# ガイダンスの解説

平成 30 年 3 月 March 2018

気 象 庁 予 報 部

今回の数値予報課報告・別冊では、初めてガイダンス特集を企画した。

数値予報ガイダンスは、数値予報の出力結果を元に、数値予報モデルの分解能より細かい 地形の効果や数値予報モデルの系統的な誤差等を主に統計的に補正する手法として 1977 年 に予報業務に導入された。これにより、予報のための定量的かつ客観的な判断材料が現場の 予報官に届けられ予報作業を強力に支援するようになり、また、1980 年の降水確率予報の 実用化に道を開くことになった。

1991年には、ガイダンス等数値予報応用技術開発の推進のため、数値予報課にアプリケーション班が設置され、その後ガイダンスの精度向上と予報要素の拡充が一段と進んだ。これにより、地方天気分布予報・地域時系列予報の運用開始(1996年)や市町村を対象とする気象警報・注意報の発表(2010年)など予報業務の高度化に大きく寄与した。

この 40 年あまりの間にガイダンスは数値予報や電子計算機の進化に足並みを揃えるように、社会の変化に応じて変革する予報業務を支えながら飛躍的な進歩を遂げてきた。今日では、ガイダンスは予報業務に欠くことのできない支援資料となっている。

そこで、こうしたガイダンスに関するこれまでの変遷や理論的基盤、技術の動向、運用中 ガイダンスの解説、今後の展望などを本誌にまとめておくことにした。

第1章ではガイダンスの概要及び利用上の留意点を述べ、当庁のガイダンスの歴史や海外気象機関での状況について紹介する。第2章ではガイダンスに用いられている様々な手法について基礎的な理論と留意点を解説する。第3章ではガイダンス設計時の留意事項や運用にあたって必要な対応等について述べる。第4章では各ガイダンスの仕様、作成手法、精度、予測事例及び利用上の留意点を個別具体的に解説する。最後に第5章では今後のガイダンスの見通しについて、気象庁での取り組みや海外機関の動向を紹介する。

近年は、多様な産業界において、IoT や人工知能(AI) ビッグデータ等に関する技術の発展を背景に、意思決定や業務プロセスの改善、生産性の向上に資するべく、気象データの利活用をいっそう促進することが求められている。数値予報の応用技術に焦点を合わせた本誌の刊行がそうした活動の一助にもなれば幸いである。

<sup>\*</sup> 松村 崇行

## ガイダンスの解説

## 目 次

## はじめに

第1章	ガイダンスの概要	1
1.1	序論	1
1.2	ガイダンス概論	3
1.3	気象庁のガイダンスの変遷	9
1.4	海外のガイダンス	15
第2章	ガイダンスの作成技術	19
2.1	はじめに	
2.2	手法の分類・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	20
2.3	ガイダンスに用いる統計の基礎・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	24
2.4	線形重回帰⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯	35
2.5	ロジスティック回帰	43
2.6	ニューラルネットワーク・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	48
2.7	カルマンフィルタ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	62
2.8	診断手法	75
2.9	頻度バイアス補正	78
2.10	) その他の統計手法	81
第3章	ガイダンスの開発と運用	86
3.1	はじめに	86
3.2	ガイダンス設計時の検討事項・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	86
3.3	モデル更新への対応・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	88
3.4	観測所の移設等の影響と対応・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	91
第4章	ガイダンスの解説	94
4.1	ガイダンス一覧	94
4.2	降水ガイダンス	95
4.3	降雪ガイダンス	119
4.4	気温ガイダンス・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	132
4.5	風ガイダンス・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	144
4.6	天気ガイダンス	154
4.7	発雷確率ガイダンス・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	165
4.8	湿度ガイダンス・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	177
4.9	視程ガイダンス・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	185
4.10	) 雲ガイダンス	197
	乱気流指数	
	<sup>2</sup> 着氷指数······	
4.13	3 積乱雲量・雲頂高度・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	226

第5章	今後の展望	233
5.1	ガイダンスの今後	233
5.2	ディープニューラルネットワーク・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	237
5.3	ガイダンスグループ	241
付録 A	数値予報課報告・別冊で用いた表記と統計的検証に用いる代表的な指標	242
付録 B	電子計算室報告、同別冊、数値予報課報告・別冊 発行履歴	247

## 第1章 ガイダンスの概要

## 1.1 序論 1

ガイダンスは、数値予報の地上気温や降水量などの予測値を補正してその誤差を軽減したり、数値予報が出力していない天気や発雷確率などを作成することによって、予報作業を支援するものである。ガイダンスの歴史は古く、1970年頃より米国気象局で運用が開始され、その後多くの国の気象機関でも運用が開始され、数値予報の応用プロダクトとして予報業務に利用されてきた。気象庁でも1970年代から開発を始め、1977年にガイダンスの配信を開始した。その後改良や新たなガイダンスの開発を継続的に行い、2018年現在まで約40年間の歴史を持つ。

本報告では、このガイダンスの基礎技術、実際の開発・運用・仕様等について総合的な解説を行う。この40年間でガイダンスの総合的な解説として、「短期間予報用ガイダンスの解説」(気象庁予報部 1986)と「天気予報ガイダンスの解説」(気象庁予報部 1991)が発刊されている。これ以降は、数値予報研修テキスト及び予報技術研修テキスト<sup>2</sup>において、ガイダンスの開発・改良について随時解説を行ってきたが、まとまった解説書は発行されなかった。本報告は 27 年ぶりの総合的な解説書となる。

この 27 年の間には、ガイダンスの作成方法に大きな変化があった。27 年前は全てのガイダンスが線形重回帰(第 2.4 節)で作成されて予測式は固定されていたが、1996 年からはカルマンフィルタ(第 2.7 節)及びニューラルネットワーク(第 2.6 節)といった逐次学習型のガイダンスが使われるようになり、予測式が日々変化するようになった。また、確率型ガイダンスに適したロジスティック回帰(第 2.5 節)も使われるようになった。さらに、数値予報モデルの高性能化に伴い、視程分布予想のように、数値予報モデルの予測値をベースとした簡易な予測手法である診断手法(第 2.8 節)も使われるようになった。

予報業務の高度化に伴い、ガイダンスの仕様も変化してきた。例えば、予報作業支援システムを使った市町村を対象とする警報・注意報 (予報部予報課 2011)の作成を支援するために、5 km や 20 km 格子の格子形式のガイダンスが作成され、最大降水量や最大降雪量など防災気象情報の作成を支援する新たなガイダンスが運用されている。また、27年前には航空気象予報用のガイダンスはなかったが、2018年現在は飛行場予報及び空域予報の高度化 (予報部予報課航空予報室 2006)を支援するために、風、視程、雲のガイダンスや乱気流指数などのガイダンスが運用されている。

1 高田 伸一

このような予報業務の高度化とそれを支援するガイダンスの高度化と多様化により、予報作業も変化してきている。予報作業支援システム (予報部予報課 2005)の運用を開始した 2004 年以降、予報作業は、多くのガイダンスや数値予報資料をベースにし、予報作業支援システム上で多種多様な予報を短時間で作成・発表するといった作業形態となっている。このシステムにおいて、ガイダンスは発表予報のベースとなる重要な資料となっている。詳細な降水量、降雪量、最大風速、最高・最低気温等の量的予報の作成にはガイダンスは必要不可欠であり、今後もその重要性はますます高まるであろう。

近年の変化として、ガイダンスは発表予報と並行して配信され、インターネット等を通じて社会に提供されるようになってきた。これは日本に限らず海外でも同様である。ガイダンスは、時間的・空間的に細かい予測を迅速に提供できるため、今後も予報官の修正を経ることなく社会に直接提供される気象予測資料としての利用が増えてゆくと考えられる。

また、気象庁内外を問わず数値予報を使った応用処理が行われることが多くなってきており、各ガイダンスの内容について詳細な解説を求められることも多くなってきている。それ以外にも、長年にわたって開発・運用してきたガイダンスの取組みについて、そのノウハウ、経緯及び課題等を総括的に解説しておくことは、数値予報の応用処理及び今後のガイダンス開発者にとっても有用と考えられる。

このような状況を鑑み、本報告は、ガイダンスの作成技術、ガイダンスの開発と運用、各ガイダンスの具体的な作成方法を重点的に解説する。前述した過去2度の総合的な解説では、気象庁の予報作業上でのガイダンスの具体的な利用方法について多くのページを割いて解説されていたが、本書ではその点には触れず、各ガイダンスの利用上の留意点のみを記述している。過去の解説はガイダンスが予報作業のベースになりつつある時代のものであり、予報作業でどうガイダンスを利用してゆくかが重要であったと考えられる。しかし、ガイダンスの運用が開始されてから約40年経ち、ガイダンスは既に予報作業に浸透していることから、今回の解説では、数値予報の応用技術としての解説に多くのページを割いた。

まず第1章では、ガイダンスの作成概念や役割、全般的な利用上の留意点を示し、ガイダンスの基本的な事項を解説する。その後に気象庁のガイダンスの歴史と海外気象機関で運用されているガイダンスについて解説する。前述したようにガイダンスは海外気象機関でも作成されているが、統計手法や作成要素などが各国で異なっている。気象庁のガイダンスと海外気象機関のそれとの違いを確認しておくことは、今後のガイ

<sup>2</sup> 平成 20 年度までは量的予報研修テキスト

ついて. 測候時報、78、1-18.

ダンスの方向性を考えるためにも重要である。

第2章では、ガイダンスに用いられている各種統計 手法について、その基礎的な理論と手法を使用する際 の留意事項を解説する。前述したように、数値予報を 用いた応用処理を行うことが多くなってきており、数 値予報データを扱う観点から見た統計及び機械学習の 基礎知識を提供する。

第3章では、ガイダンスを開発・運用する際に重要な3つの点として、ガイダンスの設計、モデル更新への対応及び観測所の移動の影響と対応を解説する。ガイダンスを開発・運用してきた経験に基いた実用的な事項を記述する。

第4章では、2018年現在運用しているガイダンスの一覧を示した上で、各ガイダンスの詳細な解説を行う。前述したように数値予報研修テキスト等において、ガイダンスの開発・改良について随時解説を行ってきたが、紙面の制限により主に変更した部分についての解説となっており、変更と関係のない部分についての解説となっており、変更と関係のない部分については、過去の数値予報研修テキスト等や過去の文献が引用されていた。このため、個々のガイダンスが複数年度のテキストにわたって解説されているといった問題があり、ガイダンスの利用者にはわかりづらい状況であった。よって、ここでは、個々のガイダンスの仕様、具体的な作成方法、統計検証及び事例検証を通じたガイダンスの利用上の留意点を全て述べ、なるべく他の資料を参照する必要のないようにした。

最後の第5章では、今後の展望として、2018年現在計画されているガイダンスの開発・改良に加え、ガイダンスの将来的な方向性を示す。また、最近話題のディープニューラルネットワークについて、海外での気象分野における取り組みを紹介する。そして最後に、平成29年度から始まった、気象庁技術開発推進本部の実況監視・予測技術開発部会の「ガイダンスグループ」の活動も合わせて紹介する。

なお、本報告は天気予報・防災気象情報用、航空気象予報用のガイダンスに限って解説する。これらのガイダンス以外にも週間予報用、季節予報用のガイダンスが気象庁で運用されているが、本報告では紙面の関係もあり、上記の2つのガイダンスに限っている。

#### 参考文献

気象庁予報部, 1986: 短期間予報用ガイダンスの解説. 予報技術資料第 35 号.

気象庁予報部, 1991: 天気予報ガイダンスの解説. 予報 技術資料第 41 号.

予報部予報課, 2005: 予報作業支援システムと新しい予報作業形態について. 測候時報, **72**, 53-63.

予報部予報課航空予報室, 2006: 新しい航空気象予報業 務について. 測候時報, **73**, 23-49.

予報部予報課, 2011: 市町村を対象とした警報・注意報 発表のための新しい予報作業支援システムの整備に

#### 1.2 ガイダンス概論 1

本節では、ガイダンス全般に関する基礎的な事項について解説する。まず、ガイダンスの概念とガイダンスの必要性を述べ、その作成方法の概要を解説する。その後、ガイダンスによる数値予報の誤差の軽減について解説し、軽減できない誤差を含めてガイダンスの限界や課題をまとめる。なお、2018年現在運用しているガイダンスの一覧や手法などについては、第4.1節に示しており、適宜そちらを参照して頂きたい。

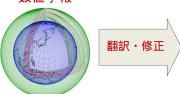
#### 1.2.1 ガイダンスの概念

ガイダンスは、数値予報の応用プロダクトの1つである。図1.2.1 に数値予報の後処理としてのガイダンスの概念図を示した。数値予報は格子点の予測値を提供しているが、これらは単に膨大な数字が入ったデータである。例えば2018 年時点での気象庁の全球モデル(GSM)、局地モデル(LFM)では一億以上の格子点がある。その格子点で風、気温、湿度など多くの予イがある。その格子点で風、気温、湿度など多くのデータが存在している。これらの膨大な予測値のうち予報に大きな影響のある要素を可視化した図を予報官が閲覧し、今後の気象状況の推移が検討される。しかし、気温、降水量などの発表すべき多くの予報を迅速に行うためには、図だけでなく、これらの量的な予測値を直接示す客観的な予測資料が必要であり、そのためにガイダンスが作成・配信されている。

ガイダンスは、米国で 1970 年頃から数値予報を用いた客観的予測資料として運用され始めた (Glahn and Lowry 1972)。「ガイダンス」という名称が使われ始めたのがいつ頃かは不明だが、1970 年代の米国の論文にはこの客観的予測資料について "provided as guidance to forecasters at NWC" (Klein and Glahn 1974)、 "used as guidance by forecasters at NWS" (Zurndorfer et al. 1979) と記述されている。つまり、予報官をガイド(手引き)する資料であることから「ガイダンス」という名称が使われたと考えられ、2018 年現在も "guidance" という名称が使われている(例えば Hamill et al. (2017) )。日本でも 1970 年代後半からガイダンスの運用が開始され、保科ほか (1982) では「ガイダンス」という名称が使われており、2018 年現在まで継続して使用されている。

一方、ガイダンスの解説 (気象庁予報部 1986) では、ガイダンスを「数値予報を用いた客観的・統計的翻訳」とも記述している。また、米国では "MOS (Model Output Statistics )" または "MOS guidance" と統計手法を使ったことを明記した呼び方が一般的に使われている <sup>2</sup>。最

#### 数值予報



計算結果は未来の大気状態に対応する様々な数値の羅列(1億以上の格子点:ビックデータ)

#### ガイダンス

晴·曇り·雨 禁 最高気温27 発雷確率30%...

天気予報、防災情報等 に必要な情報

図 1.2.1 ガイダンスの概念図

近では国際的には 主に statistical post-processing または単に post-processing と呼ばれることが多く、数値予報の統計的な後処理であることがわかる名称となっている。これらはガイダンスの本質を表した名称と言える。いずれにしろ、ここで解説するガイダンスは、数値予報を統計的に処理した客観的予測資料であり<sup>3</sup>、予報の作成をガイドするために作成されている資料である。

## 1.2.2 数値予報の後処理としてのガイダンスの必要性

前項では、ガイダンスは数値予報の統計的な後処理をした資料であることを述べた。ここでは、その数値 予報に後処理を加える必要性を2点述べる。

1つ目の必要性は、発雷確率、乱気流及び視程など、数値予報が直接予測していないが、天気予報、警報・注意報、飛行場予報などの発表に必要な気象要素を作成することである。これらの予測を行うため、数値予報データからそれぞれの要素に関係の深いものを取り出し、加工し、予測を行う。この例として、次項で発雷確率の作成方法を示す。

2 つ目の必要性は、気温や降水量など数値予報でも 予測しているが、その予測値を補正し、より精度を向 上させることである。この例として、次項で気温の作 成方法を示す。数値予報の予測値には、場所、時刻、予 報時間、季節及び気象場に応じて、それぞれ異なった 誤差を持っている。この誤差を統計的に補正すること で、予測精度を向上させることができる。この誤差の 補正については、第 1.2.4 項で詳しく解説する。

## 1.2.3 ガイダンスの作成方法(概論)

図 1.2.2 にガイダンスの作成方法の概念図を示した。 ガイダンスは数値予報データと予測対象である実況の 観測データを用い、統計手法によって予測式を作る。そ のために、まず過去の数値予報データと観測データを 用意する。用意した数値予報データから観測と因果関 係が深い要素を抽出して作成したものを説明変数、観

<sup>1</sup> 高田 伸一

 $<sup>^2</sup>$  米国では、2018 年現在、モデルを統計処理していない資料をモデルガイダンスとし、MOS ガイダンスと使い分けている。

MOS 方式でなく数値予報を統計処理しない PPM(Perfect Prognosis Method) 方式のガイダンスもあるが、ほとんどのガイダンスは MOS 方式なので、ここでは MOS 方式を基本として解説する。PPM については第 2.2 節を参照。

測データを目的変数として、予測式を作成する。説明 変数は気象及び統計の知識を基に作成するが、どのよ うな説明変数を用いるかがガイダンスの予測精度に大 きく影響する。図では、発雷確率ガイダンスの例を示 しており、過去の数値予報から CAPE (対流有効位置 エネルギー) SSI(ショワルター安定指数)など発雷 に関係する変数を作成し、それと発雷の有無の観測を 元にロジスティック回帰(第2.5節参照)で予測式を作 成する。そして、その予測式に最新初期時刻の数値予 報データを入力することによって、発雷確率ガイダン スを算出する。過去データは多いほど良いが、通常同 じ予測特性を持った数値予報データを用意できる期間 は限られており、実際に運用している発雷確率ガイダ ンスでも2~3年分の数値予報データを用意して予測式 を作成している。米国気象局においても、ガイダンス の開発には通常2年間以上の数値予報データが必要と されている (Carter 1986)。

上記の発雷確率ガイダンスのように過去数年分のデータを利用し一括して学習する型のガイダンスは、一括学習型、バッチ学習型もしくは係数固定型と呼ばれる。一方、気温ガイダンスのように逐次的に学習し、係数が変化するガイダンスもある。このようなガイダンスは、逐次学習型、オンライン学習型もしくは係数更新型と呼ばれる。図 1.2.3 に気温ガイダンスを例として、逐次学習型ガイダンスの概念図を示した。図 1.2.2 の一括学習型と大きく異なるのは、最新の観測が入ったら、学習用のデータ(最新の観測と対応する数値予報から作成した説明変数)を作成し、予測式を評価して予測式の係数を更新することである。この予測式の更新は、通常、最新初期時刻の数値予報データに基づくガイダンスを作成する直前に行われる。

気象庁では、一括学習型ガイダンスに使われる学習 (統計)手法として、ロジスティック回帰及び線形重回 帰を、逐次学習型ガイダンスの作成に使われる手法として、カルマンフィルタ及びニューラルネットワーク

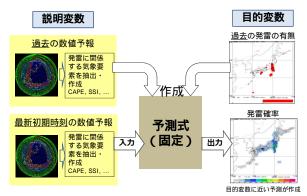


図 1.2.2 ガイダンス (一括学習型)の作成方法の概念図。発 雷確率ガイダンスの例で、予測式はロジスティック回帰で 作成される。

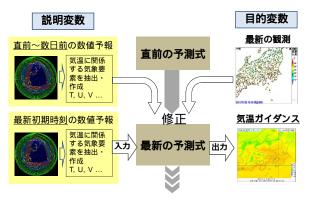


図 1.2.3 ガイダンス (逐次学習型)の作成方法の概念図。気温ガイダンスの例で、予測式の係数がカルマンフィルタで更新される。

を使っている 4。第 3.3 節でも解説するが、気象庁では、1996 年 3 月より、モデル更新に柔軟に対応できるように、初めて逐次学習型ガイダンスを気温や降水量などに導入し (田沢 1996)、その後視程や最小湿度など徐々に逐次学習型のガイダンスを拡げていった。よって、ガイダンスの多くは逐次学習型となっているが、乱気流、発雷確率など主に発現頻度の低い現象においては、逐次学習型では実況の後追い予測 5 となりやすい等の理由により、一括学習型を利用している。どのガイダンスにどのような手法が用いられているかは第 4.1 節「ガイダンス一覧」に示しているので、そちらを参照願いたい。

逐次学習型と一括学習型には表 1.2.1 に示す利点と 欠点がある。詳しくは第 2.2 節で述べるが、逐次学習 型は日々学習するため、一括学習型と違い、モデル更 新、観測所の移設・新設に自動的に対応できる。また、 逐次学習型は日々の学習により、季節変化に対応可能 である。ただし、逐次学習型は予測式が変化するため、 予測特性が日々変化して予報官が利用しにくいという 欠点や、予測式の係数が不適切に変化する場合があり 予測の安定性に劣るといった欠点がある。また、逐次 学習型には、係数の変化の大小を調整するパラメータ を設定する必要があること、その変化を監視する必要 があること、係数のデータファイルが変化するために 処理が複雑になること等開発・運用コストが高い。

## 1.2.4 ガイダンスの数値予報の誤差の補正

ここでは、数値予報の誤差とガイダンスによるその 誤差の軽減について解説する。数値予報の予測値は、 場所、予報対象時刻、予報時間、季節及び気象場に応 じて、それぞれ異なった系統誤差を持っている。ここ で言う系統誤差とは、場所毎や対象時刻毎で統計検証

 $<sup>^4</sup>$  逐次学習型でないニューラルネットワークもあり、1,3 時間最大降水量ガイダンスで使われている。

<sup>&</sup>lt;sup>5</sup> 例えば発雷確率では、雷発生後に確率が高くなるといった、 実況の後を追うように予測が変化すること。見逃しの後に空 振りとなるため、予測精度が低くなる。

表 1.2.1 逐次学習型ガイダンスと一括学習型ガイダンスの 利用面・運用面からみた優劣。優劣を不等号で示している。

モデル更新への対応	逐次 > 一括
季節変化への対応	逐次 > 一括
観測所の移設・新設への対応	逐次 > 一括
予測特性の日々の変化	逐次 < 一括
予測の安定性	逐次 < 一括
開発・運用コスト	逐次 < 一括

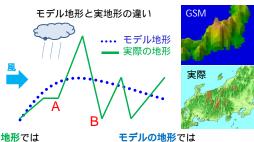
した際の期間平均の誤差という意味で使う。ガイダン スはこの系統誤差を統計的に補正することで、予測精 度を向上させることができる。数値予報の系統誤差の 要因として考えられるものとして、

- 1. 数値予報モデルの地形と実際の地形の違い
- 2. 数値予報モデルの海陸分布と実際の海陸分布の違い
- 3. 数値予報モデルの不完全性や空間代表性

がある。ガイダンスは統計処理をすることにより、こ れらの系統誤差を補正し、予測値の精度を向上させて いる。以下では3点それぞれについて、例を挙げて説 明する。

まず、「数値予報モデルの地形と実際の地形の違い」に より発生する系統誤差について、その模式図を図1.2.4 に示した。実地形に比べて数値予報モデルの地形はか なりなだらかで、この地形の違いにより、例えば図中 の A 地点では、数値予報の地形性の降水は実況より弱 くなる傾向がある。また、B地点では、実際の地形で は谷となっているが、数値予報モデルでは谷が表現さ れていないため、逆転層がない限り、数値予報の地上 気温は実況より低くなる。

