorological Agency, Tokyo, Japan. 1–32. (https://www.data.jma.go.jp/mscweb/en/product/library/note/index.html).

Palmer, T. N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. J. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and model uncertainty. *ECMWF Tech. Memo*, **598**, 42 pp.

Palmer, T. N., G. J. Shutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. *Quart. J. Roy. Meteor. Soc.*, **112**, 1001–1039.

Parrish, D. and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Peng, S. and Y. Li, 2015: A parabolic model of drag coefficient for storm surge simulation in the South China Sea. *Sci. Rep.*, **5**, 15 496, 1–6.

Persson, A. O., 1991: Kalman Filtering - A new approach to adaptive statistical interpretation of numerical meteorological forecasts. *WMO Technical Document*, **421**, XX27–XX32.

Phillips, D. S., 1984: Analytical surface pressure and drag for linear hydrostatic flow over three-dimensional elliptical mountains. *J. Atmos. Sci.*, **41**, 1073–1084.

Plant, W. J., 1982: A relationship between wind stress and wave slope. *J. Geophys. Res.*, **87**, 1961–1967.

Prather, M. J., 1986: Numerical advection by conservation of second-order memoments. *J. Geophys. Res.*, **91**, 6671–6681.

Price, J. F., J. D. Weller, and R. Pinkel, 1986: Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. *J. Geophys. Res.*, **91**, 8411–8427.

Purser, R. J., W. S. Wu, D. F. Parrish, and N. M. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535.

Ramankutty, N., A. T. Evan, C. Monfreda, and J. A. Foley, 2008: Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1003.

Randall, D. and D.-M. Pan, 1993: Implementation of the Arakawa-Schubert cumulus parameterization with a prognostic closure. *The representation of cumulus convection in numerical models, AMS Meteorological Monograph Series*, **46**, 137–144.

Randel, W. J., F. Wu, J. M. Russell III, A. Roche, and J. W. Waters, 1998: Seasonal cycles and QBO variations in stratospheric CH_4 and H_2O observed in UARS HALOE data. *J. Atmos. Sci.*, **55**, 163–185.

Ray, R. D., 1999: A Global Ocean Tide Model From TOPEX/POSEIDON Altimetry: GOT99.2. Technical Memorandum NASA/TM-1999-209478, Goddard Space Flight Center. 58pp.

Redelsperger, J.-L., F. Guichard, and S. Mondon, 2000: A parameterization of mesoscale enhancement of surface fluxes for large-scale models. *J. Climate*, **13**, 402–421.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, 1154–1158.

Reed, M., 1989: The physical fates component of the natural resource damage assessment model system. *Oil and Chemical Pollution*, **5**, 99–123.

Reynolds, R. W., 1987: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.

Roberts, N. M. and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Wea. Rev.*, **136**, 78–97.

Rothman, L. S., A. Barbe, D. C. Benner, L. R. Brown, C. Camy-Peyret, M. R. Carleer, K. Chance, C. Clerbaux, V. Dana, V. M. Devi, A. Fayt, J.-M. Flaud, R. R. Gamache, A. Goldman, D. Jacquemart, K. W. Jucks, W. J. Lafferty, J.-Y. Mandin, S. T. Massie, V. Nemtchinov, D. A. Newnham, A. Perrin, C. P. Rinsland, J. Schroeder, K. M. Smith, M. A. H. Smith, K. Tang, R. A. Toth, J. Vander Auwera, P. Varanasi, and K. Yoshino, 2003: The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001. *J. Quant. Spectrosc. Radiat. Transfer*, **82**, 5–44.

Saito, K., J. Ishida, K. Aranami, T. Hara, T. Segawa, M. Narita, and Y. Honda, 2007: Nonhydrostatic Atmospheric Models and Operational Development at JMA. *J. Meteor. Soc. Japan*, **85B**, 271–304.

Saito, K., T. Fujita, Y. Yamada, J. Ishida, Y. Kumagai, K. Aranami, S. Ohmori, R. Nagasawa, S. Kumagai, C. Muroi, T. Kato, H. Eito, and Y. Yamazaki, 2006: The Operational JMA Nonhydrostatic Mesoscale Model. *Mon. Wea. Rev.*, **134**, 1266–1298.

Sakamoto, K., H. Nakano, S. Urakawa, T. Toyoda, Y. Kawakami, H. Tsujino, and G. Yamanaka, 2023: Reference Manual for the Meteorological Research Institute Community Ocean Model Version 5 (MRI.COMv5). Technical Reports of the Meteorological Research Institute 87, Meteorological Research Institute of Japan Meteorological Agency, Ibaraki, Japan. 334pp. (https://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_87/tech_rep_mri-v5.pdf).