地形に沿って吹く風も、モデルの地形と実際の地形 によって差が出る。図 1.2.5 に冬型の気圧配置の事例 におけるアメダス地点での GSM の風の予測、風ガイ ダンスの予測、アメダスの実況の風を示した。図に示 したように、GSM の地形は 20 km 解像度のため滑ら かであり、それを反映した風となっているが、実況で は実地形を反映した風となっている。風ガイダンスは この系統誤差を補正し、実際の風向に近づけているこ



実際の地形では

A:急斜面で地形性の降水多い A: 地形がなだらかで降水少ない B:標高低く、気温高い B:標高高く、気温低い

図 1.2.4 数値予報モデル地形と実地形の違いで生じる、系統 誤差の概念図。実際の地形の図は GTOPO30 ( 米国地質研 究所作成の 1 km 格子の地形データ) から作成している。

とがわかる。例えば富山県の東部(図中A)では、冬 型の気圧配置に地形の影響で南風になることが多いが、 GSM では滑らかな地形のために西~北西風を予測す る系統誤差がある。風ガイダンスでは、GSM の風を適 切に補正して、南風を予測している。

次に「数値予報モデルの海陸分布と実際の海陸分布 の違い」により発生する系統誤差について説明する。 海上では陸上に比べて、通常、風が強く、湿度が高く、 気温の日変化が小さい、という特性を持つ。数値予報 でもこれに対応して海格子と陸格子で大きく特性が変 化する。数値予報モデルと実際の海陸分布が一致して いればそれほど大きな問題はないが、沿岸では両者が 一致しないことが多く、これによって系統誤差が発生 する。図 1.2.5 (右上)の「GSM の地形」からは新潟県 粟島 (図中B) は島ではなく単に海となっていること がわかる。また、図は省略するが、メソモデル (MSM) でも粟島は海上となっている。ただし、実際にはアメ ダスは陸上にあり、気温の日変化が大きい。図 1.2.6 に 粟島における気温ガイダンスの実例を示した。赤線で 示した実況は日変化が大きいが、青・緑の細実線で示 した GSM 及び MSM の気温は日変化が小さい。青・緑 の太実線で示した GSM 及び MSM 気温ガイダンスは この日変化を大きくし、実況に近い気温を予測してい ることがわかる。

最後に「数値予報モデルの不完全性や空間代表性」 による系統誤差について説明する。数値予報モデルは、 複雑な大気の状態を様々な仮定のもとでシミュレート しており、このため様々な原因により系統誤差が発生 する。また、モデル解像度等により、格子平均である降 水量などは実況と同じにならないものもある。図 1.2.7 に GSM 及び MSM の降水量と両平均降水量ガイダンス のバイアススコアを示した。図からわかるように GSM の降水量予測は 10 mm/3h 以上ではバイアススコアは 1より低く、並~大雨の予測頻度が過小であることが わかる。また MSM の降水量予測は 3 mm/3h 以上で バイアススコアが1以上となっており、並~大雨の予 測頻度が過剰となっている。実線の降水量ガイダンス はモデルの降水量予測の頻度の偏りをなるべく実況に 近づけるよう (バイアススコアが1になるよう)に補 正していることがわかる。

これまでガイダンスが補正できるモデルの系統誤差 を説明してきたが、次にガイダンスが補正できない誤 差について解説する。なお、ここではモデルの誤差か ら系統誤差を引いたものを「ランダム誤差」として説 明を行う。この使い方は正確ではなく、後述するよう に系統誤差とここでいうランダム誤差の違いは明確で ないが、説明のためにこの用語を用いる。

はじめに、気温ガイダンスにおける、系統誤差補正 の実例を示す。図 1.2.8 は、長野県軽井沢の最高気温 ガイダンスの予測例を2年間で時系列表示したもので ある。気温ガイダンスはモデルの気温の誤差を目的変

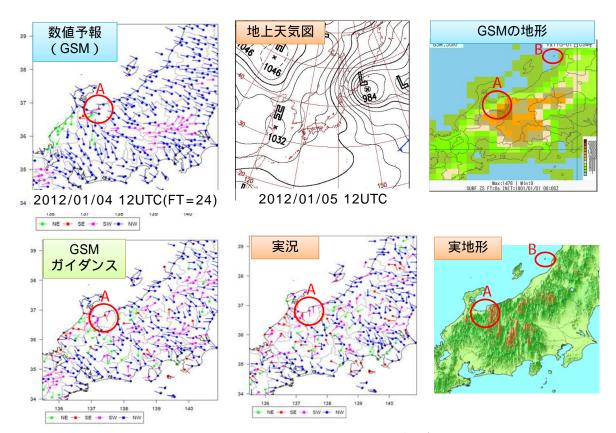


図 1.2.5 2012 年 1 月 4 日 12UTC の FT=24 の GSM の風(左上)と GSM 風ガイダンス(左下)、対象時刻の日本域地上天気図(中上)とアメダスの風の実況(中下)、GSM の地形(右上)と実地形(右下)。

数としているため、MSM の予測誤差(黒線)とガイダンスによる MSM の予測誤差の見積もり(MSM の予測値の補正量;赤線)を示している。両者の線が一致していれば、ガイダンスがモデルの誤差を正確に予測したことを示す。図からわかるように MSM の誤差は日々変動し、また緩やかな周期で季節変動している。ガイダンスの予測をみると、季節変化している MSM の平均的な誤差(系統誤差)を補正し、MSM の誤差を減じていることがわかる。一方、系統誤差を補正して残っている誤差(ランダム誤差; 黒線と赤線の差)は日々変動している。

このようなランダム誤差の要因は様々であるが、例えば、数値予報の前線や低気圧等の擾乱の位置ずれ、擾乱の発生・発達の外れによる誤差がある。これらは頻繁に起こるが、様々な誤差を生じるため、ガイダンスにとってはランダムに変化する誤差となり、基本的にはガイダンスで修正できない。擾乱の位置ずれによるランダム誤差はガイダンス全般に言えることだが、ほかにも、要素によってランダム誤差となる要因がある。例えば、気温の場合には、天気の外れ(日照や降水予測の外れ)、海陸風の入るタイミングの違い、放射冷却が起こるか否か、積雪面解析の誤りなどが考えられる。これらの誤差も、様々な要因で様々な誤差が生じ、ガイダンスにとってはランダムに変化する誤差となり、基本的にはガイダンスで修正できない。

前述したように、ここで使った系統誤差とは、場所 毎などで統計検証した際の期間平均の誤差のことであ る。このため、期間平均の仕方により、系統誤差は変 わり、その結果残差であるランダム誤差は変わる。通 常は地点(格子)毎に期間平均を取るが、ほかにも対 象時刻毎、予報時間毎、季節毎、風向毎、気圧配置毎な ど平均の仕方を変えると、系統誤差は変わり、ランダ ム誤差として見えている誤差も変化する。つまりガイ ダンスにおいて、補正できる誤差とできない誤差(系 統誤差とランダム誤差)というのは、明確に分離でき ているわけではない。

気温ガイダンスでは、地点毎、初期時刻毎、予報時間毎、季節毎に分けて予測式を作成している(第4.4節)。また、説明変数に西風成分等を取り入れて風向ごとの系統誤差を軽減している。しかし、それでも補正できていない誤差はある。例えば、気圧配置等のパターンで見られる誤差(例えば放射冷却が起こるパターンでの平均的な誤差)は、現状では補正できてはいないが、気圧配置等で層別化できれば誤差を軽減できるであろう。つまり、今までランダム誤差として扱っていた誤差を系統誤差として補正できることになる。ガイダンスは数値予報が持つ様々な要因による誤差を、学習のサンプル数を保ちながら系統付けてその誤差を見積り、補正している。それを適切に行えばガイダンスの精度向上に結びつく。

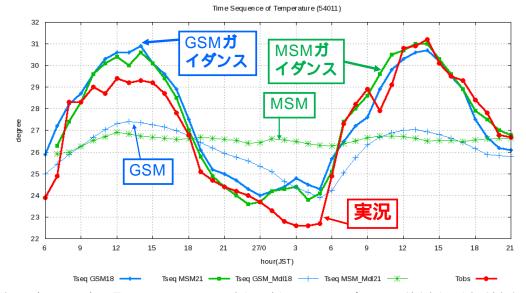


図 1.2.6 新潟県粟島の 2015 年 7 月 26~27 日にかけての気温の実況と GSM 及び MSM の地上気温予測、対応する GSM 及び MSM 気温ガイダンス。 GSM は 7 月 25 日 18UTC、MSM は 7 月 25 日 21UTC 初期値の予測。

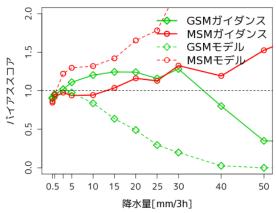


図 1.2.7 2016 年 6 月 ~ 8 月における GSM と MSM のモデル降水量のバイアススコア、及び GSM と MSM 平均降水量ガイダンスのバイアススコア。検証に利用した実況は20km 格子平均の解析雨量。 MSM 及び MSM ガイダンスは20km 格子に変換して検証している。緑破線は GSM の降水量、赤破線は MSM の降水量、緑実線は GSM 平均降水量ガイダンス、赤実線は MSM 平均降水量ガイダンスを表す。

## 1.2.5 ガイダンスの限界と課題

最後にこれまでの解説をもとにガイダンス全般の限界と課題を示す。これらの限界と課題はガイダンスを利用する際の留意点にもなる。なお、個別のガイダンスの課題は第4章をご覧頂きたい。

第 1.2.3 項では、ガイダンスは通常 2~3 年分のデータを使って作成していることを述べた。しかし、この 2~3 年で発生していない顕著現象に対しては、ガイダンスの予測式が適合していない可能性が高い。例えば、今まで起こったことのないような記録的な大雨においては、それより少ない降水事例で学習した予測式を外

挿しているだけであり、その量的予測は適切かどうかはわからない。また、学習したデータでも、事例数が少ない場合は、その予測式が適切でない可能性もある。 学習していない事例、学習のサンプル数が足りていない事例に対する予測は、信頼性が低く、誤差が大きくなる傾向がある。

逐次学習型ガイダンスの欠点として、不適切な係数の変化が起きる可能性があると述べた。実際、降水量ガイダンスにおいて、予測に比べて実況の降水量が非常に大きかった際に係数が大きく変化し、その後の予測精度が劣化した例がある。係数が大きく動かないようにカルマンフィルタ等のパラメタの設定を工夫するなどの対策は行っているが、完全に排除することはできない。

ガイダンスはモデルの擾乱の位置ずれ、擾乱の発生 や強度の外れなどの誤差を修正することができない。 これらの誤差については、擾乱について実況とモデル

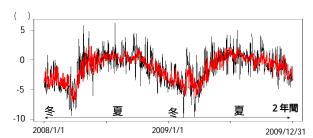


図 1.2.8 長野県軽井沢における MSM の最高気温の誤差(黒線)と MSM 最高気温ガイダンスによる MSM の補正量(赤線)。赤線と黒線が一致した所が MSM ガイダンスが MSM の誤差を適切に補正し、ガイダンスの予測の誤差がゼロになったことを示す。 MSM の最高気温は 30 分ごとの地上気温予測値から日中の最高値として算出している。期間は 2008~2009 年の 2 年間。

研究時報, 34, 239-276.

の予測を比較して補正を行う必要がある。ただし、このタイプのランダム誤差を軽減する方法として、他のモデルのガイダンスとの平均、または適切と考えたモデルのガイダンスを用いるといった方法がある。他のモデルのガイダンスと平均することによってランダム誤差が軽減できることは、第5.1節でもう少し詳しく説明する。

ガイダンスは、地点や予報対象時間などで予測式を 層別化し、また風などを説明変数として取り入れることによってモデルの誤差を軽減しており、さらに適切 な層別化や説明変数によってガイダンスの予測精度を 向上できることを述べた。現在、ガイダンスは気圧配 置や暖湿流などのパターンで生じる誤差を補正できていないが、予報現場では予報官のパターン認識を含ん だワークシートを作成し、ガイダンスを補正している がある。第5.1節でも述べるが、画像認識で有効性が 確認されているディープニューラルネットワークを使 うことにより、将来はガイダンスでこういったパター ンを平面的・立体的に認識して誤差を軽減できる可能 性がある。

### 参考文献

- Carter, G. M., 1986: Moving towards a more responsive statistical guidance system. Preprints, 11th Conf. on Weather Forecasting and Analysis, Kansas City, Amer. Meteor. Soc., 39–45.
- Glahn, H. R. and D. A. Lowry, 1972: The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.
- Hamill, T. M., E. Engle, D. Myrick, M. Peroutka, C. Finan, and M. Scheuerer, 2017: The U.S. National Blend of Models for Statistical Postprocessing of Probability of Precipitation and Deterministic Precipitation Amount. *Mon. Wea. Rev.*, 145, 3441–3463.
- 気象庁予報部, 1986: 短期間予報用ガイダンスの解説. 予報技術資料第 35 号.
- Klein, W. H. and H. R. Glahn, 1974: Forecasting Local Weather by Means of Model Output Statistics. Bull. Amer. Meteor. Soc., 55, 1217–1227.
- 田沢秀隆, 1996: 予報支援資料. 平成7年度量的予報研修テキスト、気象庁予報部, 43.
- Zurndorfer, E. A., J. R. Bocchieri, G. M. Carter, J. P. Dallavalle, D. B. Gilhousen, K. F. Hebenstreit, and D. J. Vercelli, 1979: Trends in Comparative Verification Scores for Guidance and Local Aviation/Public Weather forecasts. *Mon. Wea.* Rev., 107, 799–811.
- 保科正男, 安富裕二, 此田進, 1982: 短期間予報用最終 製品型予測資料(ガイダンス)の作成手法. 気象庁

#### 1.3 気象庁のガイダンスの変遷<sup>1</sup>

気象庁は 1970 年代からガイダンスの開発を進め、 1977 年にガイダンスの配信を開始した。2018 年現在 までの主なガイダンスの変遷を、利用する数値予報モ デルと合わせて図 1.3.1 に示した。図には主に新たな ガイダンスの配信開始、モデル更新への対応を記述し ており、ガイダンスの改良は省略している。

本節では、気象庁におけるガイダンスについて、線 形重回帰(MLR)による一括学習型ガイダンスのみを 使っていた時代である 1996 年 3 月以前と、カルマン フィルタ(KF)やニューラルネットワーク(NN)と いった逐次学習型ガイダンスが登場した1996年3月以 降の2期間に分けてそれぞれ解説する。前者について は、主に気象庁予報部 (1986) 及び気象庁予報部 (1991) に基づいて記述しているので、詳細はこれらを参照願 いたい。なお、2018年現在運用しているガイダンスに ついては第4.1節を参照願いたい。また、第4章で各 ガイダンスの変遷について記述されているので、そち らも参照願いたい。

#### 1.3.1 運用開始から 1996 年 3 月まで

米国気象局 (NWS) では、1972 年に MLR に基づく MOS ガイダンスの運用を開始した。気象庁では、これ を参考に、1970年代から独自のガイダンスの開発を開 始した。後述するように NWS との違いもみられるが、

統計手法として MLR を利用、 暖候期・寒候期(ま たは季節)に分けて予測式を作成、 MOS 方式(第 2.2節)で作成していたこと等、NWS のガイダンスの 作成方法と共通する部分が多く、これらの3点は1996 年3月まで変更がなかった。なお、NWSでは2018年 現在もこの3点が概ね維持されている。

第1.2節で述べたように、ガイダンスは数値予報を 入力とした統計処理であるため、以下では利用された モデルに分けてガイダンスの経過を説明する。

## (1) FLM-6L (1977年~1982年)

ガイダンスが最初に適用されたモデルは、FLM-6L (アジア・プリミティブ・ファインメッシュモデル、格子 間隔 150 km、鉛直 6 層) という現在と比べると低解像 度の数値予報モデルであった (保科ほか 1982)。まず降 水関係のガイダンスが開発され、1977年よりガイダン スとして配信が開始された。その後、徐々に要素を拡 充し、風ガイダンス、雷雨確率(後に発雷確率)ガイ ダンス、気温ガイダンス、日照確率(後に曇天率)ガ イダンス、雪関係のガイダンスが1982年にかけて運 用開始となった。表 1.3.1 に 1982 年時点での要素を示 した。

風及び気温ガイダンスは気象官署を対象として作成 されていたが、降水ガイダンスなどその他のガイダン

スは格子形式でモデルと同じ格子間隔であった(一部 地域形式)。この点は全てのガイダンスが地点形式で あった NWS のガイダンスとは大きく異なる $^2$ 。多くの 種類のガイダンスを格子形式で作成し、日本全体をカ バーする予測資料を作成することができたのは、1974 年から始められた全国900余のアメダスの運用開始が 大きく寄与していると思われる。つまり、モデルの格 子間隔である 150 km 格子内に観測点が多く存在し、格 子平均値の平均降水量や降水確率などの目的変数が算 出できたことによると考えられ、気象庁予報部 (1986) にもその旨が記述されている。

#### (2) FLM-10L/12L (1982年~1988年)

1982 年 3 月に利用するモデルが FLM-6L から FLM-10L (格子間隔 127 km、鉛直 10 層)に更新されたが、 FLM-10L の運用開始時は FLM-6L で作成した予測式 を継続して利用した。この際に格子間隔の違いを緩和 するために旧格子への内挿処理が行われた(保科ほか 1982)。1983年3月には鉛直層が12層に増強されたが (FLM-12L)、その際も旧予測式が継続利用された。そ の後 FLM-10L 及び FLM-12L の予測データが 2 年間 分蓄積されたことに伴い、新しいモデルに適合させる ために予測式の改訂作業が開始された。要素ごとに改 訂時期は異なるが、1985年12月までに全てのガイダ ンスの予測式が改訂された。この際に格子間隔はモデ ルの格子間隔に合わせ 150 km から 127 km に変更に なった。また、ガイダンスの予測式の更新に合わせ、ガ

表 1.3.1 1982 年時点でのガイダンスの一覧。保科ほか (1982) を基に作成。予測式の作成には全て FLM-6L (初期時刻は 00 と 12 UTC) が利用されており、作成方法は全て線形 重回帰。予報時間は概ね24時間先まで(1982年3月の FLM-10L への変更の際に 36 時間までに延長 》 格子間隔 は 150 km。内容に記述してある「地域確率」とは、ある 現象が格子・地域内のどこかで発生する確率、単に「確率」 レけ地占攻攻の投ス、地域で均でする。

こる地点領	とは地点帷率の格子・地域平均であることを示り。			
要素	作成形式	内容		
降水型	格子、地域	予測される降水が層状性か対 流性かの判別		
降水確率	格子	6, 12 時間に 1 mm 以上の雨 の降る確率		
平均雨量	格子、地域	6, 12, 24 時間に予測領域内に 降る平均降水量		
日照確率	格子	日照の確率		
雷雨確率	格子	雷の発生する地域確率(定義 は各管区で異なっていた)		
雪確率	格子	降水現象があると仮定した場合、それが雪になる確率		
最大風速	地点 (15の	日本時間でみた翌日の日最大		
以八瓜瓜	気象官署)	風速とその風向		
最高·最低 気温	地点(152 の 気象官署)	世界標準時でみた翌日の日最 高気温と最低気温		

 $<sup>^2</sup>$  米国では 2006 年より Gridded MOS として格子形式のガ イダンスも運用開始している。

<sup>1</sup> 高田 伸一

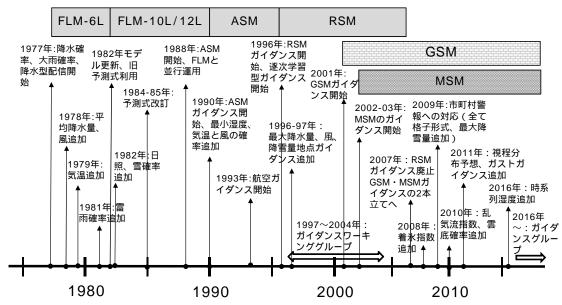


図 1.3.1 気象庁のガイダンスの変遷。上部の四角枠はガイダンスが利用している数値予報モデルを示している。

表 1.3.2 1986 年 2 月時点でのガイダンスの一覧。気象庁 予報部 (1986) を基に作成。予測式の作成には全て FLM-10/12L (初期時刻は 00 と 12 UTC) が利用されており、 作成方法は全て線形重回帰、予報時間は 36 時間、格子間 隔は 127 km。

要素	作成形式	内容
降水型	格子	予測される降水が層状性か対流 性かの判別
降水確率	格子	所定期間内に 1 mm 以上の雨 の降る確率
大雨確率	格子	3 時間内に格子内で少なくとも 1 地点で 30 mm 以上の大雨の 降る確率
平均雨量	格子	所定期間内に予測領域内に降る 平均降水量
曇天率	格子	日照率を 100 から引いた値 (地域内の平均値)
雷雨確率	格子	雷の発生する確率 ( 発雷はエ コー頂高度で定義 )
地域最大 降水量	地域	寒候期における細分地域内の最 大降水量
雪確率	地域	降水現象があると仮定した場 合、それが雪になる確率
大雪確率	地域	大雪の降る確率 (大雨確率と雪 確率の積)
最大風速	地点 (70 の 気象官署)	日本時間でみた翌日の日最大風 速とその風向
最高・最 低気温	地点(152の 気象官署)	世界標準時でみた翌日の日最高 気温と最低気温。晴天時と曇天 時の2つの予測値を出力。

イダンスの要素や予報時間の拡充が行われた。1986 年 2 月時点でのガイダンスの一覧を表 1.3.2 に示す。

#### (3) ASM (1988年~1995年)

1988 年 3 月に FLM-12L が ASM (アジアモデル、格子間隔 75 km、鉛直 16 層) に更新された。このモデル

更新の対応として、ガイダンスの精度を維持するために特別な対処が行われた。それは、更新前のモデルである FLM-12L をガイダンスのために運用を継続させたことである。ASM の予測資料が蓄積されるまで、新しい ASM と古い FLM-12L の二つのモデルが並行に運用された。これは、ガイダンスが予報作業に必要不可欠な資料であり、ガイダンスの予測精度を維持する必要があるとの判断によると思われる。しかし、ガイダンスが使うモデルと数値予報 FAX 図等が使っているモデルが異なること、二つのモデルを並行運用するコストが高いことなどの問題もあった。

前回の FLM-10L(12L) への更新と同様に、この時も ASM の予測データの蓄積を待って ASM に適合した予測式に改訂された。新旧モデルの並行運用を早く解消するために、ASM への変更から 2 年経った 1990 年 4 月から暖候期用、1990 年 10 月から寒候期用の予測式が改訂された<sup>3</sup>。この改訂の際に、モデルの高解像度化に合わせてガイダンスも高解像度化された。ただし、それまではモデル出力とガイダンスの格子間隔は同じだったが、この時からガイダンス独自の格子系である緯度経度 40 分間隔 ( $50\sim75$  km) に変更になった。これにより、モデルの格子系の変更に合わせてガイダンスの格子系を変える必要がなくなった 4。また、この予測式の改訂の際には、最小湿度、気温の確率、風確率等の追加があった一方、雪予測資料の地域最大降水量が廃止となった。表 1.3.3 に 1991 年 3 月時点でのガ

 $<sup>^3</sup>$  新しい予測式への改訂のための開発期間がないと考えるかもしれないが、当時は予測式を暖候期と寒候期に分けて作成していたため、例えば暖候期の  $^2$  年間の数値予報データが蓄積が終わった  $^1989$  年  $^10$  月から運用開始の  $^1990$  年  $^4$  月までに約半年間の開発期間がある。

<sup>4</sup> これ以降、ガイダンスの格子系は格子間隔の変化はあるものの、独自の格子系を継続している。

表 1.3.3 1991 年 3 月時点でのガイダンスの一覧。気象庁予報部 (1991) を基に作成。予測式の作成には全て ASM (初期時刻は 00 と 12 UTC) が利用されており、作成方法は全て線形重回帰、予報時間は 45 時間、格子間隔は 40 分。

エー・ストルン主		
要素	作成形式	内容
降水型	格子	予測される降水が層状性か対流 性かの判別
降水確率	格子	6 時間に 1 mm 以上の雨の降る 確率
大雨確率	格子	3 時間に 30 mm 以上の大雨の 降る確率
平均雨量	格子	所定期間内に予測領域内に降る 平均降水量
曇天率	格子	日照率を 100 から引いた値 (地域内の平均値)
雷雨確率	格子	雷の発生する確率(発雷はエ コー頂高度で定義)
気温	アメダス 地点	00 時(日本時)の気温
最高·最低 気温	アメダス 地点	日中の最高気温と朝の最低気温
気温確率	アメダス 地点	最高(最低)気温が基準値を上 回る(下回る)確率
風確率	アメダス 地点	最大風速が基準値を超える確率
最大風速	アメダス	最大風速が基準値を超えると仮
・風向	地点	定したときの最大風速・風向
最小湿度	気象官署	日中に出現すると仮定したとき の最小湿度

#### イダンスの一覧を示した。

なお、この期間には、ASM を利用した航空気象予報ガイダンス(以下、航空ガイダンス)が 1993 年 3 月より運用開始された。航空ガイダンスは一般の天気予報用ガイダンスと同じ MLR を用いて作成された。一般の天気予報用のガイダンスとの大きな違いは目的変数が航空定時・特別観測気象報 (METAR, SPECI) から算出されること、視程や雲などの航空気象予報に特有の要素がある点である。これにより、航空気象予報業務にも初めてガイダンスが導入された。表 1.3.4 に1993 年 3 月時点での航空ガイダンスの一覧を示した。

## 1.3.2 1996 年から 2018 年現在 (逐次学習型の登場)

1996年3月にガイダンスの作成方法について、大きな変化があった。前述したように、それまでのガイダンスは MLR を用い、いったん作成された予測式は固定されていた。第1.2 節で述べたように、この手法の欠点として、モデル更新への対応が難しく、実際にも更新時にガイダンス側で対応に苦慮していた。このため、予報部ではモデル更新に柔軟に対応できる KF や NN という逐次学習型の手法の開発を進め、1996年3月よりこれらの手法を使ったガイダンスの運用を開始した(田沢 1996)。ただし、この際に全てのガイダンスが逐次学習型に移行できたわけでなく、発雷確率ガイ

表 1.3.4 1993 年 3 月時点での航空ガイダンスの一覧。田崎・ 此田 (1994) を基に作成。対象空港は 68 空港、作成方法 は全て線形重回帰、予報時間は最大 45 時間までで、時間 間隔は視程と雲底が 3 時間毎、風は 6 時間毎。CIG は雲 量 5/8 以上の雲の最低雲層の雲底高度。IMC は計器飛行 気象状態(視程 5 km 未満、CIG が概ね 10,000 ft 未満)。

>1025 ( ):	元·1至 9 Kili
要素	内容
視程確率	視程が 10 km 未満になる確率
最小視程	視程が 10 km 未満になると仮定したとき の最小視程
雲底確率	CIG が 15,000 ft 未満になる確率
雲底高度	CIG が 15,000 ft 未満になると仮定したと きの CIG
IMC 確率	IMC となる確率
風確率	最大風速が基準値を超える確率
最大風速と その風向、 風向変動幅	最大風速が基準値を超えると仮定したとき の最大風速とその風向、風向変動幅
最高気温	明日の日中の最高気温
最低気温	朝の最低気温(00 UTC 初期値は明日、12 UTC 初期値は明後日)

ダンス(一括学習型 NN) 視程・雲ガイダンス(MLR) など一括学習型も継続された。これにより、2018 年現在に至るまで、逐次学習型のガイダンスが主体であるが、一括学習型のガイダンスも混在する状態が続いている。

この期間の特徴として、複数のモデルのガイダンスが運用され始めたこと、モデル・ガイダンスの統廃合が行われたことも特徴して挙げられる。よって 1996 年までのようにモデル毎に経過を記述することが難しいが、(1)ASM が RSM (領域モデル)に更新された際の対応、(2) その後 RSM が予報業務の主体となっていた期間、(3)RSM が GSM (全球モデル)に統合され GSMと MSM (メソモデル)が予報業務の主体となった時から 2018 年現在までの 3 つの期間に分けて解説する。

#### (1) ASM から RSM への更新 (1996年3月)

1996 年 3 月に ASM が RSM (領域モデル、格子間 隔 20 km、鉛直 36 層 ) に更新された。この際に逐次学 習型のガイダンスを導入することにより、それまでの ように新モデルの予測データの蓄積を待ってガイダン スの予測式を改訂することなく、RSM への更新と同時 に RSM ガイダンスの運用を開始した。ただし、ASM で開発した逐次学習型ガイダンスをモデル更新と共に 単純に入力データを RSM に差し替えたたわけでなく、 RSM の格子間隔と近く、それまで運用されていた JSM (日本域モデル、格子間隔 30 km)で2年間ほど逐次学 習してから RSM に移行された。これは、逐次学習と いっても、学習速度などのパラメータを設定するため には長期間の似た数値予報データが必要なこと、長期 間の統計検証を行ってから予報業務に導入する必要が あること、モデル更新の程度が大きいと逐次学習もす ぐに更新に対応できるわけではないためである。

この RSM ガイダンスは、当時開始された地方天気分布予報及び地域時系列予報を実現するため、降水や天気の格子形式のガイダンスは緯度 12 分×経度 15 分 (約 20 km ) と高解像度化された。この解像度は 2018 年現在も維持されている  $^5$ 。要素に関しても、地方天気分布予報の基礎資料である天気ガイダンス、地域時系列予報の基礎資料である 3 時間毎の時系列気温ガイダンスが新たに運用開始となった。一方、気温確率、風確率、降水型、最大風速のガイダンスは継続されなかった(その後 2002 年に最大風速は復活した)。また、降水量ガイダンスに初めて頻度バイアス補正(第 2.9 節)が導入された。これにより降水量ガイダンスの精度が大幅に改善した。表 1.3.5 に 1996 年 3 月時点での RSMガイダンスの一覧を示した。

なお、発雷確率は JSM のデータを用いて作成された一括学習型の NN を RSM に適用した。また、航空ガイダンスである視程及び雲底ガイダンスも、 JSM の予測データを用いて MLR の手法で予測式を作成し、これを RSM に適用する対策が行われた。ただし、 RSM と JSM ではモデルの物理過程や格子間隔が違うため、不安定な予測となる可能性が留意点として示されている(部内資料「新航空気象予報ガイダンスの解説(1996年 2 月)」より )。航空ガイダンスの気温及び風速  $^6$  については、 KF への手法変更が行われた。

表 1.3.5 1996 年 3 月時点でのガイダンスの一覧。田沢 (1996) を基に作成。予測式の作成には全て RSM (初期時刻は 00 と 12 UTC)が利用されている。発雷確率は一括学習型 NN、天気(雨雪判別)は PPM 方式(第 2.2 節参照)の一括学習型 NN。二次細分は当時の警報・注意報の細分区域(2018 年現在の市町村等をまとめた地域に相当。以下同様。)を示す。

131310 / 213	13136 7 631 7 6			
要素	形式	予報対象時間	手法	
天気	20 km	   FT=6~51、3 時間毎	NN	
平均降水量	格子	1.1—0,~91′ 9 14411111111111		
降水確率		FT=9~51、6 時間毎		
最高気温	アメダス	00 UTC:明日、12 UTC: 明日・明後日	KF	
最低気温	地点	00 UTC : 明日・明後日、 12 UTC : 明日		
時系列気温		FT=6~51、3 時間毎		
最小湿度	気象官署	00 UTC:明日、12 UTC: 明日・明後日	NN	
大雨確率	二次細分	   FT=6~51、3 時間毎	1111	
発雷確率				

(2) RSM が予報業務の主体となっていた期間 (1996 年~2007年)

この期間は、ガイダンスの改良と拡充が積極的に行われた期間であり、複数のモデルからガイダンスが作成された期間でもある。ここでは RSM ガイダンスの改良、GSM 及び  $MSM^7$  を使ったガイダンスの運用開始の 2 点について説明する。

RSM ガイダンスが運用された後も新しい要素の拡充と改良が頻繁に行われた。まず、1996年6月には大雨警報・注意報の作成を支援するために、最大降水量ガイダンスが開発された。その後、組織的・計画的にガイダンスの開発を行うため、1997年に「ガイダンスワーキンググループ」が予報部内で結成されてからは、このワーキンググループの下で、多くの取組みが行われ、3時間毎の風向風速を予測する風ガイダンスと12時間降雪量を予測する降雪量地点ガイダンスの配信が1997年12月から開始された。また、2000年にかけて、天気、降水、気温ガイダンス等の改良が次々と行われた。さらに、2001年3月の計算機更新の際には、天気、最小湿度、発雷確率、最大降水量ガイダンスが一挙に改良された(佐藤 2000a,b)。

航空ガイダンスでは視程ガイダンスと雲ガイダンスの改良が行われた。視程ガイダンスの目的変数は、それまで前 3 時間の 10 km になる確率とその時の視程だったが、前 3 時間の最小視程と最小視程が 5 km 未満になる確率に改良され、手法も MLR から KF に変更になった。雲ガイダンスの目的変数は、前 3 時間に CIG が 7000 ft になる確率とその場合の雲底高度及び上中下層雲量であったが、前 3 時間の最小 CIG 時の下から 3 層の雲量と雲底高度に改良され、手法も MLR から KF に変更になった。

この期間は RSM だけでなく、他のモデルを使った ガイダンスが運用開始となった。2001年3月より、明 後日予報を支援するために GSM を使った天気、降水 量、気温、風及び最小湿度ガイダンスが運用開始となっ た (佐藤 2000a,b; 海老原ほか 2000)。RSM の予報時間 は51時間であったため、GSM ガイダンスはその先の 51 時間から 72 または 75 時間後までをカバーした。ま た、防災気象情報の作成を支援する目的で、MSM 最 大風速ガイダンスが 2003 年に運用が開始された (松 本 2003)。その後 MSM の非静力学化 (2004 年) の際 に MSM 最大降水量ガイダンスの運用を開始した (木 村 2004)。2006 年には MSM の格子間隔が 10 km から 5 km への高解像度化された。このとき、MSM を使っ たガイダンスは、モデル更新に合わせて改良を行って いる。表 1.3.6 に 2006 年 3 月時点のガイダンスの一覧 を示した。

航空ガイダンスでも MSM が利用されるようになった。2002 年 8 月には MSM に基づく航空ガイダンスと

 $<sup>^5</sup>$  GSM ガイダンスの解像度はこの当時と同じだが、2018 年現在、MSM ガイダンスは約  $5~\mathrm{km}$  格子である。

 $<sup>^6</sup>$  風向は RSM の予測値を使っており、統計処理はされていなかった。なお、その後 1997 年に風向風速共に KF の手法に移行された。

 $<sup>^7</sup>$  MSM は 1998 年 3 月から試験運用、2001 年 3 月から本運用。

表 1.3.6 2006 年 3 月時点でのガイダンスの一覧。予報対象時間の括弧内は時間間隔。発雷確率、雪水比は一括学習型 NN。最大降水量は PPM 方式の一括学習型 NN。

要素	形式	予報対象時間	モデル	手法
定時風		$FT=3\sim51(3h)$	RSM	
72.19724		$FT = 57 \sim 75(6h)$	GSM	
しまける 最大風速		$FT=3\sim51(3h)$	RSM	
投入へ込む	アメダ	$FT=3\sim33(3h)$	MSM	KF
· 時系列気温	ス地点	$FT=6\sim51(3h)$	RSM	
F-12/12/2020		$FT = 54 \sim 72(6h)$	GSM	
最高・最		明日 (, 明後日)	RSM	
低気温		明後日 (,3 日後)	GSM	
降雪量地点		$FT=24\sim48(12h)$	RSM	
最小湿度	気象官	明日 (, 明後日)	RSM	
取小业皮	署	明後日 (,3 日後)	GSM	NN
天気		$FT=6\sim51(3h)$	RSM	
/XX		$FT=57\sim75(6h)$	GSM	
	20 km	$FT=6\sim51(3h)$	RSM	
平均降水量	格子	$FT = 57 \sim 75(6h)$	GSM	170
	143	$FT=3\sim18(3h)$	MSM	KF
降水確率		$FT=9\sim51(6h)$	RSM	
		$FT=57\sim75(6h)$	GSM	
雪水比		$FT=6\sim51(3h)$	RSM	
発雷確率		$FT=6\sim51(3h)$	RSM	
大雨確率		$FT=6\sim51(3h)$	RSM	NN
1, 3 時間最	二次細	FT=6~51(3h)	RSM	ININ
大降水量	分	FT=3~18(3h)	MSM	
24 時間最大 降水量		FT=27~51(24h)	MSM	

して、15 時間先までの視程、雲、風のガイダンスが運 用開始となった (岩倉ほか 2002)。 それまであった RSM を使った航空ガイダンスは、予報時間が45時間で、長 距離飛行用飛行場予報 (TAF-L、有効期間  $9\sim27$  時間) 及び短距離飛行用飛行場予報(TAF-S、有効期間 0~9 時間)の両方を支援していたが、MSM を使ったガイ ダンスは予報時間が 15 時間であったため、TAF-S の 支援に特化したガイダンスであった。このため、RSM を使った航空ガイダンスは TAF-L ガイダンス、MSM を使ったガイダンスは TAF-S ガイダンスと呼ばれた。 予測要素は前3時間の最小視程、視程が5km未満と なる確率、前3時間の最小CIG時の下から3層の雲量 と雲底高度、3時間毎の風、1時間毎の時系列気温及び 最高最低気温であった。その後、MSM の非静力学化 (2004年) MSM の高解像度化(2006年)のモデル更 新に対応した(高田 2004; 山田 2004; 高田ほか 2005)。

## (3) GSM と MSM が予報業務の主体となった期間 (2007年~2018年現在)

2007年は GSM 高解像度 ( 20 km 格子 ) 化及び MSM の予報時間延長という大きな変化があった。 2007年 5月に MSM の予報時間が 33 時間に延長されたことに伴い、MSM から作成されていた TAF-S ガイダンスの予報時間を延長した (阿部 2007)。 これは、2007年 11

表 1.3.7 2007 年 5 月時点での航空ガイダンスの一覧。高田 (2007) を基に作成。全て MSM から作成されており、予 報時間は、00,06,12,18 UTC 初期値は 15 時間、03,09, 15,21 UTC 初期値は 33 時間。LR はロジスティック回帰。

15, 21 010 物類直は 55 時間。上にはロンスティック目標。			
要素	内容	手法	
視程	前 1 時間の最小・平均視程	KF	
視程確率	視程が 5/1.6 km 未満になる確率	KF	
雲	前 $1$ 時間の最小 $\mathrm{CIG}$ 時の $3$ 層の雲量と雲底高度	NN	
定時風	1 時間毎の風向・風速	KF	
最大風速	前 1 時間の最大風速時の風向風 速	KF	
天気	前 1 時間の卓越天気と降水強度	お天気 マップ	
発雷確率	前 3 時間に空港周辺で発雷する 確率	LR	
時系列気温	1 時間毎の気温	KF	
最高·最低気 温	当日・明日の最高・最低気温	KF	

月に予定されていた RSM 廃止に対応するためであるが、天気予報ガイダンスが RSM から GSM を利用したガイダンスに移行したのと異なり、航空ガイダンスは、GSM を使ったガイダンスは作成されず、MSM のみから作成されるガイダンスに一本化された。これにより、TAF-L 及び TAF-S の両方の作成を MSM から作成されるガイダンスで支援することとなり、2018 年現在まで続いている  $^8$ 。この際に、確率のガイダンスを除く風、気温、視程及び雲の予報時間間隔は全て 1 時間になり、また定時風、平均視程、視程が 1600~m 未満となる確率、最大風速の新たな要素が追加された。表 1.3.7 に 2007 年 5 月時点での航空ガイダンスの一覧を示した。

2007 年 11 月に GSM が 20 km 格子に高解像度化するのに合わせて、それまで RSM が担っていた天気予報の支援を GSM が担うこととなり、RSM が廃止された。これにより、GSM ガイダンスに取り込まれる形で RSM ガイダンスが廃止となった (阿部 2007)。また、MSM ガイダンスについても、MSM の予報時間延長を受けて降水確率、定時風、時系列気温、最高最低気温及び発雷確率が新たに追加された。表 1.3.8 に 2007 年 11 月時点でのガイダンス (航空用除く)の一覧を示した。

RSM 廃止後は、2009 年 9 月には、市町村を対象とした警報・注意報の発表のために更新された予報作業支援システムにガイダンスを適合させるため、それまで二次細分単位(2018 年現在の市町村等をまとめた地域に相当)だった  $GSM \cdot MSM$  最大降水量ガイダンス及び GSM 発雷確率ガイダンスが格子単位となった。その際に、それまでは GSM 発雷確率の手法は NN であったが、MSM 発雷確率と同じロジスティック回帰となった。また、MSM ガイダンス(格子)の格子間隔は

 $<sup>^8</sup>$  なお、TAF-L, TAF-S も 2008 年に一つの TAF に統合された。

表 1.3.8 2007 年 11 月時点でのガイダンス ( 航空用除く ) の一覧。阿部 (2007) を基に作成。予報対象時間の括弧内は時間間隔。MSM は 00,06,12,18 UTC 初期値は 15 時間予報、03,09,15,21 UTC 初期値は 33 時間予報のため、ガイダンスも初期値によって予報時間が異なる。発雷確率と最大降水量は一括学習型 NN。24 時間最大降水量のFT=24 ( MSM は FT=21 ) までは実況の解析雨量も利用。

要素	形式	予報対象時間	モデル	手法	
定時風		FT=3~84(3h)	GSM	I	
最大風速		$FT=3\sim84(3h)$	RSM		
		$FT=3\sim15/33(3h)$	MSM		
時系列気	アメ ダス 地点	$FT=3\sim75(3h)$	GSM	KF	
温		$FT=1\sim15/33(1h)$	MSM		
最高・最 低気温		明日、 明後日 (、3 日後)	GSM		
		当日、明日	MSM		
降雪量地 点		FT=24~72(12h)	GSM		
最小湿度	気象官署	明日、明後日 (、3 日後)	GSM	NN	
天気		FT=6~75(3h, FT57 から 6h)	GSM		
平均降水 量	20 km 格子	FT=6~75(3h, FT57 から 6h)	GSM		
里		$FT=3\sim15/33(3h)$	MSM	KF	
		$FT=9\sim 81(6h)$	GSM		
降水確率		$FT=9\sim15/6\sim30$ (6h)	MSM		
雪水比		$FT=6\sim84(3h)$	GSM	NN	
発雷確率		$FT=3\sim15/33(3h)$	MSM	LR	
発雷確率		$FT=6\sim84 \text{ (3h)}$	GSM		
3 時間大雨 確率	二次細分	FT=6~84(3h)	GSM		
1, 3 時間		$FT=6\sim84(3h)$	GSM	NN	
最大降水 量		FT=3~15/33(3h)	MSM		
24 時間		$FT=6\sim51(24h)$	GSM		
最大降水量		FT=3~15/33 (24h)	MSM		

5 km と詳細化された。その他、雪水比ガイダンスは最大降雪量ガイダンスに変更され、大雨確率ガイダンスは廃止された。最近では、視程分布予想の追加(MSM 2011 年, GSM 2013 年)、GSM・MSM24 時間最大降水量ガイダンスの手法を NN から MLR に変更(2013年)、時系列湿度ガイダンスの追加(2016年)、MSM 降雪量地点ガイダンスの追加や仕様拡充(2016年)等が実施された。また、降水、気温、風、発雷確率ガイダンスなど定期的に改良が行われ、予測精度が改善している。

また、航空ガイダンスも、定期的に改良と拡充が行われている。主な改善としては、視程ガイダンスの改良(2010年)、雲ガイダンスの改良(2008年)、風ガイダンスの改良(2010, 2013年)、雲底確率ガイダンスの追加(2010年)が行われた。また、TAFの予報時間延長に対応す

るため、航空ガイダンスの予報時間も 39 時間に延長された (2013 年)。 さらに、空域予報を支援する資料として、積乱雲量の改良と雲頂高度の追加 (2007, 2009年)、着氷指数の追加 (2008 年)、乱気流指数の追加と改良 (2010, 2014年)が行われ、空域予報もガイダンスを基本とした予報作業に変わってきた。

気象庁のガイダンスは以上の経過を辿って 2018 年現在に至っている。2018 年現在のガイダンスの一覧については、第 4.1 節を参照願いたい。

#### 参考文献

- 阿部世史之, 2007: ガイダンス類の解説. 平成 19 年度 数値予報研修テキスト、気象庁予報部, 48-49.
- 海老原智, 酒井喜敏, 山田芳則, 2000: 天気予報ガイダンス. 平成 12 年度数値予報研修テキスト、気象庁予報部, 28-35.
- 岩倉晋, 松本逸平, 大林正典, 2002: 航空ガイダンス. 平成 14 年度数値予報研修テキスト、気象庁予報部, 46-51.
- 木村陽一, 2004: 最大降水量ガイダンス. 平成 16 年度 数値予報研修テキスト、気象庁予報部, 48-50.
- 気象庁予報部, 1986: 短期間予報用ガイダンスの解説. 予報技術資料第 35 号.
- 気象庁予報部, 1991: 天気予報ガイダンスの解説. 予報 技術資料第 41 号.
- 松本逸平, 2003: RSM 及び MSM 最大風速ガイダンス. 平成 15 年度数値予報研修テキスト、気象庁予報部, 43-46.
- 佐藤豊, 2000a: RSM 天気ガイダンスの改善と GSM 天 気ガイダンスの開発. 平成 12 年度量的予報研修テキスト, 気象庁予報部, 63-70.
- 佐藤豊, 2000b: RSM 最小湿度ガイダンスの改善と GSM 最小湿度ガイダンスの開発. 平成 12 年度量 的予報研修テキスト, 気象庁予報部, 71-78.
- 高田伸一, 2004: TAF-S 視程ガイダンス. 平成 16 年度 数値予報研修テキスト、気象庁予報部, 53-55.
- 高田伸一, 工藤淳, 新美和造, 2005: 航空気象予報. 平成17年度数値予報研修テキスト、気象庁予報部, 49-62.
- 高田伸一, 2007: 航空気象予報ガイダンス. 平成 19 年 度数値予報研修テキスト、気象庁予報部, 87-93.
- 田崎允一, 此田進, 1994: 航空気象ガイダンスの利用例 について. 航空気象ノート、気象庁航空気象管理官, 48, 39-49.
- 田沢秀隆, 1996: 予報支援資料. 平成7年度量的予報研修テキスト、気象庁予報部, 43.
- 山田雄二, 2004: TAF-S 雲ガイダンス. 平成 16 年度数値予報研修テキスト、気象庁予報部, 56-57.
- 保科正男, 安富裕二, 此田進, 1982: 短期間予報用最終 製品型予測資料(ガイダンス)の作成手法. 気象庁 研究時報, **34**, 239-276.