Sakamoto, K., H. Tsujino, H. Nakano, S. Urakawa, T. Toyoda, N. Hirose, N. Usui, and G. Yamanaka, 2019: Development of a 2-km resolution ocean model covering the coastal seas around Japan for operational application. *Ocean Dyn*, **69**, 1181–1202.

Sato, N., P. J. Sellers, D. A. Randall, E. K. Schneider, J. Shukla, J. L. Kinter III, Y-T Hou, and E. Albertazzi, 1989: Effects of implementing the simple biosphere model in a general circulation model. *J. Atmos. Sci.*, **46**, 2757–2782.

Saunders, R, J Hocking, E Turner, S Havemann, A Geer, C Lupu, J Vidot, P Chambon, C Köpken-Watts, L Scheck, and others, 2020: RTTOV-13 science and validation report. Tech. rep., EUMETSAT NWP SAF, 0 pp. https://nwp-saf.eumetsat.int/site/download/documentation/rtm/docs_rttov13/rttov13_svr.pdf.

Saunders, R, 2008: RTTOV-9 science and validation report. Tech. rep., EUMETSAT NWP SAF, 74 pp. (https://nwp-saf.eumetsat.int/site/software/rttov/).

Saunders, R., J. Hocking, P. Rayer, M. Matricardi, A. Geer, N. Bormann, P. Brunel, F. Karbou, and F. Aires,

2012: RTTOV-10 science and validation report. Tech. rep., EUMETSAT, 31 pp. (https://nwp-saf.eumetsat.int/oldsite/deliverables/rtm/docs_rttov10/rttov10_svr_1.11.pdf).

Schaaf, C. B., F. Gao, A. H. Strahler, W. Lucht, X. Li, T. Tsang, N. C. Strugnell, X. Zhang, Y. Jin, J. P. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R. P. d'Entremont, B. Hu, S. Liang, J. L. Privette, and D. P. Roy, 2002: First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sens. Environ.*, **83**, 135–148.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575.

Scinocca, J. F., 2003: An accurate spectral nonorographic gravity wave drag parameterization for general circulation models. *J. Atmos. Sci.*, **60**, 667–682.

Sekiguchi, R., Y. Ichikawa, K. Ochi, and T. Takakura, 2022: Hindcast verification of JMA's GEPS for one-month prediction. *WGNE Res. Activ. Earth Sys. Modell.*, **52**, 6.3–6.4.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.*, **43**, 505–531.

Selwood, P., 2012: The Met Office Unified Model I/O Server. *ENES Workshop: Scalable IO in climate models, available at: https://verc.enes.org/computing/hpc-collaborations/parallel-io/ workshop-scalable-io-in-climate-models/presentations/Unified_Model_IO_Server_Paul_Selwood.ppt/view*.

Shao, Y. and L. M. Leslie, 1997: Wind erosion prediction over the Australian continent. *J. Geophys. Res.*

Shao, Y. and H. Lu, 2000: A simple expression for wind erosion threshold friction velocity. *J. Geophys. Res.*

Shibata, K., M. Deushi, T. T. Sekiyama, and H. Yoshimura, 2005: Development of an MRI chemical transport model for the study of stratospheric chemistry. *Pap. Meteorol. Geophys.*, **55**, 75–119.

Shimbori, T., Y. Aikawa, and N. Seino, 2009: Operational implementation of the tephra fall forecast with the JMA mesoscale tracer transport model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **39**, 5.29–5.30.

Shimbori, T. and K. Ishii, 2021: Design of the Japan Meteorological Agency atmospheric transport model. *Tech. Rep. MRI*, **84**, 146pp.

Shimokobe, A, 2012: Improvement of the Stratocumulus Parameterization Scheme in JMA's Operational Global Spectral Model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, 4.17–4.18.

Shoji, Y., K. Sato, M. Yabuki, and T. Tsuda, 2017: Comparison of shipborne GNSS-derived precipitable water vapor with radiosonde in the western North Pacific and in the seas adjacent to Japan. *Earth, Planets and Space*, **69(1)**, 1–13.

Simmons, A. J. and D. M. Burridge, 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758–766.

Simonsen, C., 1991: Self adaptive model output statistics based on Kalman filtering. *WMO Technical Document*, **421**, XX33–XX37.

Slingo, A., 1989: A GCM parameterization for the shortwave radiative properties of water clouds. *J. Atmos. Sci.*, **46**, 1419–1427.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations I. The basic experiment. *Mon. Wea. Rev.*, **91**, 99–164.

Smith, R. N. B., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Quart. J. Roy. Meteor. Soc.*, **116**, 435–460.

Smith, S. D. and E. G. Banke, 1975: Variation of the sea surface drag coefficient with wind speed. *Quart. J. Roy. Meteor. Soc.*, **101**, 665–673.

Sommeria, G. and J. W. Deardorff, 1977: Subgrid-scale condensation in models of nonprecipitating clouds. *J. Atmos. Sci.*, **34**, 344–355.

Stanski, L. J. Wilson, H. R. and W. R. Burrows, 1989: Survey of common verification methods in meteorology. *Research Rep.*, **89-5**, Forecast Research Division, Atmospheric Envirnment Service, Environment Canada, 114pp.

Sun, Z., 2001: Reply to comments by Greg M. McFarquhar on 'Parametrization of effective sizes of cirrus-cloud particles and its verification against observations'. (October B, 1999, 125, 3037-3055). *Quart. J. Roy. Meteor. Soc.*, **127**, 267–271.

Sundqvist, H., 1978: A parameterization scheme for non-convective condensation including prediction of cloud water content. *Quart. J. Roy. Meteor. Soc.*, **104**, 677–690.

Sundqvist, H., E. Berge, and J. E. Kristjánsson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1657.

Suzuki, T., 1983: A theoretical model for dispersion of tephra. *Arc Volcanism: Physics and Tectonics*, Shimozuru, D. and I. Yokoyama, Eds., TERRAPUB, Tokyo, 95–113.

Takakura, T. and T. Komori, 2020: Two-tiered sea surface temperature approach implemented to JMA's Global Ensemble Prediction System. *WGNE Res. Activ. Earth Sys. Modell.*, **50**, 6.15–6.16.

Takasa, S, M. Chikasawa, and H. Mori, 2011: A new method of tide prediction along the Japanese coastline and its accuracy. *Weather Bulletin, JMA*, **78**, S33–S42, (in Japanese).

Takaya, Y., T. Yasuda, Y. Fujii, S. Matsumoto, T. Soga, H. Mori, M. Hirai, I. Ishikawa, H. Sato, A. Shimpo, M. Kamachi, and T. Ose, 2017: Japan Meteorological Agency/Meteorological Research Institute-Coupled Prediction System version1 (JMA/MRI-CPS1) for operational seasonal forecasting. *Clim, Dyn*, **48**, 313–333.

Takeuchi, Y. and T. Kurino, 1997: Document of algorithm to derive rain rate and precipitation with SSM/I and AMSR. *Algorithm description of PIs for SSM/I and ADEOS-II/AMSR, 2nd AMSR Workshop*, Tokyo, Japan, 61.1–61.9.

Takeuchi, Y. and T. Tsuyuki, 2002: The operational 3D-Var assimilation system of JMA for the Global Spectrum Model and the Typhoon Model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **32**, 1.59–1.60.

Tanaka, T. Y. and M. Chiba, 2005: Global Simulation of Dust Aerosol with a Chemical Transport Model, MASINGAR. *J. Meteor. Soc. Japan*, **83A**, 255–278.

Tanaka, T. Y., K. Orito, T. T. Sekiyama, K. Shibata, and M. Chiba, 2003: MASINGAR, a global tropospheric aerosol chemical transport model coupled with MRI/JMA98 GCM: Model description. *Pap. Meteorol. Geophys.*, **53 (4)**, 119–138.

Tanguay, M., E. Yakimiw, H. Ritchie, and A. Robert, 1992: Advantage of spatial averaging in semi-implicit semi-Lagrangian schemes. *Mon. Wea. Rev.*, **120**, 113–123.

Taylor, A. and B. Glahn, 2008: Probabilistic guidance for hurricane storm surge. *19th Conference on probability and statistics, Vol. 74*, 1–10.

The WAMDI Group, 1988: The WAM model—A third generation ocean wave prediction model. *J. Phys. Oceanogr.*, **18**, 1775–1810.

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.*, **121**, 3040–3061.

Tokioka, T., K. Yamazaki, A. Kitoh, and T. Ose, 1988: The equatorial 30-60 day oscillation and the Arakawa-Schubert penetrative cumulus parameterization. *J. Meteor. Soc. Japan*, **66**, 883–901.