#### 1.4 海外のガイダンス<sup>1</sup>

ここでは海外の気象機関が 2018 年現在運用または 試験運用しているガイダンスについて簡単に紹介する。まず、1972 年から既に MOS ガイダンスの運用を開始し、気象庁が参考にした米国気象局のガイダンスについて述べる。次に、多くのガイダンスを運用し、かつ先進的な統合型ガイダンス<sup>2</sup> の開発を早くから進めているドイツ気象局のガイダンスについて紹介する。また、気象庁と同様に様々な機械学習の手法を用いているフランス気象局のガイダンスについて述べる。その後、英国気象局、カナダ気象局、香港天文台、韓国気象庁のガイダンスについて簡単に触れる。最後に気象庁と他国との相違点について簡単に触れる。最後に気象庁と他国との相違点について解説する。なお、ここでの情報は、各国の気象機関のホームページ (HP)、論文及び WMO の Technical Progress Report <sup>3</sup> 等からの情報を基にしている。

## 1.4.1 米国気象局のガイダンス

表 1.4.1 に米国気象局 (NWS) で運用中または試験 運用中のガイダンス (2018年現在)を示す。 NWS で は運用当初から全て地点形式のガイダンスを継続して いたが、2006年より地形を考慮して地点ガイダンスを 格子に配分した格子形式ガイダンス GMOS (Gridded MOS) の運用を開始している (Glahn et al. 2009)。2018 年現在、格子間隔は2.5 km で、かつ15 種類の予測要 素があり、気温や降水のみならず風や相対湿度等の格 子形式ガイダンスを運用しているのが特徴である。ま た、2012年5月よりアンサンブル予報のガイダンスで ある EKDMOS (Ensemble Kernel Density MOS) の 正式運用を開始している (Veenhuis 2012)。 NWS とカ ナダ気象局(CMC)の全球アンサンブル予報を合わせ た41メンバー毎の最高・最低気温、定時の気温と露点 温度の予測を作成し、それらの予測と予測誤差を用い てカーネル密度推定<sup>4</sup>を行い、かつスプレッド/スキル の関係式で補正して確率密度関数 (PDF) を作成してい る。さらに、最近では GMOS, EKDMOS, NWS の全 球モデル (GFS) と全球アンサンブルシステム (GEFS) 等の多くのガイダンスとモデル予測値をブレンドし、1 時間先から10日先までのシームレスなガイダンスを提 供する NBM (National Blend of Models) の開発を行っ ており、利用するモデルや予測要素を徐々に拡充して いる (Ruth et al. 2017)。このほかに、飛行場予報のた

めに、GFS ガイダンスを METAR 等の観測で毎時修正 し、1 時間から 25 時間先までの予測を提供する LAMP (Localized Aviation MOS Program) を 2005 年から運 用し、改良が続けられている (Ghirardelli and Glahn 2010)。

また、ガイダンスの予測要素も多く、3・6・12 時間の複数の時間幅の発雷確率、シビアな発雷の確率(雷があった場合の条件付確率)、着氷性降水の確率(降水があった場合の条件付確率)、視程障害種別(煙霧、霧、雪等のカテゴリ)など他国ではあまり見られない要素も作成されている。地点形式のガイダンスでは降水量や降雪量等は値そのものを直接予測するのではなく、幅を持ったカテゴリで作成されていることが特徴である。

ガイダンスに利用する統計手法としては、線形重回帰 (MLR) のみが使われているのが大きな特徴である。 KF 等の逐次学習型は利用されていないため、モデル 更新への即時対応が難しいが、同一手法を用いている ため、開発・運用コストは低いと考えられる。

## 1.4.2 ドイツ気象局のガイダンス

ドイツ気象局 (DWD) は地点形式ガイダンスと警報用の確率型格子形式ガイダンス、TAF 用ガイダンス、アンサンブル型ガイダンスなど数種類のガイダンスを運用している (Hess et al. 2015)。どれも DWD 及び欧州中期予報センター (ECMWF) のモデルの個々の数値予報から MOS 方式のガイダンスを作成し、それらを結合して1つのガイダンスとして作成されている。

地点形式ガイダンスは、DWD の全球モデル (ICON) から作成したガイダンス、ECMWF の全球モデル (IFS) から作成したガイダンスの 2 つがあり、かつ 2 つの予測結果を最適に結合した MOSMIX という統合型ガイダンスを作成している (Veira et al. 2017)。世界中の約5400 の地点について、気温、最高・最低気温、露点温度、風向風速、ガスト、ガスト確率、天気、6・24 時間降水量、6・24 時間降水確率、全雲量、上・中・下層雲量、日照時間などの多くの要素を作成している。統計手法は、確率にはロジスティック回帰 (LR) を利用し、ほかの要素には MLR を利用している。240 時間先まで 3 時間毎の予測が提供され、最新の観測でガイダンスを毎時修正している。

警報用の確率型格子形式ガイダンスは、自動警報を作成するための予測資料の一つとして開発されている。前述した IFS, ICON, DWD の領域アンサンブルモデル、ECMWF の全球アンサンブルモデルに LR を適用して各々の確率型ガイダンスを作成し、それを最適に組み合わせて時空間の解像度を統一し、ModelMIX という統合型ガイダンスが作成される (Reichert et al. 2017)。格子間隔は1km、予報時間は24及び75時間予報、要素はガスト、降雪量・長雨・吹雪・大雨・雷雨・霜・融雪・霧など

<sup>1</sup> 高田 伸一

<sup>&</sup>lt;sup>2</sup> 複数の数値予報結果 (アンサンブル予報含む) に対してそれぞれガイダンスを作成し、それらを統合して一つのガイダンスとするもの。

<sup>&</sup>lt;sup>3</sup> WMO の以下の URL から取得 http://www.wmo.int/pages/prog/www/DPFS/ProgressReports/2015/GDPFS-NWP-2015.html

<sup>4</sup> 確率変数(ここでは予測値)の確率密度を推定する手法の 一つ.

表 1.4.1 NWS のガイダンス。2017 年 10 月時点の NWS の HP (http://www.weather.gov/mdl/) の情報を基に作成。全ガイダンスが線形重回帰の手法を使っている。NWS ではガイダンスのことを単に MOS (Model Output Statistics) と呼ぶことも多いため、表ではその略称を使用する。表中のシビア確率は雷がある場合における条件付確率、着氷性降水の確率と雪の確率は降水がある場合における条件付確率である。NAEFS (North American Ensemble Forecast System) は NWS と CMC の全球アンサンブル予報 (41 メンバー)。なお、表以外にも Marine、River Basin、Coop Station の観測を使った GFS ガイダンス、NAM (North American Mesoscale Forecast System; NWS の領域モデル)を使った GMOS (試験運用中)等があるが、ここでは省略する。

種別	形式、利用モデル等	要素
GFS MOS (短期)	地点(2040 地点) FT=6~84 (3 時間毎) GFS(00・06・12・ 18 UTC 初期値)	定時気温、最高・最低気温、定時露点温度、雲量、定時風(風向・風速) 6・12 時間降水確率、6・12 時間降水量(カテゴリ) 24 時間降雪量(カテゴリ) 6・12 時間発雷確率、6・12 時間シビア発雷確率、6 時間着氷性降水確率、6 時間雪確率、降水種別、シーリング、視程(カテゴリ) 視程障害種別
GFS MOS (延長予報)	地点(2022 地点) FT=24~192 (12 時間毎) GFS (00・12 UTC 初期値)	定時気温、最高・最低気温、定時露点温度、雲量、最大風速、6・12 時間降水確率、6・12 時間降水量(カテゴリ) 12・24 時間発雷確 率、12 時間着氷性降水確率、12 時間雪確率、降水種別、24 時間降 雪量(カテゴリ)
NAM MOS	地点(1687 地点) FT=6~84 (3 時間毎) NAM(00・12 UTC 初期値)	定時気温、最高・最低気温、定時露点温度、雲量、定時風(風向・風速)、降水確率(6、12 時間)、6・12 時間降水量(カテゴリ)、24時間降雪量(カテゴリ)、発雷確率・条件付きシビア発雷確率(6、12 時間)、シーリング、視程(カテゴリ)、視程障害種別
GMOS	格子(2.5 km 格子) FT=6~192(3 時間毎) GFS (00・12 UTC 初期値)	定時気温、最高・最低気温、定時露点温度、雲量、定時風(風向・ 風速 ) ガスト、6・12 時間降水確率、6・12 時間降水量、24 時間 降雪量、相対湿度、3・6・12 時間発雷確率、卓越天気
EKDMOS	格子 ( 2.5 km 格子 )・地点 ( 2884 地点 )、FT=6~192 ( 3 時間毎 ) NAEFS (00・12 UTC 初期値)	気温、露点温度、最高・最低気温の PDF(確率密度関数)と各 41 メンバーの予測。試験運用中として、風速、降水量の 6 つの閾値を 超える 6・12 時間確率。
LAMP	格子(2.5 km 格子)・地点(1678 地点) FT=1~25(1 時間毎) GFS(00・06・12・18 UTC 初期 値)を毎時観測で更新	地点形式は気温、露点温度、風向・風速・ガスト、降水確率、6時間降水確率、降水有無、降水タイプ、発雷確率、発雷の強さ、対流雲の確率、対流雲の強さ、着氷性降水確率、雪確率、降水種別、全雲量、シーリング、降水があった場合のシーリング、視程、降水があった場合の視程、視程障害種別。格子形式は対流雲、発雷、気温、露点温度、シーリング、視程、全雲量、風向・風速
NBM (開発中)	格子(2.5 km 格子)、 FT=1~10days、GFS, GEFS, GMOS, EKDMOS など 15 の予 測を入力し毎時更新	定時気温、最高・最低気温、定時露点温度、12 時間降水確率、6 時間降水量、定時風(風向・風速 )、ガスト、雲量、シーリング、視程、1 時間降水確率インデックス、1 時間降水量、相対湿度、最大・最小相対湿度、卓越天気、降雪量

の確率がある。

#### 1.4.3 フランス気象局のガイダンス

表 1.4.2 にフランスのガイダンスの概要を示す。フランス気象局 (Météo France) は、KF を気温、湿度及び雲量に利用している。また、MLR, KF, DA ( Discrimination Analysis; 判別分析)、ランダムフォレスト(第2.10 節参照)等の多くの手法を採用している。

形式は一部格子形式があるが、基本的には地点形式で作成されている。また、利用モデルは基本的に全球モデルであるが、2015年時点で、自国の領域モデルAROMEで気温と湿度を試験的に作成している。

## 1.4.4 その他の国のガイダンス

#### (1) 英国気象局

英国気象局 (Met Office) のガイダンスに関する情報は HP 上でほとんど公開されていないため、論文等から収集した情報から概要のみを記述する。 Met Office では、数値予報のバイアスの補正のために、気温、湿度及び風に対して KF が適用されている。また、天気

カテゴリを決めるためにファジーロジック  $^5$  が利用されている。視程、降水種別等には気象庁と同様に診断手法が利用されている (Bennett et al. 2014; Haywood et al. 2008)。

Met Office でも、多くのモデルの予測結果を最適な 重みで結合し、予測値と確率を作成している(Optimal blended forecasts と呼ばれている)。地点形式について は、KF を使ってモデルバイアスを除去することによ り、モデル解像度の違いによる影響を軽減してから結 合している。降水量、風、気温の格子形式については、 粗い解像度のモデルの予測を結合する際に、地形を考 慮したダウンスケーリングを行って適切に結合が行わ れている(Bennett et al. 2014; Sheridan et al. 2010)。

## (2) カナダ気象局

カナダ気象局 (CMC) も MLR または DA を使っているが、大きな特徴として、UMOS (Upgradable MOS) という MLR 及び DA の予測式を月に 4 回アップグレー

<sup>&</sup>lt;sup>5</sup> ある現象の有無を条件式や関数等を用いて 0−1 の中間的な 値で出力する方法。

表 1.4.2 Météo France のガイダンス一覧。2015 年の WMO の Technical Progress Report を基に作成。MLR は線形重回帰、KF はカルマンフィルタ、DA は判別分析。ARPEGE は Météo France の全球モデル(4 日先、水平格子間隔は約 7.5–37 km の可変)、PEARP は Météo France のアンサンブル予報システム、IFS は ECMWF の全球モデル、MIX は ARPEGE と IFS のミックス、AROME は Météo France の領域モデル(水平格子間隔は 1.3 km )。

モデル	形式、予報時間、初期値	要素	手法
ARPEGE	地点(要素で可変) FT = 3~ 最 大 102 (3 時間毎) 00・06・12・	気温、日最高・最低気温、湿度、日最大・最 小湿度、全雲量、日射量	MLR+KF
	18 UTC 初期値	露点温度、風向・風速	MLR
		全雲量、視程確率、ガスト確率	DA
	格子、FT、初期値は地点に同じ	全雲量	ランダムフォレスト
PEARP	地点(要素で可変) FT = 3~ 最 大 102(3 時間毎) 06・18 UTC	気温、日最高・最低気温、風速	MLR ( 各メンバー )、 KF ( アンサンブル平 均 )
	初期値	風速	MLR と更正処理
		6 時間降水量	更正処理
IFS	地点(要素で可変) FT = 1~ 最 大 180(1, 3, 6 時間毎) 00・12 UTC 初期値	気温、日最高・最低気温、湿度、日最大・最 小湿度	MLR+KF
		全雲量	DA
		風向・風速	MLR
MIX	地点、FT = 1~ 最大 102 (1,3 時間毎)、00・06・09・12・18・21 UTC 初期値	気温、日最高・最低気温	MLR+KF
EPS	地点(要素で可変) FT = 1~ 最 大 180(1,3,6 時間毎)、00・12 UTC 初期値	気温、日最高・最低気温、風速、湿度、日最 大・最小湿度	MLR (各メンバー ) KF (アンサンブル平 均)
AROME (試験中)	地点、FT=1~最大42(1時間毎) 00・06・12・18 UTC 初期値	気温、日最高・最低気温、湿度、日最大・最 小湿度	勾配ブースティング+ 単純フィルタリング

ドできるシステムを導入している (Wilson and Vallée 2002)。モデル更新の場合には新旧モデルが混在したデータで回帰を行うが、旧モデルのデータを徐々に減らし、最後には新モデルのデータのみで回帰を行うといった手法である。これによりモデル更新に徐々に対応でき、かつ逐次学習型のような完全な変動型ではないため、より安定な予測が可能と思われる。

この UMOS は主に CMC の領域モデル (RDPS) のガイダンスに使われており、3 時間毎の気温と露点温度、風向・風速、6・12 時間降水確率、全雲量の4カテゴリ毎の確率、海上風が作成されている。全球モデル (GDPS) においては、気温のみに利用されている。

気温以外の GDPS ガイダンスとして降水確率、全雲量、風向風速が運用されている。これらは、PPM 方式を利用しているが、PPM 方式で問題となるモデルバイアス(第 2.2 節)を軽減するために、ある期間で計算したバイアスを減じる後処理を行っている。

#### (3) 香港天文台

香港天文台は、気温、湿度、露点温度、風、天気、雲量のガイダンスを作成しており、KF, MLR, LR が利用されている。また、ECMWFのIFSとENS、気象庁のGSM、NWSのGFS、Meso-NHM(気象庁の非静力学モデルを香港付近で利用)の予測を結合し、9日先までの統合的な予測であるOCF(Objective Consensus Forecast)を作成している。

#### (4) 韓国気象庁

韓国気象庁では、領域モデル (RDAPS) の予測を用いて、気温、最高・最低気温、全雲量、降水確率、降水種別、相対湿度、風向・風速が作成されている。気温には KF が用いられているが、その他は MLR が用いられている。全球モデル (GDAPS) では、最高・最低気温、降水種別、全雲量の3つが作成されている。最高・最低気温には KF が用いられている。

## 1.4.5 気象庁との比較

日本の気象庁と海外気象機関との大きな相違点は、 気象庁は多くの要素に KF 及び NN といった逐次学習 型のガイダンスを運用していることである。外国では Météo France では気温と湿度、雲量の3要素に KF を、 Met Office では気温と湿度、風に KF を利用している が、気象庁のように気温、湿度、風のみならず降水量、 降水確率、降雪量、視程等の多くの要素に逐次学習型 のガイダンスを利用している国はない。

利用モデルに関しては、NWS, DWD, Météo France など多数のガイダンスを運用している国では全球モデルを使ったガイダンスが主流である。領域モデルを使ったガイダンスは、他国でも見られるものの、気象庁は全球モデルのガイダンスよりむしろ領域モデルを使ったガイダンスの方が多く、他国と異なっている。

また、NWS, DWD, Météo France, Met Office などでは複数のモデルを使ったガイダンスを最適に結合し、均一な空間・時間解像度の統合型ガイダンスを開発・運

用している。特に DWD の ModelMIX では、自国及び ECMWF の決定論予報とアンサンブル予報を全て使って多くのガイダンスを作成した後に結合し、警報の作業に利用されている。また、NWS や Met Office でも顕著現象の予測に多くのモデルやガイダンスを結合しており、これらの統合型ガイダンスは、今後のガイダンスの主流となると考えられる。

予測要素の種類については、NWS が一番豊富であるが、気象庁は世界的にみれば、かなり多くの要素のガイダンスを作成している。また、2018 年現在は他国も格子形式のガイダンスを作成しているが、地点形式のガイダンスが主である。前述したように気象庁はガイダンスの運用開始当初から格子形式のガイダンスを運用しており、地点形式は少ないことが大きな特徴である。

## 参考文献

- Bennett, A., I. Pearman, S. Moseley, C. Jones, J. Canvin, S. Jackson, and B. Wright, 2014: Met Office post-processing of operational NWP forecasts. Conference WWOSC 2014, Montreal, August 2014, https://www.wmo.int/pages/prog/arep/wwrp/new/wwosc/documents/ACB-WWOSC2014-v2.pdf.
- Ghirardelli, J. E. and B. Glahn, 2010: The Meteorological Development Laboratory 's Aviation Weather Prediction System. Wea. Forecasting, 25, 1027–1051.
- Glahn, B., K. Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, 2009: The Gridding of MOS. Wea. Forecasting, 24, 520–529.
- Haywood, J., M. Bush, S. Abel, B. Claxton, H. Coe, J. Crosier, M. Harrison, B. Macpherson, M. Naylor, and S. Osborne, 2008: Prediction of visibility and aerosol within the operational Met Office Unified Model. II: Validation of model performance using observational data. Quart. J. Roy. Meteor. Soc., 134, 1817–1832.
- Hess, R., J. Glashof, and C. Primo, 2015: Calibration with MOS at DWD. *ECMWF* Calibration Meeting 12 February 2015, https://software.ecmwf.int/wiki/display/ FCST/Calibration+Meeting+12+February+2015.
- Reichert, B., R. Hess, T. Hirsch, P. James, C. Lenhart, J. Paller, W. Raatz, T. Schleinzer, G. Schroder, and R. Thehos, 2017: The operational warning decision support system AutoWARN for forecasting and nowcasting severe weather at DWD. EMS Annual Meeting: European Conference for Applied Meteorology and Climatology 2017, Dublin, Ireland, 14, EMS2017-849, http://meetingorganizer.

- copernicus.org/EMS2017/EMS2017-849.pdf.
- Ruth, D. P., D. T. Myrick, and M. Peroutka, 2017: Developing the National Blend of Models for National Weather Service Operations. 97th American Meteorological Society Annual Meeting, Seattle, WA., https://ams.confex.com/ams/97Annual/webprogram/Paper313249.html.
- Sheridan, P., S. Smith, A. Brown, and S. Vosper, 2010: A simple height-based correction for temperature downscaling in complex terrain. *Meteorol. Appl.*, **17**, 329–339.
- Veenhuis, B. A., 2012: Spread Calibration of Ensemble MOS Forecasts. *Mon. Wea. Rev.*, **141**, 2467–2482.
- Veira, A., R. Hess, S. Trepte, G. Vogt, and B. Reichert, 2017: Model Output Statistics for Point Forecasts at Deutscher Wetterdienst:Current Status and Future Developments. EMS Annual Meeting: European Conference for Applied Meteorology and Climatology 2017, Dublin, Ireland, 14, EMS2017-378-1, http://meetingorganizer.copernicus.org/EMS2017/EMS2017-378-1.pdf.
- Wilson, L. J. and M. Vallée, 2002: The Canadian Updateable Model Output Statistics (UMOS) System:
  Design and Development Tests. Wea. Forecasting, 17, 206–222.