Toyoda, T., Y. Fujii, T. Yasuda, N. Usui, T. Iwao, T. Kuragano, and T. Kamachi, 2013: Improved analysis of the seasonal interannual fields by a global ocean data assimilation system. *Theoretical and Applied Mechanics Japan*, **61**, 31–48.

Toyoda, T., Y. Fujii, T. Yasuda, N. Usui, K. Ogawa, T. Kuragano, H. Tsujino, and M. Kamachi, 2016: Data assimilation of sea ice concentration into a global ocean-sea ice model with corrections for atmospheric forcing and ocean temperature fields. *J. Oceanogr.*, **72**, 235–262.

Toyoda, T., T. Awaji, N. Sugiura, S. Masuda, H. Igarashi, Y. Sasaki, Y. Hiyoshi, Y. Ishikawa, T. Mochizuki, T. Sakamoto, H. Tatebe, Y. Komuro, T. Suzuki, T. Nishimura, M. Mori, Y. Chikamoto, S. Yasunaka, Y. Imada, M. Arai, M. Watanabe, H. Shiogama, T. Nozawa, A. Hasegawa, M. Ishii, and M. Kimoto, 2011: Impact of the assimilation of sea ice concentration data on an atmosphere-ocean-sea ice coupled simulation of the arctic ocean climate. *SOLA*, **7**, 37–40.

Trémolet, Y, 2008: Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus*, **60**.

Tsujino, H., T. Motoi, I. Ishikawa, M. Hirabara, H. Nakano, G. Yamanaka, T. Yasuda, and H. Ishizaki, 2010: Reference Manual for the Meteorological Research Institute Community Ocean Model (MRI.COM) Version 3. Technical Reports of the Meteorological Research Institute 59, Meteorological Research Institute of Japan Meteorological Agency, Ibaraki, Japan. 241pp. (http://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_59/59_en.html).

Tsujino, H., H. Nakano, K. Sakamoto, S. Urakawa, M. Hirabara, H. Ishizaki, and G. Yamanaka, 2017: Reference manual for the Meteorological Research Institute Community Ocean Model version 4 (MRI.COMv4). Technical Reports of the Meteorological Research Institute 80, Meteorological Research Institute of Japan

Meteorological Agency, Ibaraki, Japan. 284pp. (https://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_80/index_en.html).

Tsujino, H, S. Urakawa, H. Nakano, R. J. Small, W. M. Kim, S. G. Yeager, G. Danabasoglu, T. Suzuki, J. L. Bamber, M. Bentsen, C. W. Boning, A. Bozec, E. P. Chassignet, E. Curchitser, F. Boeira Dias, P. J. Durack, S. M. Griffies, Y. Harada, M. Ilicak, S. A. Josey, C. Kobayashi, S. Kobayashi, Y. Komuro, W. G. Large, J. Le Sommer, S. J. Marsland, S. Masina, M. Scheinert, H. Tomita, M. Valdivieso, and D. Yamazaki, 2018: JRA-55 based surface dataset for driving ocean-sea-ice models (JRA55-do). *Ocean Modelling*, **130**, 79–139, URL https://www.sciencedirect.com/science/article/pii/S146350031830235X.

Ueno, K., 1998: An energy dissipation term in a numerical prediction wave model. *Sokkou-Jihou*, **65**, S181–S187, (in Japanese).

Ueno, K. and N. Kohno, 2004: The development of the third generation wave model MRI-III for operational use. *Proceedings of the 8th International Workshop on Wave Hindcasting and Forecasting*, North Shore, Oahu, Hawaii, 14 - 19 November 2004, G2, (http://www.waveworkshop.org).

Ujiie, M., M. Higuchi, T. Kadowaki, Y. Kuroki, K. Miyaoka, M. Oda, K. Ochi, R. Sekiguchi, H. Shimizu, S. Yokota, and H. Yonehara, 2021: Upgrade of JMA's operational global NWP system. *Res. Activ. Earth. Sys. Modell,*, 6.9–6.10.

Umlauf, L. and H. Burchard, 2003: A generic length-scale equation for geophysical turbulence models. *JMR*, **61**, 235–265.

Usui, N., Y. Fujii, K. Sakamoto, and M. Kamachi, 2015: Development of a four-dimensional variational assimilation system toward coastal data assimilation around Japan. *Mon. Wea. Rev.*, **143**, 3874–3892.