## 第2章 ガイダンスの作成技術

## 2.1 はじめに<sup>1</sup>

第 1.2 節でも述べたように、ガイダンスでは目的変数と説明変数の関係式を何らかの方法で事前に作成しておき、それを最新初期時刻の数値予報モデルから算出した説明変数に適用することで予測値を作成する。この関係式を導く方法には様々なものがあるが、代表的な方法は本章で述べる線形重回帰やニューラルネットワーク、カルマンフィルタなどの統計手法である。

統計手法を用いてガイダンスを開発する場合、その 手法に関する理論を理解し、実装し、検証する必要が あるが、これには一般に多くの時間と労力を要する。ま た、このような処理を独自に実装した場合には、バグ や非効率的なプログラムが含まれる可能性もあり、開 発効率の低下や、将来的にその手法を維持・管理して いくためのコストの増加につながる。そこで近年のガ イダンスの開発においては、統計処理を行うためのソ フトウェアやパッケージといった統計ツールが用いら れている。統計ツールを利用すれば、新たな手法によ るガイダンスを比較的容易に導入できるため、開発に 掛かる時間を大幅に削減できるとともに、プログラム の可読性が向上し、バグが混入する可能性も低くなる。 近年では、統計処理を行うための様々なツールが無償 で利用できるようになっており、これらのツールはガ イダンスの開発を効率的に進める上で不可欠な存在と なっている。

一方で、統計ツールを利用すれば、その手法に関す る知識がなくても予測値を作成できるため、不適切な 手法を用いたり誤った入力データを与えたりすること や、開発の方向性を限定してしまうことにつながる。 竹内ほか (1992) は、手法の意味を理解せずに統計ツー ルを利用したとしても、その結果を正しく利用するこ とはできない上に、誤った判断をくだす危険があるこ とを述べている。久保 (2012) は中身を理解せずに統計 ツールを利用する手法をブラックボックス統計学と呼 び、誤った手法を無自覚に用いたり統計や検証の結果 を都合良く解釈したりするのではなく、データをよく 見て目的に沿った統計モデルを構築することの重要性 を述べている。開発においては、十分に注意していた としても、手法の妥当性よりも目先の予測精度を重視 してしまいがちである。しかしそれでは、数値予報モ デルの予測特性の変化などにより、予測精度が大きく 低下したり、不自然な予測値が出力されたりする可能 性が高まり、将来的な維持・管理コストの増加につな がる。ガイダンスの予測精度を長期的に維持・向上さ せるためには、統計ツールを有効に活用するとともに、 用いられている統計手法について正しく理解し、正し

く使用することが重要である。

本章では、ガイダンスの開発・管理に携わっている者、またはこれから携わる者を対象として、本稿執筆時(2018年現在)の気象庁のガイダンスに用いられている作成技術について、統計手法の理論を中心に述べる。統計手法を理解する上では統計や数学の高度な知識が必要となる場合もあるが、ここではガイダンスに利用する上で必要となる知識に絞って簡潔に述べることにしたい。より詳しく知りたい方は統計学や機械学習、時系列解析等に関する解説書を適宜参照していただきたい。

本章の構成は以下のとおりである。まず第 2.2 節では、ガイダンスに用いられている様々な手法を分類しながら、ガイダンスに用いられている手法を概観する。続いて第 2.3 節では、本章で述べる統計手法を理解する上で必要となる統計の基礎について述べる。第 2.4 節から第 2.9 節では、気象庁のガイダンスに用いられている手法と、各手法をガイダンスに利用する上での留意点を述べる。最後に第 2.10 節では、気象庁のガイダンスにはまだ利用されていないが、海外の気象機関では利用されているなど、今後新たな手法の導入を検討する上で候補となりうる手法について、その一部を紹介する。

#### 参考文献

久保拓弥, 2012: データ解析のための統計モデリング入 門 — 一般化線形モデル・階層ベイズモデル・MCMC. 岩波書店, 267 pp.

竹内啓, 矢島美寛, 廣津千尋, 藤野和建, 竹村彰通, 縄田和満, 松原望, 伏見正則, 1992: 自然科学の統計学. 東京大学出版会, 366 pp.

<sup>1</sup> 工藤 淳

#### 2.2 手法の分類 1

ガイダンスの関係式を作成する手法には様々なものがあり、目的変数や説明変数の特性やガイダンスの利用形態等を考慮して選択される。1つのガイダンスに使用される手法は1つとは限らず、多くのガイダンスでは複数の手法を組み合わせて予測を行っている。本節ではガイダンスの作成に利用されている手法を分類しながら各手法の特徴を述べる。

#### 2.2.1 統計手法と診断手法

過去データを統計的に処理することによって目的変 数と説明変数の関係式を作成する手法を統計手法と呼 ぶ。代表的な統計手法として線形重回帰やニューラル ネットワークなどが挙げられる。統計手法では、統計モ デル(目的変数と説明変数の関係は線形である等)を 事前に設定しておき、過去データを用いて統計モデル の最適な係数を決定する(係数を学習するともいう)。 これに対して診断手法では、目的変数と説明変数の関 係式は過去の研究や目的変数の定義などに基づいて決 定される。例えば第一圏界面の定義に従って圏界面気 圧の予測を行う手法や、消散係数の調査結果と視程の 定義に基づいて視程の予測を行う手法、パーセル法に 基づいて積乱雲の雲頂高度を予測する手法などが診断 手法に分類される。圏界面気圧の予測のような純粋な 診断手法では関係式の作成に過去データを必要としな いというメリットがあるが、実際には視程分布予想や 積乱雲頂高度など多くの診断手法型ガイダンスでは、 予測精度を向上させるために過去データも用いて係数 やパラメータの調整を行っている。

統計手法では平均的な精度が良くなるように係数が 学習されるため、稀な現象は予測されにくいという特 徴がある。これに対して診断手法では数値予報モデル の結果が直接的に予測に反映されるため、シャープな 予測を行うことになり、数値予報モデルが極端な値を 予測すれば、ガイダンスも極端な値を予測しうる、と いう特徴がある。ただし、予測精度としては統計手法 を用いた方が良くなりやすい。

## 2.2.2 MOS ≥ PPM

統計手法を用いて関係式を作成するとき、係数の学習に用いる説明変数が数値予報モデルから算出される場合を MOS (Model Output Statistics, Glahn and Lowry (1972)) といい、実況値や解析値から算出される場合を PPM<sup>2</sup>(Perfect Prognosis Method<sup>3</sup>) という。例えば発雷確率を予測する場合、MOS では過去の数値予報モデルによる降水量や安定度などの予測と過去の発雷の有無の実況を用いて係数を学習するのに対し、PPM で

は過去の実況の降水量や、高層気象観測や解析値から 算出した安定度などと過去の発雷の有無の実況を用い て係数を学習する。PPM は数値予報モデルの予測が 完璧で誤差がないと考えた場合の予測であるのに対し、 MOS では数値予報モデルに含まれる誤差(系統誤差、 ランダム誤差)も考慮される4ため、予測精度として は一般に MOS の方が高くなる。これは例えば、地上 気温と湿度の予測値を入力として降水の雨雪を判別す るような場合を考えればわかりやすい。仮に数値予報 モデルの地上気温に正バイアス(気温を高めに予測す る系統誤差)があり、湿度はバイアスが0であったと すると、PPM を用いた場合には実況と比べて雨と予測 する頻度が多くなる。これに対し MOS を用いた場合 には、地上気温の正バイアスを考慮して雨雪が判別さ れるため、予測精度は高くなる。また、PPM では数値 予報モデルが直接予測する要素(地上の風や気温、降 水量など)を予測することには意味をなさない。なぜ ならば、もし数値予報モデルの予測が完璧ならば統計 手法を用いる必要はなく、数値予報の出力そのものを 用いれば良いからである。このため、これらの要素を 予測するガイダンスは必然的に MOS を用いることに なる。

2018 年現在の気象庁のガイダンスの多くは予測精度 や予測対象の理由から MOS 方式を採用しているが、 PPM 方式には以下のようなメリットがあり、モデルが 精緻化されて誤差が小さくなってきた現在の数値予報 モデルにおいては PPM も考慮する価値のある手法で ある。まず、PPM で学習した係数は数値予報モデル の誤差特性に影響されないため、数値予報モデルが改 良されて予測誤差が減少していけば自然にガイダンス の精度も向上していくことが期待できる。これに対し て、MOS で学習した係数は数値予報モデルの誤差特 性に強く影響されるため、数値予報モデルの更新の影 響を受けやすく、数値予報モデルの誤差特性が大きく 変わる場合にはガイダンスの予測精度が大きく低下す ることもある。また、PPM は過去の実況や解析値を元 に係数を学習するため、データが保存されている限り、 10年や20年といった長期間のデータに基づいてガイ ダンスを作成することもできる。このため PPM では、 比較的稀な現象に対してもある程度多くのサンプル数 を用いて係数を学習することができる。これに対して MOS で学習した係数は、誤差特性が異なる数値予報モ デルに適用することは不適切であるため、数値予報モ デルの特性が大きくは変化していないと考えられる期 間(通常はせいぜい3,4年程度)しか過去に遡って学 習することができない。

一般に予報時間が進むほど数値予報モデルの予測誤

<sup>1</sup> 丁醛 渲

<sup>&</sup>lt;sup>2</sup> perfect prog, perfect prognostic または PP と呼ばれることも多い (Wilks (1995) など)。

http://glossary.ametsoc.org/wiki/Main\_Page

<sup>4</sup> 第 1.2 節でも述べたように、ガイダンスは系統誤差は補正できるがランダム誤差は補正できない。ただし、ランダム誤差がどれくらいあるか(予測と実況の相関がどれくらい弱いか)、という情報は考慮される。

差は大きくなるため、数値予報モデルのランダム誤差を 正しく反映した予測では、初期時刻に近い時刻はシャー プな予測を行い、予報時間が進むにつれてメリハリの ない(気候値に近い)予測をすることになる。実際、予 報時間で層別化 (第 2.2.5 項を参照 ) された MOS はこ のような予測を行う。このため、予報時間で層別化さ れた MOS の予測を初期時刻順に同じ対象時刻で並べ た場合、気象場の予測が変わらなかったとしても、初期 時刻が新しくなるにつれて予測がシャープになる。こ れは予報後半ほどランダム誤差が大きいという数値予 報モデルの特徴を正しく反映させた確率予測を行うよ うな場合には望ましい特性であるが、そうでない場合 には、初期時刻が新しくなるにつれて対象とする現象 が強まったかのような印象を与えてしまう。これに対 して PPM では、数値予報モデルのランダム誤差を考 慮しないため、常に数値予報モデルと同程度にシャー プな予測をすることになる。初期時刻順に同じ対象時 刻の PPM の予測を並べた場合、気象場の予測が変わ らなければ、初期時刻に関わらず常に同じような強度 の予測をすることになるため、PPM は現象の推移の把 握に適している。

数値予報モデルに系統誤差がある場合、PPMを用いると予測誤差が大きくなってしまう。そこで予測時の入力データに数値予報モデルの出力値を用いる代わりに系統誤差が補正されたガイダンス値を用いることがある。例えば最大降水量ガイダンスでは、学習時には解析雨量から算出した平均降水量と最大降水量を用いてその比を表す関係を学習し、予測では数値予報モデルの降水量の代わりに平均降水量ガイダンスを入力として最大降水量を求めている。また降水種別ガイダンスや雪水比ガイダンスでは、地上気温の実況を用いて関係式を学習し、予測では格子気温ガイダンスを入力としている。ガイダンス値を予測時の入力データに用いることで、PPMの利点を活かしつつ系統誤差を軽減することができる。

MOSとPPMの中間的な方法として、数値予報の初期時刻に近い時刻で係数を学習し、その係数をそれ以外の時刻の予測にも適用するという手法もある<sup>5</sup>。空域予報で利用されている乱気流指数と着氷指数などはこの手法を採用している。空域予報では短時間の予測が特に重要であるため、初期時刻に比較的近い時刻のデータを利用して係数を学習し、その係数をほかの時刻の予測にも適用している。これにより、注目したい時刻では最適な予測がされつつ、それよりも先の予測時刻でも着目している時刻と同程度のメリハリを持って予測されるため、乱気流や着氷の予測域の推移が把握しやすくなっている。ただし数値予報モデルの系統誤差が予報時間とともに変化する場合には、予報時間によって予測特性が変化することに注意が必要である。

#### 2.2.3 一括学習と逐次学習

統計手法を用いて予測式の係数を学習する場合、過去の一定期間のデータを用意し、過去データによる予測値とそれに対応する実況値から求められる誤差関数を最小にするような係数を求めることになる。この時、用意した全ての過去データを用いて算出した誤差関数を最小にするように係数を決定する手法を一括学習(またはバッチ学習)という。これに対して、過去データを時系列に並べて1組ずつ与え、データが与えられる度に係数を更新する手法を逐次学習(またはオンライン学習)という。また、一括学習に近い手法としてミニバッチ学習と呼ばれる学習方法もある。ミニバッチ学習は過去データの中から一定数のデータをランダムに抽出して学習する方法で、主にニューラルネットワークの学習に用いられる。これについては第2.6.6 項などで述べる。

本章で述べる統計手法の中では、線形重回帰とロジスティック回帰では一括学習が、ニューラルネットワークとカルマンフィルタでは逐次学習が用いられる場合が多いが、各統計手法の解説で述べるように、線形重回帰とロジスティック回帰でも逐次学習は可能で、ニューラルネットワークでも一括学習は用いられる。カルマンフィルタは基本的には係数を逐次更新することを前提とした手法であるが、一定期間で逐次学習してその後は係数を学習しないという手法を取るならば、一括学習ともいえる。ただしその場合には学習の最終日に最適化された係数で予測を行うことになるため、予測精度は線形重回帰や通常のカルマンフィルタと比べて低くなる。

一括学習を用いた場合、再学習するまで係数は変化 せず、ガイダンスの予測特性は変わらない。このため、 利用者にとってはガイダンスの特性が把握しやすいと いうメリットがある。しかし、季節変化や数値予報モ デルの特性変化に追随できないため、季節による層別 化(第2.2.5項を参照)が必要であったり、数値予報モ デルの更新によりモデルの特性が大きく変化する場合 には、ある程度長期間のデータを用いて係数を再学習 したりする必要がある。これに対し逐次学習を用いた 場合は、新しい実況データが得られる度に係数が変化 しガイダンスの予測特性も変わる。このため、一括学 習と比べるとガイダンスの特性は把握しにくくなるが、 季節変化や数値予報モデルの変化に追随することがで き、数値予報モデルの更新時には比較的短い期間で係 数を最適化できる場合がある。ただし必ずしも速やか に追随するわけではなく、例えば大雨や強風など頻度 の少ない現象を予測対象としている場合には、短い期 間では学習に必要なサンプル数が得られないため、数 値予報モデルの変化に適応するまでに数か月以上掛か る。学習速度を速めるように調整することもできるが、 単に速めただけでは実況を後追いするだけになってし まい、予測精度は低下する。

<sup>&</sup>lt;sup>5</sup> この手法は pseudo-PP (Météo-France 2015) とも呼ばれている。PP は perfect prognosis の略。

逐次学習型のガイダンスでは、稀な現象を予測できなかった場合、次回はその現象を予測しようとして係数を大きく変化させる。しかしそのような稀な現象は続いては起きないことが多いため、係数が元の状態に戻るまで過剰な予測を続けてしまう、というようなことが生じうる。このため、稀な現象に対して逐次学習を用いると予測精度が低くなる場合がある。一括学習ではこのようなことは起こらないため、稀な現象に対してもある程度の予測精度を持つことになる。このように、比較的頻度が高い現象を予測する場合は一括学習が適している。

逐次学習型のガイダンスは、数値予報ルーチンを運用する場合や、過去のデータを用いて予報実験を行う場合には、数値予報モデルのデータと実況データ、前回の係数を適切な順番で与える必要があり、データの取り回しが複雑になる。また逐次学習を行う中で、係数が異常な値になってしまうこともあるため、係数や予測値をモニターし、異常な場合には係数を差し替える等の対応が必要になる。一括学習型ガイダンスの場合は数値予報モデルの更新時を除いては係数を更新する必要がないため、逐次学習型ガイダンスに比べて維持・管理する手間は少ない。

#### 2.2.4 補正手法

降水量の実況の頻度分布は、弱い降水の頻度が非常 に多く、強い降水は少ないという偏った形をしている。 このような分布を持つデータに対して統計手法を用い ると、平均的なスコアを良くするために、頻度の低い 現象が予測されにくくなるという傾向がある 6。しか しながら気象予測においては、頻度が低い現象(大雨、 強風、低視程など)の予測が重要である場合が多く、こ のような現象に対しても実況と同程度の頻度で予測さ れることが望ましい。そこで、頻度バイアス補正と呼 ばれる補正手法が用いられている。頻度バイアス補正 では、統計手法により予測値を求めた後に、実況の頻 度と予測の頻度が等しくなるように予測値の補正を行 う。これにより、統計手法による予測では全く予測さ れない (頻度がゼロの)現象も学習期間内の実況と同 程度に予測されるようになるため、空振りは増えるが 的中も増え、稀な現象に対する精度は向上する。頻度 バイアス補正は、平均降水量ガイダンス、風ガイダン ス、視程ガイダンスなど、逐次学習型のガイダンスで かつ、実況の分布に偏りがあり、頻度が低い現象の予 測が重要なガイダンスに利用されている。頻度バイア ス補正について詳しくは第2.9節で述べる。

頻度バイアス補正以外にも、降水確率ガイダンスを

用いた平均降水量ガイダンスの補正、上空の気温予測等を用いた降水種別の補正、予測に最適な閾値を一定に近づけるための乱気流指数の補正など、各ガイダンスの特性や使用目的に応じて独自の補正を行っているガイダンスもある。これらについて詳しくは、第4章の各ガイダンスの解説で述べる。

#### 2.2.5 層別化

層別化は、学習データを地点や時刻、季節などに分 割し、それぞれに対して係数を学習し予測に利用する 手法である。地点で層別化した場合には、地点毎に異 なる係数を持つことになる。例えば気温予測の場合、東 風が吹くと A 地点では気温が下がるが、B 地点では逆 に気温が上がるなど、対象とする地点によって気温に 影響を及ぼす現象や及ぼし方が変わる。このような場 合、A 地点とB 地点で同じ係数を用いるとどちらの地 点に対しても予測が不十分になってしまうため、地点 で層別化して別々の係数を用いることが望ましい。同 じ地点でも時刻によって数値予報モデルの特性が異な る場合には時刻(予報時間または予報対象時刻)で層 別化する。例えば図 1.2.6 で示した新潟県粟島の気温 予測のように、数値予報モデルの気温予測が日中は負 バイアス、夜間は正バイアスを持つような場合、時刻 で層別化することで気温を適切に予測できる。

特殊な層別化を用いている例として視程ガイダンスがある。視程ガイダンスでは地点や時刻のほかに天気(雨、雪、無降水)でも層別化している。これは視程の予測式が天気によって大きく変わることを考慮するためであるが、天気の予測が外れると視程の予測も大外れする可能性があることに注意が必要である。

一般に、統計手法を用いる場合、係数の学習に用い るデータは同一の特性を持つことが望ましく、適切な 層別化は予測精度の向上をもたらす。ガイダンスが予 測対象とする現象は、地点や時刻、季節などによって 特性が変わるため、層別化は多くのガイダンスで利用 されている。一方、層別化の数を増やせば増やすほど 学習に利用できるデータ数は少なくなるため、係数の 推定精度が低くなったり、稀な現象に対する学習が不 十分になったりする。十分なサンプル数を確保するた めには、単純には学習期間を長くすればよいが、MOS を用いる場合には学習期間をいくらでも長くすること はできない。この対策として発雷確率ガイダンスでは クラスター分析が用いられている。クラスター分析を 用いることで似た特性を持つ地点を一つにまとめるこ とができるため、適切な層別化を行いつつ学習のため のサンプル数を増やすことができる。

#### 2.2.6 アンサンブル手法

アンサンブル予報 (気象庁予報部 2016) や複数の数値予報モデルが利用できる場合、各予測に対するガイダンス値のアンサンブル平均等を利用することで、決

<sup>&</sup>lt;sup>6</sup> 予測に位置ずれがあった場合、実況は頻度が多い側(降水の場合は降水量が少ない方)にずれる事例が多くなる。統計手法では平均的なスコアを良くするため頻度が多い現象を予測しやすくなるため、頻度が低い現象の予測は少なくなってしまう。

定論予測に基づくガイダンスよりも予測精度が高くなる場合が多い(第5.1.2項)。ただし最大降水量ガイダンスや視程ガイダンスのように、現象の最大値や最小値を予測するガイダンスの場合には、各予測に対するガイダンス値の単純な平均を用いた場合には予測精度が低下する。アンサンブル平均を用いる代わりに各メンバーの最大または最小値を用いることもできるが、一般に予測頻度が過大になり予測精度も低下する。このような場合には各メンバーの予測の分布に基づいた補正や予測を行う必要がある。

アンサンブル予報のスプレッドが適切で十分なメンバー数が得られる場合には、気温や風など量的な予測の誤差幅を示すことができる。また、ガイダンス値がある閾値を上回る割合などを利用することで、ロジスティック回帰やニューラルネットワーク等を用いなくても、予測の不確実性に応じた確率予測を行うこともできる。スプレッドが不十分な場合には誤差幅や確率予測も不十分になるが、第1.4節で示した米国のEKDMOSのように、スプレッドスキルの関係を利用した統計式で予測誤差を補正する手法も開発されている。

アンサンブル予報が利用できない場合でも、初期値の異なる複数の予測値をアンサンブルメンバーのように利用することで、アンサンブル平均が得られる。これを LAF (Lagged Average Forecast) 法 (Hoffman and Kalney 1983) と呼ぶ。LAF 法はアンサンブル予報を実行する必要がないため低コストで実施でき、予測精度の向上も期待できる。また、初期値方向に平滑化するため、初期値変わりの影響を軽減する効果がある。LAF 法は発雷確率ガイダンスで用いられており、決定論予測に基づくガイダンスと比べて予測精度や信頼度が向上することが確かめられている (高田 2008)。

#### 参考文献

- Glahn, H. R. and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.
- Hoffman, R. N. and E. Kalney, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, 35A, 100–118.
- 気象庁予報部, 2016: 確率的な気象予測のためのアンサンブル予報の課題と展望. 124 pp.
- Météo-France, 2015: WMO Technical Progress Report on the Global Data-Processing and Forecasting System and Numerical Weather Predition Research Activities for 2015. WMO/GDPFS.
- 高田伸一, 2008: 発雷確率ガイダンス. 平成 20 年度数値予報研修テキスト, 気象庁予報部, 88-89.
- Wilks, D. S., 1995: Statistical methods in the atmospheric sciences. Academic press, 467 pp.