Usui, N., S. Ishizaki, Y. Fujii, H. Tsujino, T. Yasuda, and M. Kamachi, 2006: Meteorological Research Institute multivariate ocean variational estimation (MOVE) system: Some early results. *Adv. Space Res.*, **37**, 806–822.

van de Berg, L., J. Gustafsson, and A. Yildirim, 2002: Reprocessing of atmospheric motion vectors from Meteosat image data. ECMWF ERA-40 Project Report Series 3, European Centre for Medium-Range Weather Forecasts, Reading, UK. 159–168. (https://www.ecmwf.int/en/publications).

Van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S. J. Smith, and S. K. Rose, 2011: The representative concentration pathways: an overview. *Climatic Change*, **109**, 5–31.

Veersé, F., D. Auroux, and M. Fisher, 2000: Limited-memory BFGS diagonal preconditioners for a data assimilation problem in meteorology. *Optimization and Engineering*, **1**, 323–339.

Vialard, J., C. Menkes, J.-P. Boulanger, P. Delecluse, E. Guilyardi, M. J. McPhaden, and G. Madec, 2001: A model study of oceanic mechanisms affecting equatorial Pacific Sea surface temperature during the 1997-98 El Niño. *J. Phys. Oceanogr*, **31**, 1649–1675.

Vosper, S. B., 2015: Mountain waves and wakes generated by South Georgia: Implications for drag parametrization. *Quart. J. Roy. Meteor. Soc.*, **141**, 2813–2827.

Wan, Z., 2014: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sensing of Environment*, **140**, 36–45.

Wee, T. K. and Y. H. Kuo, 2004: Impact of a digital filter as a weak constraint in MM5 4DVAR: An observing system simultion experiment. *Mon. Wea. Rev.*, **132**, 543–559.

Westerink, J. J., R. A. Luettich, and J. C. Feyen, 2008: A Basin-to Channel-Scale Unstructured Grid Hurricane Storm Surge Model Applied to Southern Louisiana. *Mon. Wea. Rev.*, **136**, 833–864.

Wicker, L. J. and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences, Second Edition*, International Geophysics Series, Vol. 91. Academic Press, 627pp.

Wilson, D. R. and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK meteorological office unified model. **125**, 1607–1636.

WMO, 1993: *Guide on the Global Data-processing System*. WMO-No.305, World Meteorological Organization.

WMO, 2019: Manual on the Global Data-processing and Forecasting System. *Annex IV to the WMO Technical Regulations*, World Meteorological Organization, Geneva, Switzerland.

WMO, 2021: The state of Greenhouse Gases in the Atmosphere Based on Global Observations through 2020. *WMO GREENHOUSE GAS BULLETIN*, **17**, 1–10.

Woodruff, S.D., H.F. Diaz, J.D. Elms, and S.J. Worley, 1998: COADS release 2 data and metadata enhancements for improvements of marine surface flux fields. *Physics and Chemistry of the Earth*, **23**, 517–526.

Woodruff, S.D., S.J. Worley, S.J. Lubker, Z. Ji, J.E. Freeman, D.I. Berry, P. Brohan, E.C. Kent, R.W. Reynolds, S.R. Smith, and C. Wilkinson, 2011: ICOADS release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967.

Xie, S. C.. and M. H. Zhang, 2000: Impact of the convection triggering function on single-column model simulations. *J. Geophys. Res.*, **105**, 14 983–14 996.

Yabu, S., 2013: Development of longwave radiation scheme with consideration of scattering by clouds in JMA global model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **43**, 4.07–4.08.

Yabu, S., T. Y. Tanaka, and N. Oshima, 2017: Development of a multi-species aerosol-radiation scheme in JMA's global model. *WGNE blue book, Res. Activ. Atmos. Oceanic Modell*, **47**, 4–15.

Yamaguchi, H., J. Chiba, Y. Ichikawa, and T. Takakura, 2023: Hindcast verification of JMA's GEPS for one-month prediction with a globally expanded two-tiered sea surface temperature approach. *WGNE Res. Activ. Earth Sys. Modell.*, **53**, 6.15–6.16.

Yamaguchi, H., M. Higaki, and M. Kyouda, 2014: Upgrade of JMA's One-Week Ensemble Prediction System. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **44**, 6.17–6.18.

Yamaguchi, H., M. Ikegami, T. Iwahira, K. Ochi, R. Sekiguchi, and T. Takakura, 2021: Upgrade of JMA's Global Ensemble Prediction System. *WGNE Res. Activ. Earth Sys. Modell.*, **51**, 6.13–6.14.