#### 2.3 ガイダンスに用いる統計の基礎1

ここでは次節以降で述べる手法を記述する上で必要 となる統計の基礎について書く。本章で利用する主な 変数と添字の定義を付録 2.3.A にまとめる。

#### 2.3.1 確率変数と確率分布

サイコロを振るというような試行を N 回行ったとき、ある事象(この場合はサイコロの目)y が m 回起きたとする。N を大きくした場合に、y が起きた相対頻度 m/N がある一定値 p(y) に近づくと考えられるとき、p(y) を事象 y の確率という。

$$p(y) = \frac{m}{N} \tag{2.3.1}$$

サイコロを振る場合、それぞれの目が出る確率は例えば表 2.3.1 のようになる。確率は全て正の値を取り、起りうる全ての事象について和を取ると 1 になる。

サイコロを振る場合、どの目が出るかは確率的にしか知ることはできない。このように得られる結果が確率的に決まる変数を確率変数という。ある変数 y が確率変数である場合、y の具体的な値はまだ決まっておらず(サイコロを振る前の状態)、y の値は何らかの分布に従って確率的に与えられることになる。これを y の確率分布という。一方、変数 y が非確率変数である場合、y の値はサイコロを振って出た目のように既に決まった値となる。ある確率変数 Y が具体的な値 y を取る確率を  $P_x(Y=y)$  または単に p(y) と書く。

y が確率変数でその分布が  $\phi$  であるとき、「 y の分布 が  $\phi$  に従う」という意味で  $y\sim\phi$  と書く。例えば、 y が平均  $\mu$ 、分散  $\sigma^2$  の正規分布  $N(\mu,\sigma^2)$  に従うならば、  $y\sim N(\mu,\sigma^2)$  となる。確率変数 Y の分布がある連続的 な関数で与えられる場合、その関数を確率密度関数といい、以下の式で定義される。

$$f(y) = \lim_{h \to 0} \frac{P_r(y - h \le Y \le y + h)}{2h}$$
 (2.3.2)

例えばY が平均 $\mu$ 、分散 $\sigma^2$  の正規分布に従うならば、その確率密度関数は、

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$$
 (2.3.3)

となる。本章では記述を簡単にするため、以下では確率値と確率密度関数を同じ p(y) を使って書くこととする。

表 2.3.1 サイコロの目と確率

		U. 1	. – –	' ' '   C	- РД —		
目の値 $y$	1	2	3	4	5	6	合計
確率 $p(y)$	1/6	1/6	1/6	1/6	1/6	1/6	1

<sup>1</sup> 工藤 淳

#### 2.3.2 標本平均と期待値

例えば成人男性の平均身長を知りたい場合、成人男性全員の身長を測ることは困難であるため、一部の成人男性の身長を測ることで全体の平均を推定することになる。このような場合、調査対象全体の集合を母集団といい、調査のために母集団から抽出した一部を標本(またはサンプル)という。このように、標本を抽出する目的は、母集団がもつ何らかのパラメータ(平均など)を推定することである。

ある試行を N 回行い  $y_1,\ y_2,\ \cdots,\ y_N$  が得られたとき、得られた値の平均を標本平均という。標本平均  $\bar{y}$  を式で書くと、

$$\bar{y} = \frac{1}{N} \sum_{n=1}^{N} y_n \tag{2.3.4}$$

となる。表 2.3.1 のサイコロを N 回振った結果、1 が  $m_1$  回、2 が  $m_2$  回、 $\cdots$ 、6 が  $m_6$  回出たとする。この 時、サイコロの目の標本平均  $\bar{y}$  は、

$$\bar{y} = \frac{1}{N} \sum_{y=1}^{6} y \, m_y = \sum_{y=1}^{6} y \frac{m_y}{N}$$
 (2.3.5)

となる。

確率変数 Y が離散的な値  $y_1,\cdots,y_N$  を取り、その確率が  $p(y_1),\cdots,p(y_N)$  であったとすると、Y の期待値 E(Y) は次の式で定義される。

$$E(Y) = \sum_{n=1}^{N} y_n \, p(y_n) \tag{2.3.6}$$

また、確率密度が連続分布である場合の期待値は、

$$E(Y) = \int y \, p(y) dy \tag{2.3.7}$$

と定義される。X と Y が確率変数で、a と b が非確率 変数である場合、

$$E(X + Y) = E(X) + E(Y)$$
 (2.3.8)

$$E(aX + b) = aE(X) + b (2.3.9)$$

という関係が成り立つ。

サイコロの場合、 $y_n$  をそれぞれの目の値、 $p(y_n)$  をそれぞれの目が出る確率とすれば、サイコロの目の期待値は、

$$E(Y) = \sum_{n=1}^{6} y_n \, p(y_n) \tag{2.3.10}$$

となる。サイコロを振る回数を多くした場合、(2.3.1)式より  $m_y/N \to p(y)$  となるので、この場合  $\bar{y} \to E(Y)$  となる。この例からも分かる通り、期待値は平均(無限回試行した場合の標本平均)を表す値である。期待値と平均値は同じような意味を持つが、平均といった

場合には標本平均を指す場合があることに注意する。 期待値は「サイコロを1回振ったときに出る目の期待 値」など、1回の試行に対して定義できるのに対し、標 本平均は複数回試行したときに定義できるという違い がある。

サイコロを N 回振り、その標本平均を求めるとする。サイコロを振る前の時点ではどの目が出るか決まっていないため、n 回目にサイコロを振ったときに出る目  $y_n$  や N 回振った結果の標本平均  $\bar{y}$  も確率変数である。このとき、同じサイコロを N 回振るならば、各試行の期待値は等しい値  $\mu$  になる。すなわち、n によらず  $E(y_n)=\mu$  であるから、標本平均の期待値は

$$E(\bar{y}) = \frac{1}{N} \sum_{n=1}^{N} E(y_n) = \mu$$
 (2.3.11)

となり、各試行の期待値と一致する。この例のように、平均が既知の値  $\mu$  である母集団から N 個の標本を抽出してその標本平均をとる場合、標本平均の期待値は母集団の平均(母平均)と一致する。母集団のパラメータ  $\theta$  を標本から推定したとき、その推定値を  $\hat{\theta}$  とすると、 $E(\hat{\theta})=\theta$  であるならば、 $\hat{\theta}$  は  $\theta$  の不偏推定量であるという。(2.3.11) 式からわかるように、標本平均は母平均の不偏推定量である。

#### 2.3.3 標本分散と分散

ある試行を N 回行い  $y_1, y_2, \cdots, y_N$  が得られたとき、得られた値に対する分散を標本分散という。標本分散  $s^2$  を式で書くと、

$$s^{2} = \frac{1}{N} \sum_{n=1}^{N} (y_{n} - \bar{y})^{2}$$
 (2.3.12)

となる。また、

$$s = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (y_n - \bar{y})^2}$$
 (2.3.13)

を標本標準偏差という。分散と標準偏差は分布の広がりを表す値である。確率変数 y の分散 V(y) は次のように定義される。

$$V(y) = E \left[ (y - E(y))^{2} \right]$$
  
=  $E (y^{2}) - E(y)^{2}$  (2.3.14)

期待値と同様に、分散は1回の試行に対して定義できる値であるのに対し、標本分散は複数回試行したときに定義できる値である。Xを確率変数、aとbを非確率変数とすると、

$$V(aX + b) = a^{2}V(X)$$
 (2.3.15)

が成り立つ。また、2 つの確率変数 X と Y が独立、すなわち、X の結果に Y が依存しない場合、

$$V(X+Y) = V(X) + V(Y)$$
 (2.3.16)

が成り立つ。

(2.3.11) 式と同様に、平均と分散が既知の値  $\mu,\sigma^2$  である母集団から N 個の標本  $y_1,\cdots,y_N$  を抽出して標本分散  $s^2$  を求めたとすると、その期待値は、

$$E(s^2) = \frac{N-1}{N}\sigma^2$$
 (2.3.17)

となる(詳細は付録 2.3.B )。標本分散の期待値は母集団の分散(母分散)と一致しないため、標本分散は母分散の不偏推定量ではない。不偏推定量にするためには標本分散に N/(N-1) を掛ければよい。

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{n=1}^{N} (y_n - \bar{y})^2$$
 (2.3.18)

これを不偏分散という。

#### 2.3.4 共分散と相関係数

共分散と相関係数(相関関数ともいう)は対応のある 2 つのデータの関係性を表す値である。2 つの変数 x と y について N 組のデータ  $(x_1,y_1),\cdots,(x_N,y_N)$  が得られたとき、標本共分散  $\gamma_{xy}$  は次のように定義される。

$$\gamma_{xy} = \frac{1}{N} \sum_{n=1}^{N} (x_n - \bar{x}) (y_n - \bar{y})$$
 (2.3.19)

x と y の標本標準偏差を  $s_x, s_y$  としたとき、x と y の標本相関係数  $\rho_{xy}$  は次のように定義される。

$$\rho_{xy} = \frac{\gamma_{xy}}{s_x s_y} \tag{2.3.20}$$

相関係数は [-1,1] に規格化された共分散であり、x と y が完全な相関を持つ場合に +1、相関が全くない場合 に 0、完全な逆相関を持つ場合に -1 となる。

x と y が確率変数の場合、共分散 Cov(x,y) と相関係数 R(x,y) は次のように定義される。

$$Cov(x, y) = E[(x - E(x))(y - E(y))]$$
 (2.3.21)

$$R(x,y) = \frac{Cov(x,y)}{\sqrt{V(x)V(y)}}$$
(2.3.22)

## 2.3.5 自己共分散と自己相関係数

 $y_n$  が確率変数で時系列データである場合、n 番目のデータから見て時間方向に k だけ離れたデータとの共分散と相関係数をそれぞれラグ k の自己共分散  $C_{n,n+k}$  と自己相関係数  $R_{n,n+k}$  といい、以下のように定義される。

$$C_{n,n+k} = Cov(y_n, y_{n+k})$$
 (2.3.23)

$$R_{n,n+k} = \frac{C_{n,n+k}}{\sqrt{V(y_n)V(y_{n+k})}}$$
(2.3.24)

 $y_n$  の平均、分散、共分散が時間変化しない場合、すなわち、

$$E(y_n) = E(y_{n+k}) (2.3.25)$$

$$V(y_n) = V(y_{n+k}) (2.3.26)$$

$$Cov(y_n, y_m) = Cov(y_{n+k}, y_{m+k})$$
 (2.3.27)

である場合、 $y_n$  は定常であるという。この時、

$$C_{n,n+k} = C_{0,k} (2.3.28)$$

$$R_{n,n+k} = R_{0,k} (2.3.29)$$

であり、 $C_k \equiv C_{0,k}, R_k \equiv R_{0,k}$  と書くことにすると、

$$R_k = \frac{C_k}{\sqrt{V(y_0)V(y_k)}} = \frac{C_k}{C_0}$$
 (2.3.30)

となる。

 $y_n$  が非確率変数で定常な時系列データである場合、

$$\gamma_k = \frac{1}{N} \sum_{n=1}^{N} (y_n - \bar{y}) (y_{n+k} - \bar{y})$$
 (2.3.31)

をラグ k の標本自己共分散という。また、

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{2.3.32}$$

をラグ k の標本自己相関係数という。

定常な時系列データが平均0で時間方向に相関を持たない $(k \ge 1$  で $\rho_k = 0)$ 場合、その時系列はホワイトノイズであるという。ホワイトノイズの自己相関係数は、

$$R_k = \begin{cases} 1 & (k=0) \\ 0 & (k \ge 1) \end{cases}$$
 (2.3.33)

となる。時系列がホワイトノイズでサンプル数(時系列の長さ)N が大きい場合、標本自己相関係数は平均0、分散 1/N の正規分布に従う(例えば Shumway and Stoffer 2000 の第 1 章 )。

#### 2.3.6 同時確率、条件付確率、ベイズの定理

ある事象 x が起きたという条件の下で事象 y が起きる確率を条件付確率といい、p(y|x) と書く。このとき、y は確率変数であるのに対し、x は既に値が決まっているため非確率変数である。例えば図 2.3.1 の左図のように、x が起きる確率が 60%、起きない確率が 40%などと与えられたとき、p(y|x)=12/(12+48)=1/5 となる。

2つの事象 x と y があって、x が起きた場合と起きなかった場合で y の起きる確率が等しいとき、事象 y は事象 x と独立であるという。例えば、サイコロを 2 回振ったとき、2 回目に出た目 y は 1 回目に出た目 x の影響を受けないだろうから、この場合 x と y は独立である。 x と y が独立であるならば、p(y|x)=p(y) であり、逆に、 $p(y|x)\neq p(y)$  ならば x と y は従属であるという。図 2.3.1 の例では、左の図では x の発生の有無に関わらず y が起きる確率は p(y)=1/5 であり x と y は

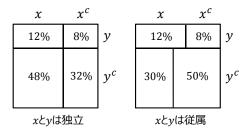


図 2.3.1 独立と従属の例。左は事象 x と y が独立、右は従属の場合。 $x^c, y^c$  はそれぞれ x と y の余事象を、数字は各事象が起きる確率を表す。

独立である。一方、右の図ではxの発生の有無によってyが起きる確率が変わるため、xとyは従属である。

事象 x と y が同時に起きる確率を同時確率といい p(x,y) と書く。p(x,y) は、事象 x が起きた場合に事象 y が起きる確率であると同時に、事象 y が起きた場合に事象 x が起きる確率でもあるため、

$$p(x,y) = p(y|x)p(x) = p(x|y)p(y)$$
 (2.3.34)

と書ける。これを確率の乗法定理という。図 2.3.1 の 左図の例では p(x,y)=12/100=3/25 で、これは  $p(y|x)p(x)=12/(12+48)\times(12+48)/100=3/25$  と 等しい。このことは図 2.3.1 の右図に対しても成り立つことが確かめられる。

x と y が独立であるならば p(y|x) = p(y) なので、

$$p(x,y) = p(x)p(y)$$
 (2.3.35)

と書ける。つまり、独立な事象の同時確率は各事象の確率の積になる。これはシンプルな関係式ではあるが、多くの統計手法では観測データが独立であることを仮定することで N 個の観測データ  $y_1,\cdots,y_N$  が得られた同時確率  $p(y_1,\cdots,y_N)$  を各確率の積  $p(y_1)\cdots p(y_N)$  として記述している。解説書の中には各観測が独立であると仮定していることを明示していない場合もあるので注意が必要である。当然だが、観測データが独立と見なせない場合にはこの関係を用いることはできない。

条件付確率の表記において、p(x|y,z) と書いた場合には、y と z が同時に起きたという条件の下で x が起きる確率を意味し、p(x,y|z) と書いた場合には z が起きたという条件の下で x と y が同時に起きる確率を意味する。この例のように、"|"の左右に書かれる変数は複数あっても良い。(2.3.34) 式に z が起きたという条件を追加すると、

$$p(x, y|z) = p(y|x, z)p(x|z) = p(x|y, z)p(y|z)$$
(2.3.36)

と書ける。(2.3.34) 式および (2.3.36) 式からすぐに、ベイズの定理と z で条件付けられたベイズの定理

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$
 (2.3.37)

$$p(x|y,z) = \frac{p(y|x,z)p(x|z)}{p(y|z)}$$
 (2.3.38)

#### が得られる。

同時確率 p(x,y) を x の全域で積分すると p(y) になる。これを周辺化という。例えば図 2.3.1 の左の例では、p(x,y)=12% と  $p(x^c,y)=8\%$  を足すと 20% となり、y が起きる確率になることが確かめられる。p を確率密度関数とし、(2.3.34) 式および (2.3.36) 式を x で積分すると、確率密度関数の周辺化の式

$$p(y) = \int p(x,y)dx$$
$$= \int p(y|x)p(x)dx = \int p(x|y)p(y)dx \qquad (2.3.39)$$

## および

$$p(y|z) = \int p(x,y|z)dx$$

$$= \int p(y|x,z)p(x|z)dx = \int p(x|y,z)p(y|z)dx$$
(2.3.40)

が得られる。

## 2.3.7 正規分布の確率密度関数

確率変数 w が平均  $\mu$ 、分散  $\sigma^2$  の正規分布  $N(\mu,\sigma^2)$  に従うとき、その確率密度関数は、

$$p(w|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} (w - \mu)^2\right] (2.3.41)$$

と書ける。また、確率変数 w が K 次元ベクトルで、平均  $\mu$ 、分散共分散行列  $\Sigma$  の正規分布に従う場合の確率密度関数は、

$$p(\boldsymbol{w}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$= (2\pi)^{-\frac{K}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} (\boldsymbol{w} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{w} - \boldsymbol{\mu}) \right]$$
(2.3.42)

と書ける。ここで  $| \ |$  は行列式を、T は転置を表す。 (2.3.42) 式を全区間で積分すると 1 になることから、

$$\int \exp\left[-\frac{1}{2}\left(\boldsymbol{w}-\boldsymbol{\mu}\right)^{T} \Sigma^{-1}\left(\boldsymbol{w}-\boldsymbol{\mu}\right)\right] d\boldsymbol{w} = (2\pi)^{\frac{K}{2}} |\Sigma|^{\frac{1}{2}}$$
(2.3.43)

となる。

#### 2.3.8 回帰分析

目的変数 y と説明変数 x の組が与えられたとき、

$$\hat{y} = f(\boldsymbol{x}, \boldsymbol{w}) \tag{2.3.44}$$

という y を近似する関係 (以下ではこれを統計モデルまたは単にモデルと呼ぶ)を仮定し、係数 w を統計的

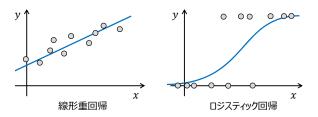


図 2.3.2 回帰の例

に推定することを回帰分析または回帰という。ガイダンスの場合、例えばyは地上気温の観測値、xは数値予報で予測された地上気温、風、雲量などになる。どのような統計モデルを仮定するかによって、線形重回帰、ロジスティック回帰、ポアッソン回帰など様々な回帰手法がある(図 2.3.2)。どの統計モデルを採用するかは、データの特性から判断する。

#### 2.3.9 尤度と最尤法

統計モデルが (2.3.44) 式で与えられており、1 組の目的変数と説明変数のデータ (y,x) が与えられたときに、係数 w を推定することを考える。このとき、w の推定値がどのような値かは分かっていないのだが、何か適当な値を与えれば、y が取る確率密度 p(y|x,w) を求めることができる。今、データ (y,x) が与えられているので、p(y|x,w) は w のみの関数であり、これをL(w) と書く。

$$L(\boldsymbol{w}) = p(y|\boldsymbol{x}, \boldsymbol{w}) \tag{2.3.45}$$

このとき、L(w) を尤度 (または尤度関数 ) という。 (2.3.45) 式は、左辺では w は変数として扱われているのに対し、右辺では確定した値として与えられているため、やや奇妙に見えるかもしれない。しかし、例えば p(y|x,w) が正規分布で書けると仮定した場合には、

$$p(y|\boldsymbol{x}, \boldsymbol{w}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y - f(\boldsymbol{x}, \boldsymbol{w}))^2}{2\sigma^2}\right] (2.3.46)$$

のように書けることから、数学的には p(y|x,w) は y,x,w の関数であり、どれを変数と見なすかは自由 に決めてよい。

例えばwを $w_1$ と $w_2$ に設定し、それぞれに対して尤度を求めたとき、 $L(w_1) > L(w_2)$  であったとする。すなわち、 $p(y|x,w_1) > p(y|x,w_2)$  であったとする。これは、wとして $w_2$ よりも $w_1$ を採用した方が目的変数の値yが得られる確率が高いことを意味している。例えばある1回の試行によってyという結果が得られた場合、yが得られる確率は低かったがたまたまyが得られたと考えるよりも、yが得られる確率が高かったのでyが得られた、と考える方が自然だろう。今は $p(y|x,w_1) > p(y|x,w_2)$ としているので、wとしては $w_1$ の方、すなわち尤度が大きい方が尤もらしいという

ことができる。このような考えの下で尤度が最大になる時のwを係数の推定値とする手法を最尤法という。

次に、与えられたデータセットが1組ではなく、N組 $(y_1, x_1), \cdots, (y_N, x_N)$ ある場合を考える。この場合の尤度は、

$$L(\boldsymbol{w}) = p(y_1, \cdots, y_N | \boldsymbol{x}_1, \cdots, \boldsymbol{x}_N, \boldsymbol{w}) \qquad (2.3.47)$$

という同時確率で表されるが、それぞれの観測が独立 であるならば、同時確率は各確率の積で表されること から、

$$L(w) = \prod_{n=1}^{N} p(y_n | x_n, w)$$
 (2.3.48)

と書ける。最尤法ではこれを最大にする w を求める。確率密度は常に正の値を取ることから、尤度を最大にする w を求めることは尤度の対数 (対数尤度)を最大にする w を求めることと等しい。(2.3.48) 式の対数を取ると、

$$\ln L(\boldsymbol{w}) = \sum_{n=1}^{N} \ln p(y_n | \boldsymbol{x}_n, \boldsymbol{w})$$
 (2.3.49)

となる。対数を取ることで積が和になるため、次項で述べる最急降下法や確率的勾配降下法において微分の扱いが容易になる。

#### 2.3.10 最急降下法と確率的勾配降下法

統計手法において、一定期間の過去データを用いて統計モデルの係数 w を決定することを学習という。係数を決定する場合、予測値 f(x,w) と実況値 y から計算される誤差関数 E を設定し、E を最小にするような係数を最適な係数とする、という手法がしばしば用いられる。誤差関数としては、平均二乗誤差 (MSE) や負の対数尤度((2.3.49) 式に負号を付けた値)が用いられることから、一般に以下のような形をしている。

$$E = \sum_{n=1}^{N} E_n \tag{2.3.50}$$

$$E_n = h(y_n, \boldsymbol{x}_n, \boldsymbol{w}) \tag{2.3.51}$$

ここで h は MSE や負の対数尤度を表す関数である。求めたい係数は、上記の誤差関数を最小にする w であるから、E を w の各成分  $w_k$  で微分して 0 になるときの w を求めれば良い。

$$\frac{\partial E}{\partial w_k} = \sum_{n=1}^{N} \frac{\partial E_n}{\partial w_k} = 0 \tag{2.3.52}$$

通常はこれを解析的に解くことはできないため、ニュートン・ラフソン法などを使って数値的に解くことになる。またはもっと単純に、E の  $w_k$  方向の傾きに従ってある一定の正の割合  $\eta$  で  $w_k$  を繰り返し変化させて

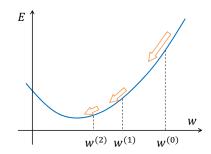


図 2.3.3 最急降下法による係数更新

もよい。このとき、繰り返しのステップを s 回行った 時の係数を  $oldsymbol{w}^{(s)}$  と書くと係数更新のステップは以下の ように書ける。

$$w_k^{(s+1)} = w_k^{(s)} - \eta \sum_{n=1}^{N} \frac{\partial E_n}{\partial w_k} \bigg|_{\boldsymbol{w} = \boldsymbol{w}^{(s)}}$$
(2.3.53)

これを適当な初期値  $w^{(0)}$  から始めて繰り返し実行し、 $w_k^{(s+1)}$  を用いて計算される誤差関数 E の値が s に依らず一定に近い値に収束したときの係数を w の推定値とする。これを最急降下法と呼ぶ。 (2.3.53) 式のように、全学習データを用いて係数を学習する方法を一括学習という。最急降下法による係数更新のイメージを図 2.3.3 に示す。

最急降下法の場合、1 回のステップ毎に学習データの数 N だけ微分計算を行う必要があるため計算量が多くなる。そこで、全学習データの中から各ステップ毎にサンプル  $\mathcal{D}$  (サンプル数  $N_{\mathcal{D}}$  個)をランダムに選択し係数を更新することを考える。すなわち、

$$w_k^{(s+1)} = w_k^{(s)} - \eta \sum_{n \in \mathcal{D}} \frac{\partial E_n}{\partial w_k} \bigg|_{\boldsymbol{w} = \boldsymbol{w}^{(s)}}$$
(2.3.54)

とする。この方法を確率的勾配降下法と呼ぶ。(2.3.54) 式のように、ランダムに選ばれたサンプルを用いて係 数を学習する方法をミニバッチ学習という。

確率的勾配降下法では  $N_{\mathcal{D}}$  個のサンプルを抽出するが、抽出するサンプル数は 1 個でも良い。すなわち、

$$w_k^{(s+1)} = w_k^{(s)} - \eta \left. \frac{\partial E_n}{\partial w_k} \right|_{w \to w(s)}$$
 (2.3.55)