Yamaguchi, H., M. Ikegami, K. Ochi, Y. Ota, R. Sekiguchi, and T. Takakura, 2020: Upgrade of JMA's Global Ensemble Prediction System. *WGNE Res. Activ. Earth Sys. Modell.*, **50**, 6.17–6.18.

Yamaguchi, H., Y. Adachi, S. Hirahara, Y. Ichikawa, T. Iwahira, Y. Kuroki, C. Matsukawa, R. Nagasawa, K. Ochi, R. Sekiguchi, T. Takakura, M. Ujiie, and H. Yonehara, 2022: Upgrade of JMA's Global Ensemble Prediction System. *WGNE Res. Activ. Earth Sys. Modell.*, **52**, 6.9–6.10.

Yamaguchi, M., R. Sakai, M. Kyoda, T. Komori, and T. Kadowaki, 2009: Typhoon Ensemble Prediction System developed at the Japan Meteorological Agency. *Mon. Wea. Rev.*, **137**, 2592–2604.

Yamashita, K., 2016: ASSIMILATION OF HIMAWARI-8 ATMOSPHERIC MOTION VECTORS INTO THE NUMERICAL WEATHER PREDICTION SYSTEMS OF JAPAN METEOROLOGICAL AGENCY. *Proceedings for the 13th International Winds Workshop*, Monterey, California, USA, 27 June - 1 July 2016.

Yamazaki, D. and T. Yamaguchi F. O'Loughlin J. C. Neal C. C. Sampson S. Kanae P. D. Bates D. Ikeshima, R. Tawatari, 2017: A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, **44**, 5844–5853.

Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. OLoughlin, J. C. Neal, C. C. Sampson, S. Kanae, and P. D. Bate, 2017: A high-accuracy map of global terrain elevations. *Geophys. Res. Lett.*, **44**, 5844–5853.

Yanagino, K. and S. Takada, 1995: Quantitative Analysis and Application to Weather Prediction by Neural Networks. *Technical Report of IEICE*, **NC95-37**, 63–70, (in Japanese).

Yano, S, Johan C. Winterwerp, A Tai, and T. Saita, 2010: Numerical Experiments on Features of Nonlinear Tide and Its Influences on Sediment Transport in the Ariake Sea and the Yatsushiro Sea. *J. of JSCE, Ser. B2 (Coastal Engineering)*, **66**, 341–345, (in Japanese).

Yokota, S., T. Banno, M. Oigawa, G. Akimoto, K. Kawano, and Y. Ikuta, 2022: Implementation of hybrid 3DVar in JMA's local analysis. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **52**, 01.19–01.20.

Yokota, S., T. Kadowaki, M. Oda, and Y. Ota, 2021: Improving ensemble-based background error covariances of the hybrid 4DVar in JMA's global analysis. *WGNE Res. Activ. Earth System Modell.*, **51**, 01.27–01.28.

Yonehara, H. and M. Ujiie, 2011: A stochastic physics scheme for model uncertainties in the JMA one-week ensemble prediction system. *WGNE blue book, Res. Activ. Atmos. Oceanic Modell*, **41**, 6–9.

Yonehara, H., M. Ujiie, T. Kanehama, R. Sekiguchi, and Y. Hayashi, 2014: Upgrade of JMA's Operational NWP Global Model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, 6.19–6.20.

Yonehara, H., T. Tokuhiro, R. Nagasawa, M. Ujiie, A. Shimokobe, M. Nakagawa, R. Sekiguchi, T. Kanehama, H. Sato, and K. Saitou, 2017: Upgrade of parameterization schemes in JMA's operational global NWP model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, 4.17–4.18.

Yonehara, H., R. Sekiguchi, T. Kanehama, K. Saitou, T. Kinami, A. Shimokobe, D. Hotta, R.Nagasawa, H. Sato, M. Ujiie, T. Kadowaki, S. Yabu, K. Yamada, M. Nakagawa, and T. Tokuhiro, 2018: Upgrade of JMA's operational global NWP system. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, 6.15–6.16.

Yonehara, H., C. Matsukawa, T. Nabetani, T. Kanehama, T. Tokuhiro, K. Yamada, R. Nagasawa, Y. Adachi, and R. Sekiguchi, 2020: Upgrade of JMA's operational global model. *WGNE blue book, Res. Activ. Earth Sys. Modell*, **50**, 6–19.

Yonehara, H., Y. Kuroki, M. Ujiie, C. Matsukawa, T. Kanehama, R.Nagasawa, K. Ochi, M.Higuchi, Y. Ichikawa, R. Sekiguchi, and S. Hirahara, 2023: Upgrade of JMA's operational global Numerical Weather Prediction system. *Res. Activ. Earth. Sys. Modell,*, 6.15–6.16.

Yoshida, M., M. Kikuchi, T. M. Nagao, H. Murakami, T. Nomaki, and A. Higurashi, 2018: Common Retrieval of Aerosol Properties for Imaging Satellite Sensors. *J. Meteor. Soc. Japan*, **96B**, 193–209.

Yoshida, M., K. Yumimoto, T. M. Nagao, T. Y. Tanaka, M. Kikuchi, and H. Murakami, 2021: Satellite retrieval of aerosol combined with assimilated forecast. *Atmos. Chem. Phys.*, **21**, 1797–1813.

Yoshimura, H., 2002: Development of a Semi-Implicit Semi-Lagrangian Global Model using Double Fourier Series. *The 4th International Workshop on Next Generation Climate Models for Advanced High Performance Computing Facilities.*, NCAR, Boulder, Colorado, 12 - 14 March 2002.

Yoshimura, H. and T. Matsumura, 2003: A Semi-Lagrangian Scheme Conservative in the Vertical Direction. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.*, **33**, 03.19–03.20.

Yoshimura, H. and T. Matsumura, 2004: Semi-Lagrangian Toitsu model. *Report of Numerical Prediction Division (Suuchiyohouka Houkoku Bessatsu Houkoku)*, **50**, 51–60, (in Japanese).

Yoshimura, H. and S. Yukimoto, 2008: Development of a Simple Coupler (Scup) for Earth System Modeling. *Pap. Meteor. Geophys.*, **59**, 19–29.

Yukimoto, S., H. Yoshimura, M. Hosaka, T. Sakami, H. Tsujino, M. Hirabara, T. Y. Tanaka, M. Deushi, A. Obata, H. Nakano, Y. Adachi, E Shindo, S. Yabu, T. Ose, and A. Kitoh, 2011: Meteorological Research Institute-Earth Systemi Model Version 1 (MRI-ESM1) –Model Description–. *Techical Reports of the Meteorological Research Institute*, **64**, 1–96, doi: 10.11483/mritechrepo.64.

Yukimoto, S., Y. Adachi, M. Hosaka, T. Sakami, H. Yoshimura, M. Hirabara, T. Y. Tanaka, E. Shindo, H. Tsujino, M. Deushi, R. Mizuta, S. Yabu, A. Obata, H. Nakano, T. Koshiro, T. Ose, and A. Kitoh, 2012: A New Global Climate Model of the Meteorological Research Institute: MRI-CGCM3 — Model Description and Basic Performance —. *J. Meteor. Soc. Japan*, **90A**, 23–64.

Yukimoto, S., H. Kawai, T. Koshiro, N. Oshima, K. Yoshida, S. Urakawa, H. Tsujino, M. Deushi, T. Tanaka, M. Hosaka, S. Yabu, H. Yoshimura, E. Shindo, R. Mizuta, A. Obata, Y. Adachi, and M. Ishii, 2019: The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *J. Meteor. Soc. Japan*, **97**, 931–965.

Yumimoto, K., T. Y. Tanaka, N. Oshima, and T. Maki, 2017: JRAero: the Japanese Reanalysis for Aerosol v1.0. *Geosci. Model Dev.*, **10**, 3225–3253.

Yumimoto, K., T. Y. Tanaka, M. Yoshida, M. Kikuchi, T. M. Nagao, H. Murakami, and T. Maki, 2018: Assimilation and Forecasting Experiment for Heavy Siberian Wildfire Smoke in May 2016 with Himawari-8 Aerosol Optical Thickness. *J. Meteor. Soc. Japan*, **96B**, 133–149.

Zeng, X. and A. Beljaars, 2005: A prognostic scheme of sea surface skin temperature for modeling and data assimilation. *Geophys. Res. Lett.*, **32**, L14 605.

Zeng, X. and A. Wang, 2007: Consistent Parameterization of Roughness Length and Displacement Height for Sparse and Dense Canopies in Land Models. *J. Hydrometeorol.*, **8**, 730–737.