とする。このとき学習データの番号を表す添字 n はランダムに選ばれたサンプルの番号を表す。ここで、学習データをランダムに選ぶのではなく、時系列順に与えることを考える。すなわち、(2.3.55) 式において、n を時刻 t で置き換える。さらに、係数の更新ステップは各時刻で 1 回だけ行うこととし、s も時刻 t で置き換えると、

$$w_k^{(t+1)} = w_k^{(t)} - \eta \left. \frac{\partial E_t}{\partial w_k} \right|_{\boldsymbol{w} = \boldsymbol{w}^{(t)}}$$
 (2.3.56)

となり、時刻 t の係数と誤差関数を用いて次の時刻 t+1 の係数を求める式、すなわち係数を逐次更新する式と

なる。(2.3.56) 式のように、新たな学習データが得られる度に係数を学習する手法を逐次学習という。

最急降下法、確率的勾配降下法のいずれを用いる場合でも、 $\eta$  を適切に与えるとともに、説明変数を規格化しておく必要がある。 $\eta$  は、一括学習やミニバッチ学習においては学習の効率を決定するパラメータで、小さずぎると収束するまでに時間が掛かり、大きすぎると極小値付近で振動してしまいいつまでも収束しなくなる。一方、逐次学習においては1回の学習での係数の変動量を決めるパラメータで、小さすぎると季節変化や数値予報モデルの変更への追随が遅くなり、大きすぎると一つの事例で係数が大きく変わってしまい係数が安定しなくなる。このように、最急降下法や確率的勾配降下法において $\eta$  は重要なパラメータであり、適切な値となるように調整する必要がある。

ηが適切であったとしても、説明変数のオーダーに差 があると学習がうまくいかない。例えば目的変数 y と 2つの説明変数  $x_1,\ x_2$  に線形モデルを仮定し、 $y,\ x_1,\ x_2$ のオーダーがそれぞれ、 $1, 1, 10^{-2}$  であったとする。 各説明変数の寄与量  $(x_1w_1 \, lacksymbol{\mathsf{L}} \, x_2w_2)$  が同程度であると すれば、 $w_1$ と $w_2$ のオーダーはそれぞれ 1と $10^2$ にな る。誤差関数に平均二乗誤差を用いた場合、誤差関数 のオーダーは1で、 $\partial E/\partial w_1$ と $\partial E/\partial w_2$ のオーダーは それぞれ  $1 \ge 10^{-2}$  になる。このとき  $\eta = 10^{-2}$  を用い たとして(2.3.53)式などに当てはめると、1 ステップの 更新でのwの変化量のオーダーは、 $w_1, w_2$  それぞれ に対して、 $10^{-2}$ ,  $10^{-4}$  となる。 $w_1$ ,  $w_2$  のオーダーは  $1 \ge 10^2$  であったから、 $w_1$  に対して  $10^{-2}$  倍程度の幅 で係数が更新されるのに対して、 $w_2$  に対しては  $10^{-6}$ 程度の幅でしか係数が更新されなくなってしまう。こ のため、 $w_2$  の更新速度が非常に遅くなり、計算の効率 が悪くなる。逆に、 $w_2$  の更新速度を速めようとして  $\eta = 10^2$  とすると、 $w_1$  に対しては更新速度が速すぎて いつまでも収束しなくなってしまう。このようなこと を避けるために、説明変数を何らかの方法で規格化し、 説明変数のオーダーを揃えておく必要がある。説明変 数のオーダーを揃えるのではなく、係数毎に $\eta$  を調整 してもよいのだが、通常、 $\eta$  の最適値を探すために  $\eta$ を変えながら何度も調整する必要があるため、その度 に全ての説明変数に対する $\eta$ を調整すると煩雑になっ てしまう。説明変数を規格化しておけば、調整するパ ラメータは1つだけで良いため取扱いが容易になる。

説明変数を規格化する方法としては、次の2つがよく用いられる。

$$x_k \to \frac{x_k - \bar{x}_k}{s_k} \tag{2.3.57}$$

$$x_k \to \frac{x_k - x_{\min}}{x_{\max} - x_{\min}} \tag{2.3.58}$$

ここで、 $\bar{x}_k$ と  $s_k$  は説明変数の標本平均と標本標準偏差、 $x_{\min}$ と  $x_{\max}$  は説明変数の最大値と最小値で、それぞれの値は学習データから求められるか、気候値や理

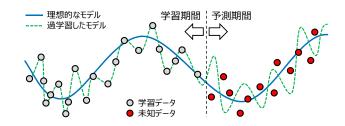


図 2.3.4 過学習のイメージ図

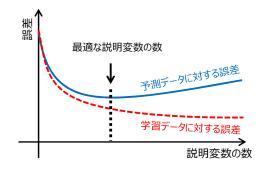


図 2.3.5 説明変数の数を変化させた場合の、学習データと 予測データそれぞれに対する誤差のイメージ。

論値、または実用的な値を設定している。(2.3.57) 式を用いた場合には平均0、標準偏差1に、(2.3.58) 式を用いた場合には $0\sim1$ の値に規格化される。

## 2.3.11 説明変数の選択

回帰分析において、目的変数 y が x の多項式  $y=a+bx+cx^2+\cdots$  で表されると仮定した場合、多項式の次元を増やすほど学習データに最適化されたモデルが得られる(図 2.3.4 の破線 )。しかしこのような回帰式を未知のデータに適用すると、図に示したように未知データに対する誤差(予測誤差)は大きくなってしまう。このように、学習データに対してモデルを最適化させ過ぎたために予測誤差が大きくなることを過学習(オーバーフィッティング)という。ガイダンスの目的は予測誤差を小さくすることであるから、図 2.3.4 の実線のようなある程度滑らかなモデルが理想的だと考えられる。

回帰分析では、過学習は説明変数の数が多すぎる場合に起きる。図 2.3.5 に、説明変数の数を変化させたときの学習データと予測データ(未知データ)それぞれに対する誤差のイメージを示す。一般に説明変数の数が少なすぎると、モデルの表現力が低いために学習データ・予測データのいずれに対しても誤差が大きくなる。そこで説明変数を増やしていくと、ある所までは学習データ・予測データとも誤差が小さくなっていくが、説明変数が多くなりすぎると、学習データに対する誤差は小さくなっていく一方で、予測データに対する誤差は増えていく。

予測誤差が小さいモデルを作成するためには、図2.3.5 の矢印で示したように説明変数を適切に選択する必要がある。ただし予測誤差が最小になる説明変数の組み合わせを事前に知ることはできないため、学習データだけを用いて最適と思われる説明変数の組み合わせを選択することを考える。説明変数を選択する方法としては主に以下の3つが挙げられる。

- 赤池情報量基準 (Akaike 1973) 等の情報量基準を 用いる方法
- 交差検証を用いる方法
- 主成分分析を用いる方法

以下ではそれぞれについて解説する。

(1) 赤池情報量基準を用いる方法 赤池情報量基準 (AIC) は以下の式で定義される。

$$AIC = -2\ln L + 2K (2.3.59)$$

ここで L は尤度、K はモデルの自由度である。回帰分 析の場合、モデルの自由度は係数の数(説明変数の数 にバイアス項の分の1を加えた数)である。右辺第1項 は学習データに対する当てはまりの良さを表し、学習 データに対する誤差が小さいほど値が小さくなる。右 辺第2項はモデルの自由度を表し、回帰分析の場合は 係数の数が少ないほど値が小さくなる。AICは、学習 データに対する誤差が小さく、係数の数が少ないほど 値が小さくなる指標で、AIC が相対的に小さいモデル ほど未知のデータに対する予測能力(汎化能力)が高 いと考える。対数尤度 (2.3.49) 式は確率の対数 (正の 値)を学習データについて和を取った値であることか ら、学習データの数が多いほど負の対数尤度は小さく なり AIC も小さくなるが、これは学習データが多いほ ど予測能力が高いという意味ではない。AIC では同じ 学習データに対して大小を比較することに意味がある。

学習データに対して、AIC が最も小さくなる説明変 数の組み合わせを選択する方法としては、総当たり法 や変数減少法、変数増加法、変数増減法が用いられる。 総当たり法は、説明変数の全ての組み合わせに対して AIC を算出し、最も AIC が小さい組み合わせを選択す る方法である。変数減少法では、初めに全ての説明変 数を利用して AIC を算出し、次に説明変数を1つずつ 除いてそれぞれの AIC を算出し、最も AIC が小さくな る説明変数を1つ削除する。これを繰り返し、どの説明 変数を除いてもそれ以上 AIC が小さくならなくなった 場合、それを最適な説明変数の組み合わせとする。変数 増加法は変数減少法とは逆に説明変数を1つずつ増や しながら説明変数の組み合わせを決定する方法で、変 数増減法は変数の増加・減少を行いながら最も AIC が 小さくなる組み合わせを決定する方法である。説明変 数の候補の数が計算機資源と比べてそれほど多くない 場合には総当たり法を用いればよい。説明変数の候補 が多すぎて総当たりでは計算に時間が掛かる場合は変 ④ ①~③を全学習データに対して行い誤差の平均を求める

図 2.3.6 LOOCV の手順の模式図

数減少法等を用いる。AIC と同様の基準として、サンプル数が少ない場合に AIC を修正した AICc (Sugiura 1978) やブートストラップ法 (第 2.3.12 項を参照)を用いて AIC を拡張した EIC (Ishiguro et al. 1997)、ベイズ統計学に基づく BIC (Schwarz 1978) など様々な情報量基準が提案されているが、ガイダンスに用いる場合は、経験的には AIC か次に述べる交差検証を用いれば十分である。

## (2) 交差検証を用いる方法

交差検証(クロスバリデーション((CV))は学習データを用いて未知データに対する誤差を推定する手法の一つである。ガイダンスの開発においては、十分な数の検証用データが得られない場合に学習データから予測誤差を評価する手法として用いられることが多いが、説明変数の選択に利用することもできる。最も単純な(CV)として、LOOCV(Leave-One-Out (CV))がある。図 (CV)2.3.6に示すように、LOOCVでは(CV)1 を取っておいて、残りの(CV)3 を取っておいて、残りの(CV)4 個のデータで予測式(CV)5 を取っておいた(CV)6 で予測式(CV)7 の作成を行い、取っておいた(CV)7 個のデータで検証を行う(二乗誤差等を求める)。これを(CV)8 個のデータ全てに対して行って誤差の平均を求めることで全体の誤差を評価する。式で書くと、LOOCVで平均二乗誤差 MSE を求める場合、

$$MSE_{LOOCV} = \frac{1}{N} \sum_{n=1}^{N} \left( y_n - \hat{f}^{-n}(\boldsymbol{x}_n) \right)^2 \quad (2.3.60)$$

となる。この計算を総当たり法や変数減少法等を用いて説明変数の組み合わせを変えながら実行し、CVで求められた誤差が最も小さくなる説明変数の組み合わせを選ぶ。

LOOCV のほかにも、k-fold CV とホールドアウト 法と呼ばれる手法もある。k-fold CV は LOOCV と同様だが、学習データをデータ数の等しい k 個のグループに分けておき、1 つのグループを除いて予測式を作成したのち、取っておいた 1 つのグループで検証を行って誤差を求める。これを k 個全てのグループについて行い、誤差を平均することで全体の誤差を評価する。

k-fold CV と呼ぶ代わりに「1 か月抜き CV」のように、取っておくデータの期間を用いて呼ぶこともある。ホールドアウト法では N 個の学習データの中から k 個のデータをランダムに抽出して検証用データとし、残りの N-k 個のデータで予測式を作成する。ホールドアウト法では予測式の作成と検証は 1 回のみ行い、LOOCV や k-fold CV のように全データをカバーするように検証を繰り返し行うことはしないため、厳密には交差検証とは呼ばれないが、シンプルで計算コストが少ないという利点がある。ただし繰り返し計算を行わないため、ランダム抽出に伴う検証結果の誤差(ばらつき)が大きくなってしまうので、計算時間に問題がない限りは LOOCV か k-fold CV を用いる。

説明変数の選択では、AIC を用いた方法と CV (LOOCV, k-fold CV) を用いた方法はどちらも同じような結果を与えることが多い。AIC を用いて説明変数を選択する場合はモデルの良さを評価する基準は 1 つしかないが、CV の場合は何でも良く、スレットスコアや RMSE など開発者が任意に設定できるという利点がある。ただし CV では予測式を繰り返し作成する必要があるため計算に時間が掛かる。

CV は学習データだけを用いて予測誤差を推定する手法であるが、検証用に取っておいたデータと予測式作成用のデータに相関がある場合には、実際に未知のデータを用いて検証を行った場合と比べて誤差がやや小さく見積もられてしまうことには注意しなければならない。日々の気象データの場合には一般に時間方向に強い相関があるため、LOOCV を用いた場合には上記の傾向は特に強くなる。このような場合にはLOOCVではなく k-fold CV を用いる。

CV を用いた場合には、CV による誤差を最小にす る説明変数の組み合わせが1つだけ選ばれることにな る2。この説明変数の組み合わせは全ての学習データ を用いて選ばれており、その意味で、学習データに特 化した説明変数の組み合わせが選ばれたということが できる。しかし、もし異なる期間の学習データを用い ることができたならば、異なる組み合わせの説明変数 が選ばれる可能性もあるだろう。すなわち、説明変数 の組み合わせを1つに限定してしまうことは、選択し た説明変数の組み合わせが予測データに対しては最適 ではないかもしれないという可能性を排除することに なる。このため、CV によって学習データに特化され たモデルで誤差を推定した場合、誤差を過小評価して しまうことにつながる。このような場合、選択された 説明変数による予測誤差を実際の値に近い値で推定す る方法として、DCV(ダブル・クロスバリデーション) がある。LOOCV と同様に検証用にデータを1つ取っ ておく場合の DCV は以下の手順で行う。

- 1. N 個の学習データの中から検証用に n 番目のデータを一つ取り除く。
- 2. 残りの N-1 個のデータに対して、AIC や CV に基づいて説明変数の選択を行い予測式を作成する。 この予測式を  $\hat{d}^{-n}$  と書く。
- 3. 全ての n に対して上記 1, 2 の手順を繰り返す。
- 4. 得られたそれぞれの予測式  $\hat{d}^{-n}$  と検証用に取っておいた n 番目のデータを用いて平均二乗誤差などを求め、それを予測誤差の推定値とする。

この手続きを式で書くと以下のようになる。

$$MSE_{DCV} = \frac{1}{N} \sum_{n=1}^{N} \left( y_n - \hat{d}^{-n}(\boldsymbol{x}_n) \right)^2 \qquad (2.3.61)$$

上記の手順では検証用データを 1 つだけ取っておいたが、k-fold CV と同様に学習データを k 個のグループに分けて DCV することもできる。DCV は CV と似た手法であるが、取り除いたデータとは独立に最適な説明変数が毎回選択されるため、(2.3.61) 式で用いられる予測式は n によって異なる。DCV ではモデル選択の誤差も含めて予測誤差を推定することができるため、CV や AIC 等で選択されたモデルを未知データに適用した場合の誤差を CV と比べて精度よく見積もることができる。

### (3) 主成分分析を用いる方法

主成分分析では K 個の説明変数を M 個 ( $M \le K$ ) の合成変数に変換することで、元の説明変数が持つ情報をより少ない数の説明変数に縮約する。主成分分析は原理的には次のような手法である。ここでは簡単のために説明変数は  $x_1$  と  $x_2$  の 2 つだけだとする。 2 つの説明変数を学習データについてプロットすると例えば図 2.3.7 のようになる。このデータを  $x_1$   $-x_2$  空間内の新たな軸に射影した時、射影したデータの不偏分散が最も大きくなる軸を  $z_1$  とし、 $z_1$  に直行する軸を  $z_2$  とする。この例では  $x_1$  と  $x_2$  の相関が強く、 $x_1$  と  $x_2$  が持っていた情報はほぼ  $z_1$  で説明することができるため、説明変数として  $z_1$  だけ用いることで、説明変数の数を減らすことができる。

説明変数の数がK個ある場合は、射影したデータの不偏分散が最も大きくなる軸を $z_1$ とし、 $z_1$ に直行する軸の中で射影したデータの不偏分散が最も大きくなる軸を $z_2$ とする、ということを繰り返して $z_K$ までの軸を決定する。m番目の合成変数は次のように書ける。

$$z_m = \sum_{k=1}^{K} a_{mk} x_k , \quad \sum_{k=1}^{K} a_{mk}^2 = 1$$
 (2.3.62)

ここで  $a_m$  は x を  $z_m$  に変換するための係数である。 実際の主成分分析では、N 個の学習データについて、 規格化した説明変数の分散共分散行列の固有ベクトル と固有値を求めることになる。このとき、固有ベクト ルが係数ベクトル  $a_m$  に、固有値  $\lambda_m$  がその軸に射影し

<sup>&</sup>lt;sup>2</sup> AIC を用いた場合でも、説明変数の組み合わせが 1 つだけ選ばれるという事情は同じである。

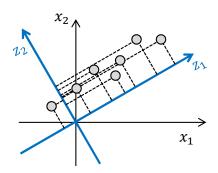


図 2.3.7 主成分分析の例

たデータの標本分散となるので、固有値が大きい順番に第一主成分、第二主成分、 $\cdots$ 、第K 主成分とする。固有値が大きい順番に $m=1,2,\cdots,K$  としたとき、

$$C = \frac{\sum_{m=1}^{M} \lambda_m}{\sum_{k=1}^{K} \lambda_k}$$
 (2.3.63)

を第 M 主成分までの累積寄与率という。主成分分析では、累積寄与率がある値(例えば 0.7 など)になるまでの主成分を用いたり、あらかじめ設定した数(例えば元の説明変数の数の 7 割など)の主成分を用いたりすることで、元の説明変数が持つ情報をできるだけ保持しつつ説明変数の数を減らすことができる。

固有ベクトルと固有値を求める際に、分散共分散行列を用いる代わりに、説明変数の相関行列を用いる場合もある。相関行列を用いる場合は説明変数を事前に規格化する必要がないため扱いがシンプルになる。この場合の累積寄与率は以下のとおりである。

$$C = \frac{1}{K} \sum_{m=1}^{M} \lambda_m \tag{2.3.64}$$

# 2.3.12 ブートストラップ法

ブートストラップ法 (Efron 1979; Efron and Tibshirani 1986) はリサンプリングによって推定量の分布を見積もる手法である。例えば 100 人の成人男性の身長が得られたとして、これを元に母集団 (成人男性全体)の身長の平均や分散 (母平均、母分散)を推定することを考える。もし母集団の分布が分かっているならば、その分布に応じて母平均や母分散などを推定したり検定したりすることができる。これをパラメトリック法という。しかしながら母集団の分布が分かっているとは限らない。そのような場合に分布を推定する手法をノンパラメトリック法という。ブートストラップ法はノンパラメトリック法の一種であり、シンプルな手法で応用範囲が広く、様々な手法で利用されている。

ブートストラップ法によるパラメータ推定は以下のように行う。まず観測された 1 組 ( データ数 N ) のデータセットから復元抽出によりランダムに N 個のデータを抽出 ( リサンプリング ) する。復元抽出とは、例え

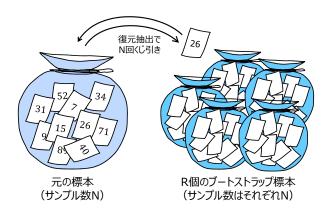


図 2.3.8 プートストラップ法によるデータセット生成のイ メージ

ばくじ引きをする場合であれば、一度引いたくじを戻して再びくじを引くことを意味する。このようにして抽出された N 個のデータからなるデータセットをブートストラップ標本という。これを R 回繰り返すことで R 組のブートストラップ標本を生成する (図 2.3.8)。 得られた R 組のブートストラップ標本と元の 1 組の標本それぞれについて求めたい推定量を算出する。この R+1 個の推定量により推定量の分布が得られる。また、R+1 個の推定量の平均から、求めたい平均や分散などの推定量が得られる。R の大きさは計算に掛かる時間と推定の精度を考慮して設定することになるが、経験的には  $500\sim5000$  程度で十分である。

プートストラップ法により、推定値の信頼区間(真値がある確率で含まれる区間)を求めることや、仮説検定を行うことができる。例えば平均値の 95%信頼区間を求める場合、R+1 個のブートストラップ標本による平均値を昇順に並べたとき、下限から 2.5%に位置する値と、上限から 2.5%に位置する値を信頼区間の下限・上限とすることができる。同様の考え方で仮説検定を行うこともできる。

スレットスコアやバイアススコアなど、分割表から 算出される検証スコアの信頼区間もブートストラップ 法で求めることができる (Kane and Brown 2000)。 サ ンプル数が N の検証スコアの信頼区間をブートスト ラップ法で求める場合には、FO, FX, XO, XX(定義 は巻末の付録 A を参照) をくじと見なし、それぞれの くじが FO/N、FX/N、XO/N、XX/N の割合で含まれ ると考えて復元抽出により合計 N回くじを引く。これ を R 回繰り返し、それぞれのブートストラップ標本に 対してスレットスコアなどを算出し、信頼区間を求め る。このとき、1 つのブートストラップ標本を生成する ために実際に乱数をN回発生させてくじ引きすると、 N に比例して計算量が多くなってしまう。そこで、く じを N 回引く代わりに FO, FX, XO, XX の割合に応 じたサイズが N の 4 項分布に従う乱数を 1 回発生させ る。これにより、サンプル数に関わらず1つのブート ストラップ標本を一度に生成することができるため処 理を高速化できる。

表 2.3.2 本章で利用する主な変数と添字の定義。表中の NN はニューラルネットワークを意味する。

文字	意味
x, X	説明変数、NN ではユニットへの入力データ
y, Y	目的変数、実況値、NN では教師データ
$\hat{y}, p$	ガイダンスの予測値(出力値) 確率値
w, W	回帰係数、係数、重み係数 (バイアス項を含める)
E	誤差関数(損失関数と呼ばれることも多い) $E()$ は期待値の意味で用いる
V()	分散
L	尤度
s	学習のステップ番号、標本標準偏差
t	時刻
f, g	関数、NN では活性化関数
$\phi$	加重和
$\eta$	学習率
$\lambda, \beta$	正則化パラメータ
$\mu$	平均
$\sigma^2, \Sigma$	分散、分散共分散行列

文字	最大値	意味
$\overline{k}$	K	説明変数の番号、 NN では入力層のユニット番号
l	L	NN では中間層のユニット番号
m	M	NN では出力層のユニット番号
$\overline{n}$	N	学習データの番号

# 付録 2.3.A 主な変数と添字の定義

本章で利用する主な変数と添字の定義を表 2.3.2 にまとめる。複数の統計手法を可能な限り同じ変数で表記するため、一般的なテキストやこれまでのガイダンスの解説に用いられてきた変数とは異なる場合があるので注意していただきたい。ここに書いていない文字やここでの定義と異なる意味で利用する際は本文中で追宜解説する。式の記述において、添字が同じ場合は和を取るというアインシュタインの縮約記法が用いられる場合があるが、統計手法においては添字が同じでも必ずしも和を取るとは限らないため、和を取る必要がある場合には和の記号  $\Sigma$  を省略せずに表記する。説明変数 x と係数 w は、大文字で書いた場合は行列を、小文字の太字で書いた場合はベクトルを、小文字で書いた場合は行列またはベクトルの成分を表すものとする。

#### 付録 2.3.B 標本分散の期待値

ここでは平均と分散が既知の値  $\mu$ ,  $\sigma^2$  である母集団 から N 個の標本  $y_1,\cdots,y_N$  を独立に抽出して標本分散  $s^2$  を求める場合、その期待値が (2.3.17) 式

$$E(s^2) = \frac{N-1}{N}\sigma^2$$

となることを示す。

初めに標本平均  $\bar{y}$  の分散を求める。

$$V(\bar{y}) = \sum_{n=1}^{N} V\left(\frac{y_n}{N}\right) = \sum_{n=1}^{N} \frac{\sigma^2}{N^2} = \frac{\sigma^2}{N}$$
 (2.3.65)

ここで、独立な確率変数に対する分散の性質 (2.3.15) 式 および (2.3.16) 式を用いている。標本平均  $\bar{y}$  の分散はまた、分散の定義 (2.3.14) 式と標本平均の期待値 (2.3.11) 式より、

$$V(\bar{y}) = E\left[ (\bar{y} - E(\bar{y}))^2 \right] = E\left[ (\bar{y} - \mu)^2 \right] (2.3.66)$$

とも書ける。これらの式と標本平均の定義 (2.3.4) 式、分散の定義 (2.3.14) 式などを用いると、標本分散の期待値は

$$E(s^{2}) = \frac{1}{N} \sum_{n=1}^{N} E\left[ (y_{n} - \bar{y})^{2} \right]$$

$$= \frac{1}{N} \sum_{n=1}^{N} E\left[ \{ (y_{n} - \mu) - (\bar{y} - \mu) \}^{2} \right]$$

$$= \frac{1}{N} \sum_{n=1}^{N} E\left[ (y_{n} - \mu)^{2} \right] + \frac{1}{N} \sum_{n=1}^{N} E\left[ (\bar{y} - \mu)^{2} \right]$$

$$- \frac{2}{N} E\left[ (\bar{y} - \mu) \sum_{n=1}^{N} (y_{n} - \mu) \right]$$

$$= \sigma^{2} - E\left[ (\bar{y} - \mu)^{2} \right]$$

$$= \sigma^{2} - \frac{\sigma^{2}}{N} = \frac{N-1}{N} \sigma^{2}$$
(2.3.67)

となり、(2.3.17) 式が導かれる。

# 参考文献

Akaike, H., 1973: Information theory and an extension of the maximum likelihood principle. B.N. Petrov, F. Csaki (Eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest, 267–281.