Zeng, X., M. Zhao, and R. E. Dickinson, 1998: Intercomparison of Bulk Aerodynamic Algorithms for the Computation of Sea Surface Fluxes Using TOGA COARE and TAO Data. *J. Climate*, **11**, 2628–2644.

Zhang, H., T. Nakajima, G. Shi, T. Suzuki, and R. Imasu, 2003: An optimal approach to overlapping bands with correlated k distribution method and its application to radiative calculations. *J. Geophys. Res.*, **108**, D20, 4641.

Zhong, W. and J. D. Haigh, 1995: Improved broadband emissivity parameterization for water vapor cooling rate calculations. *J. Atmos. Sci.*, **52**, 124–138.

Zsótér, E., 2006: Recent developments in extreme weather forecasting. *ECMWF Newsletter*, **107**, 8–17.

Zweng, M. M., J. R. Reagan, D. Seidov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, A. V. Mishonov, O. K. Baranova, K. Weathers, C. R. Paver, and I. Smolyar, 2018: *World Ocean Atlas 2018, Volume 2: Salinity*. NOAA Atlas NESDIS 82, 50 pp.

List of Authors

SATO Yoshiaki	Preface
ARANAMI Kohei	Section 3.5 , Section 3.7 , Section 3.8
ASAI Hiroaki	Section 5.3
FUKUURA Takashi	Section 5.5
HASEGAWA Hiroshi	Section 5.1
HAYASHI Yosuke	Section 3.12
HIGUCHI Mayuko	Section 2.2 , Section 2.3
HIRA Sakie	Section 5.6
HIRABARA Mikitoshi	Section 5.3
HIRAHARA Youichi	Subsection 4.8.2.1, Subsection 4.8.2.4
IOKA Yusuke	Section 2.2 , Section 2.3
ISHII Kyosuke	Section 2.2 , Section 2.3
ITO Shingo	Subsection 4.8.1
KADOWAKI Takashi	Section 2.5 (except Subsection 2.5.7) , Section 2.8 , Section 2.9
KAMADA Akane	Section 3.9 , Section 3.10
KANEHAMA Takafumi	Section 3.2
KAWANO Kohei	Section 2.6 , Section 2.7 , Section 3.6 , Section 4.5
KAZUMORI Masahiro	Section 2.4 , Section 2.10
KOBAYASHI Shinya	Section 2.11
KOIZUMI Kou	Section 2.8
KOMORI Takuya	Section 3.4
KUBO Yutaro	Section 3.4
KURAUCHI Toshihiro	Section 2.1, Section 3.1
KUSABIRAKI Hiroshi	Section 3.5
MOGI Atsushi	Section 2.1 , Section 3.1
MURAKAMI Yasutaka	Section 2.2 , Section 2.3
MURATA Hidehiko	Section 2.2 , Section 2.3 , Subsection 2.5.7
NAKAMURA Yuki	Section 2.2 , Section 2.3
NARITA Masami	Section 4.1
NISHIMOTO Shusuke	Section 3.5
NISHIMURA Akio	Subsection 4.8.2.3 , Section 5.1 , Subsection 5.2.3
NONAKA Kenichi	Section 2.2 , Section 2.3
OCHI Kenta	Section 2.9
OTA Yoichiro	Section 3.3
OTA Yukinari	Section 3.11 , Appendix
SAKAMOTO Kei	Subsection 5.2.1, Subsection 5.2.2
SAKAMOTO Masami	Section 3.8
SAKANASHI Takanori	Section 4.4
SAKURAGI Tomoaki	Section 5.4
SAKURAI Toshiyuki	Section 5.1
SANNOHE Yosuke	Section 1.1, Section 1.2, Section 1.3, Section 4.2, Section 4.3
SAWADA Masahiro	Section 3.5
SUGANO Junpei	Section 5.5
SUMITOMO Masashi	Subsection 4.8.2.2 , Subsection 4.8.2.5
TAKAKURA Toshinari	Section 3.9
TOGUCHI Ryo	Section 2.2 , Section 2.3
UEDA Manabu	Section 4.9
UJIIE Masashi	Section 2.9 , Section 3.2
YAMADA Yuji	Section 4.6 , Section 4.7
YAMAGUCHI Haruki	Section 3.3 , Section 3.9 , Section 3.10
YONEHARA Hitoshi	Section 3.2
YOSHIDA Takuma	Section 3.4

Outline of the Operational Numerical Weather Prediction at the Japan Meteorological Agency

25 January 2024
Japan Meteorological Agency
3-6-9 Toranomon, Minato City, Tokyo 105-8431, Japan