Efron, B., 1979: Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, **7(1)**, 1–26.

Efron, B. and R. Tibshirani, 1986: Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistic Science*, 1, 54–75.

Ishiguro, M., Y. Sakamoto, and G. Kitagawa, 1997: Bootstrapping log likelihood and EIC, an extension of AIC. Annals of the Institute of Statistical Mathematics, 49, 411–434.

Kane, T. L. and B. G. Brown, 2000: Confidence intervals for some verification measures - a survey of several methods. Preprints, 15th Conference on Probability and Statistics in the Atmospheric Sciences, 46–49.

- Schwarz, G., 1978: Estimating the dimension of a model. *The Annals of Statistics*, **6(2)**, 461–464.
- Shumway, R. H. and D. S. Stoffer, 2000: *Time series analysis and its applications*. Springer, 549 pp.
- Sugiura, N., 1978: Further analysis of the data by Akaike's information criterion and the finite corrections. Communication in Statistics Theory and Methods, 7(1), 13–26.

### 2.4 線形重回帰 1

## 2.4.1 はじめに

線形重回帰は、2018年現在の気象庁のガイダンスではあまり利用されていないものの、シンプルで最も基礎的な統計手法であり、正則化や説明変数の選択、利用上の注意点など線形重回帰に関連した周辺知識は、次節以降で述べるほかの統計手法を理解する上で有効である。本節では線形重回帰の基礎的な理論と周辺知識を述べる。

# 2.4.2 線形重回帰の基礎

線形重回帰は目的変数 y が説明変数 x と係数 w の線形結合で表されるとする回帰手法である。

$$y_n = w_0 + x_{n1}w_1 + \dots + x_{nK}w_K + \epsilon_n$$
  
=  $\sum_{k=0}^K x_{nk}w_k + \epsilon_n$  (2.4.1)

ここで、n は学習データの番号、k は説明変数の番号、K は説明変数の数、 $\epsilon_n$  は $x_{nk}w_k$  と独立な n 番目の観測  $y_n$  のノイズである。 $w_0$  はバイアス項で、 $x_{n0}=1$  とする。線形重回帰に関する多くのテキストでは  $w_kx_{kn}$  というように、係数、説明変数の順番で表記することが多いが、本章では (2.4.1) 式のように説明変数、係数の順番で記述する。線形重回帰では、N 個の学習データが与えられたときに、次の仮定の下で係数を推定する。

## 仮定 1. モデルの線形性

モデルが  $y_n = \sum_{k=0}^K x_{nk} w_k + \epsilon_n$  と表されること

仮定 2. 説明変数は非確率変数

仮定 3. 観測ノイズの不偏性

 $E(\epsilon_n) = 0$  であること

仮定 4. 観測ノイズの等分散性

各事例 n に依らず  $V(\epsilon_n) = \sigma^2$  であること

仮定 5. 観測ノイズの独立性

事例  $n, m (n \neq m)$  に対して、

 $E(\epsilon_n \epsilon_m) = E(\epsilon_n) E(\epsilon_m)$  であること

仮定 6. 説明変数に完全な多重共線性はない

説明変数  $x_{nk}, x_{nl}$   $(k \neq l)$  に対して、 $x_{nk} \neq ax_{nl} + b$  (a, b は n に依らない定数 ) であること

仮定  $2\sim5$  は、観測ノイズの二乗和を最小にする w が係数の分散を最小にする不偏推定量 (最良線形不偏推定量)になる  $^2$  という、最小二乗原理 (ガウス・マルコフの定理、廣津 (1992) など)を満たすために必要な

仮定である。最小二乗原理に基づいて係数を決定する 方法を最小二乗法という。

後で利用するために、上記の  $\epsilon_n$  に対する仮定を  $y_n$  に対する仮定に書き換えておく。仮定 2,3 と、係数 w は何らかの方法で確定させた値(非確率変数)であることより、

$$E(y_n) = \sum_{k=0}^{K} x_{nk} w_k \tag{2.4.2}$$

が成り立つ。また、 $\epsilon_n$  は  $x_{nk}w_k$  と独立であることから、 $V(y_n)=V\left(\sum_{k=0}^K x_{nk}w_k\right)+V\left(\epsilon_n\right)$  となる。ここで x と w が非確率変数であることより  $V\left(x_{nk}w_k\right)=0$ なので、仮定 4 を用いると、

$$V(y_n) = \sigma^2 \tag{2.4.3}$$

が成り立つ。また、 $n \neq m$  のとき、仮定 5 などを用いることで、

$$Cov(y_n, y_m) = 0 (2.4.4)$$

が成り立つ。

以下では最小二乗法を用いて係数 w を決定する方法を示す。目的変数  $y_n$  の推定値  $\hat{y}_n$  は

$$\hat{y}_n = E(y_n) = \sum_{k=0}^K x_{nk} w_k \tag{2.4.5}$$

であり、観測値  $y_n$  と推定値  $\hat{y}_n$  の差(残差)の二乗和 に 1/2 を掛けた値を誤差関数 E とする。

$$E = \frac{1}{2} \sum_{n=1}^{N} \epsilon_n^2$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left( y_n - \sum_{k=0}^{K} x_{nk} w_k \right)^2$$
(2.4.6)

最小二乗法では観測ノイズの二乗和を最小にする係数、 すなわち上記の E が最小になる時の係数を求めたいの で、E を係数の各成分  $w_k$  で微分する。

$$\frac{\partial E}{\partial w_k} = -\sum_{n=1}^{N} \left( y_n - \sum_{i=1}^{K} x_{ni} w_i \right) x_{nk} \tag{2.4.7}$$

これが0になるときのwを係数の推定値 $\hat{w}$ とすると、

$$\sum_{n=1}^{N} \left( x_{nk} y_n - x_{nk} \sum_{i=1}^{K} x_{ni} \hat{w}_i \right) = 0$$
 (2.4.8)

となる。これを行列形式で書くと次のようになる。

$$X^T \mathbf{y} - X^T X \hat{\mathbf{w}} = 0 \tag{2.4.9}$$

ここで X は  $x_{nk}$  を n, k 成分とする N 行 K 列の行列 で、 $X^T$  は X の転置行列である。X は正方行列ではな

<sup>1</sup> 工藤 淳

 $<sup>^2</sup>$  観測ノイズの二乗和を最小にする w が最適な推定値であることは自明ではない。例えば絶対値の和や四乗の和を最小にする w でも良いように思える。しかしそうではなく、二乗和を最小にする w が最適な推定値であるという定理がガウス・マルコフの定理である。

いため逆行列を持たないが、 $X^TX$  は K 行 K 列の正方行列であるため逆行列を持ちうる。(2.4.9) 式を  $\hat{w}$  について解くと、

$$\hat{\boldsymbol{w}} = \left(X^T X\right)^{-1} X^T \boldsymbol{y} \tag{2.4.10}$$

となる。このとき、説明変数間に完全な多重共線性があると X がランク落ちしして  $X^TX$  が逆行列を持てなくなる ((2.4.10) 式の解が一意に決まらなくなる) ため、前述の仮定 6 が必要になる。(2.4.10) 式を計算することで線形重回帰の係数の推定値  $\hat{w}$  が得られる。この  $\hat{w}$  を用いて、

$$\hat{\mathbf{y}} = X\hat{\mathbf{w}} 
= X (X^T X)^{-1} X^T \mathbf{y} 
\equiv H \mathbf{y}$$
(2.4.11)

と書いたとき、H をハット行列と呼ぶ。ハット行列の 転置は、

$$H^{T} = \left[ X (X^{T} X)^{-1} X^{T} \right]^{T}$$
$$= X (X^{T} X)^{-1} X^{T} = H$$
(2.4.12)

であることから、ハット行列は対称行列である。

# 2.4.3 最小二乗法と最尤法の関係

第2.3.9 項では、最尤法で係数が推定できることを述べた。線形重回帰は最小二乗法で係数を推定するが、最尤法を用いても同じ結果が得られることを以下で示す。 線形重回帰では観測ノイズの分布に対して、

- 不偏性:  $E(\epsilon_n) = 0$
- 等分散性: n に依らず  $V(\epsilon_n) = \sigma^2$
- 独立性:  $n \neq m$  のとき  $E(\epsilon_n \epsilon_m) = E(\epsilon_n) E(\epsilon_m)$

を仮定している。このような  $\epsilon_n$  の分布として、正規分布  $N(0,\sigma^2)$  を考える。今、 $y_n=\sum_{k=0}^K x_{nk}w_k+\epsilon_n$  であり、w と  $x_n$  が非確率変数として与えられたならば、 $\epsilon_n$  は w,  $x_n$  と独立であるから、 $\epsilon_n|w,x_n\sim N(0,\sigma^2)$  である。 $\epsilon_n$  を  $y_n-\sum_{k=0}^K x_{nk}w_k$  で置き換え、非確率変数  $\sum_{k=0}^K x_{nk}w_k$  を右辺に移項すれば、

$$y_n | \boldsymbol{w}, \boldsymbol{x}_n \sim N\left(\sum_{k=0}^K x_{nk} w_k, \sigma^2\right)$$
 (2.4.13)

となる。よって $y_n$ の確率密度関数は、

$$p(y_n|\boldsymbol{w},\boldsymbol{x}_n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} \left(y_n - \sum_{k=0}^K x_{nk} w_k\right)^2\right]$$
(2.4.14)

であり、観測ノイズが独立であることから、対数尤度は

$$\ln L(\boldsymbol{w}) = \sum_{n=1}^{N} \ln p(y_n | \boldsymbol{w}, \boldsymbol{x}_n)$$

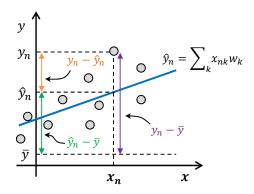


図 2.4.1 実況値  $y_n$ 、予測値  $\hat{y}_n$ 、実況の標本平均値  $\bar{y}$  の関係

$$= -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=1}^{N} \left(y_n - \sum_{k=0}^{K} x_{nk} w_k\right)^2$$
(2.4.15)

となる。N と  $\sigma^2$  は定数なので、(2.4.15) 式を最大にする w は  $\sum_{n=1}^N \left(y_n - \sum_{k=0}^K x_{nk} w_k\right)^2$  を最小にする w であることから、最尤法を用いた場合も最小二乗法を用いた場合と同じ結果が得られることになる。

# 2.4.4 決定係数

線形重回帰において、モデルの適合性を表す指標の一つに決定係数  $R^2$  がある。一般に用いられている決定係数の定義は下記の通りである。

$$R^{2} \equiv \frac{\sum_{n=1}^{N} (\hat{y}_{n} - \bar{y})^{2}}{\sum_{n=1}^{N} (y_{n} - \bar{y})^{2}}$$
(2.4.16)

ここで  $\hat{y}_n$  と  $\bar{y}$  はそれぞれ、 $y_n$  の推定値(予測値)と標本平均である。決定係数の平方根 R は重相関係数と呼ばれる。 $y_n$ ,  $\hat{y}_n$ ,  $\bar{y}$  の関係を図 2.4.1 に示す。もし回帰式によって実況値が完全に再現できた場合には、 $y_n=\hat{y}_n$  であるから、 $R^2=1$  となる。逆に、説明変数を全く用いなかった場合には回帰式は  $\hat{y}_n=\bar{y}$  となるため、 $R^2=0$  となる。すなわち決定係数は、いくつかの説明変数を用いて線形重回帰を行った場合、回帰式がどれくらい実況値を再現するかという指標である。決定係数の最大値は 1 で、1 に近いほどモデルが実況に適合しているといえる。(2.4.16) 式は以下のように変形できる ( 付録 2.4.A を参照 )。

$$R^{2} = 1 - \frac{\sum_{n=1}^{N} (y_{n} - \hat{y}_{n})^{2}}{\sum_{n=1}^{N} (y_{n} - \bar{y})^{2}}$$
(2.4.17)

この式の右辺第 2 項の分子は残差の二乗和である。残差の二乗和を最小にすることは  $R^2$  を最大にすることと等しいので、最小二乗法では  $R^2$  を最大にするように係数が決められるといえる。

(2.4.17) 式の右辺第 2 項の分母分子を N で割ると、

$$R^{2} = 1 - \frac{\frac{1}{N} \sum_{n=1}^{N} (y_{n} - \hat{y}_{n})^{2}}{\frac{1}{N} \sum_{n=1}^{N} (y_{n} - \bar{y})^{2}}$$
(2.4.18)

である。右辺第 2 項の分母は定義より  $y_n$  の標本分散である。また、(2.4.5) 式より  $\hat{y}_n = E(y_n)$  であるから、 $y_n - \hat{y}_n$  の標本平均は 0 であることを考慮すれば、右辺第 2 項の分子は  $y_n - \hat{y}_n$  の標本分散を表していることになる。サンプル数 N が限られている場合には、標本分散ではなく不偏分散を用いるように、決定係数の計算時にもサンプル数で割るのではなく、サンプル数から自由度(パラメータ数)を引いた数で割る。すなわち、

$$R_{adj}^{2} = 1 - \frac{\frac{1}{N-K-1} \sum_{n=1}^{N} (y_n - \hat{y}_n)^2}{\frac{1}{N-1} \sum_{n=1}^{N} (y_n - \bar{y})^2}$$
(2.4.19)

とし、 $R^2_{adj}$  を自由度調整済み決定係数という。

### 2.4.5 係数の推定精度

線形重回帰に限ったことではないが、少ない学習データを用いて学習した場合と多数の学習データを用いた場合とでは、係数の推定精度に差が出るであろうことは容易に想像できる。ここでは係数の推定精度に何が影響するかを示す。

係数の推定精度への影響を調べるため、最小二乗法で推定した係数  $\hat{w}$  の分散共分散行列を求める。A が非確率変数からなる行列、x を確率変数からなるベクトルとしたとき、 $V(Ax)=AV(x)A^T$  であることを利用すると、(2.4.10) 式より、

$$V\left(\hat{\boldsymbol{w}}\right) = \left(X^{T}X\right)^{-1} X^{T} V(\boldsymbol{y}) X \left(X^{T}X\right)^{-1} \quad (2.4.20)$$

となる。 $V\left(\hat{w}_{k}\right)$  は  $V\left(\hat{w}\right)$  の k 行 k 列成分  $V\left(\hat{w}\right)_{kk}$  であることから、

$$V\left(\hat{w}_{k}\right) = \sum_{i,j} \left[ \left(X^{T}X\right)^{-1}X^{T}\right]_{ki} V(\boldsymbol{y})_{ij} \left[X\left(X^{T}X\right)^{-1}\right]_{jk}$$

$$(2.4.21)$$

と書ける。ここで、(2.4.3) 式と (2.4.4) 式より、 $V({m y})_{ij}=\sigma^2\delta_{ij}$  なので、

$$V(\hat{w}_k) = \sigma^2 \left[ \left( X^T X \right)^{-1} X^T X \left( X^T X \right)^{-1} \right]_{kk}$$
$$= \sigma^2 \left[ \left( X^T X \right)^{-1} \right]_{kk}$$
(2.4.22)

となる。続いて  $\left(X^TX\right)^{-1}$  の具体的な値を求めてみるのだが、ここでは簡単のために説明変数が 2 個だけでバイアス項のないモデル

$$y_n = x_{n1}w_1 + x_{n2}w_2 + \epsilon_n$$

$$\sum_{n=1}^{N} x_{n1} = \sum_{n=1}^{N} x_{n2} = 0$$
(2.4.23)

を考えると、

$$X^{T}X = \begin{pmatrix} x_{11} & \cdots & x_{N1} \\ x_{12} & \cdots & x_{N2} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ \vdots & \vdots \\ x_{N1} & x_{N2} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{n} x_{n1}^{2} & \sum_{n} x_{n1} x_{n2} \\ \sum_{n} x_{n2} x_{n1} & \sum_{n} x_{n2}^{2} \end{pmatrix}$$
 (2.4.24)

であるので、

$$(X^{T}X)^{-1} = \frac{1}{D} \begin{pmatrix} \sum_{n} x_{n2}^{2} & -\sum_{n} x_{n1} x_{n2} \\ -\sum_{n} x_{n2} x_{n1} & \sum_{n} x_{n1}^{2} \end{pmatrix}$$
$$D = \sum_{m,n} x_{m1}^{2} x_{n2}^{2} - \sum_{m,n} x_{m1} x_{m2} x_{n1} x_{n2}$$
$$(2.4.25)$$

となることから、例えば $w_1$ の推定値の分散は、

$$V(\hat{w}_1) = \frac{\sigma^2}{D} \sum_{n=1}^{N} x_{n2}^2$$

$$= \frac{\sigma^2}{N} \frac{1}{\frac{1}{N} \sum_{n=1}^{N} x_{n1}^2} \frac{1}{1 - \rho_{12}^2}$$
(2.4.26)

$$\rho_{12}^2 \equiv \frac{\sum_{m,n} x_{m1} x_{m2} x_{n1} x_{n2}}{\sum_{m,n} x_{m1}^2 x_{n2}^2}$$

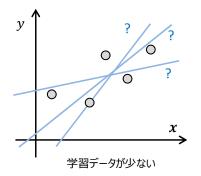
となる。ここで  $\rho_{12}$  は説明変数  $x_1,x_2$  間の標本相関係数である。また、 $\sum_{n=1}^N x_{n1}=0$  であることから、 $\frac{1}{N}\sum_{n=1}^N x_{n1}^2$  は  $x_1$  の標本分散である。 $V\left(\hat{w}\right)$  が大きい場合、学習データのサンプリング方法が少し変わったり、学習データの一部に誤差の大きなデータが含まれるだけで推定値が大きく変わる可能性がある。これは係数の推定精度が低いことを意味する。(2.4.26) 式より、係数の推定値の分散を大きくする要因は以下の 4 点である。

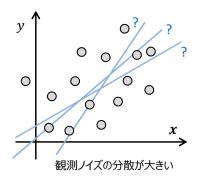
- N が小さい(学習データが少ない)
- $\sigma^2$  が大きい (観測ノイズの分散が大きい)
- $\frac{1}{N}\sum_{n=1}^N x_{n1}^2$  が小さい(説明変数の分散が小さい)
- $\rho_{12}^2$  が 1 に近い(説明変数間の相関係数が 1 に近い)

これらのうち始めの 3 点が要因で係数推定の精度が悪くなる例を図 2.4.2 に示す。このような場合には、学習データを増やす、実況と相関が強い説明変数を導入する、学習データのサンプリングの妥当性を検討するなどの対応が必要となる。次に 4 点目の要因について考える。今、2 つの説明変数  $x_1$ と  $x_2$  に強い多重共線性があり、 $x_{n1} \approx ax_{n2}$  (a は定数 )と書けるとすると、 $\rho_{12}^2 \approx 1$  となることから、説明変数に強い多重共線性がある場合も係数の推定精度が悪くなるといえる。ガイダンスの開発を行う場合、一般に説明変数間にはある程度の多重共線性がある。このこと自体は大きな問題ではないが、多重共線性が強い場合には係数の推定精度が悪くなることに注意しなければならない。

### 2.4.6 正則化法

説明変数間に強い多重共線性がある場合でも係数を もっともらしく推定する方法の一つがここで述べる正 則化法である。正則化法は線形重回帰に限らず様々な 統計手法で利用される。





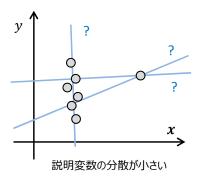


図 2.4.2 係数の推定精度が悪くなる場合の例

ここでは簡単のために、説明変数が 2 つでバイアス 項が 0 のモデル (2.4.23) 式を考える。説明変数に強い 多重共線性があり、 $x_{n1} \approx x_{n2}$  とすると、

$$\hat{y}_n = x_{n1}$$

$$\hat{y}_n = 2x_{n1} - x_{n2}$$

$$\hat{y}_n = 100x_{n1} - 99x_{n2}$$
(2.4.27)

はいずれもほぼ同じ結果を与える。十分な学習データが与えられれば 1 番目の回帰式が得られることが期待されるが、学習データが不十分な場合には係数の推定精度が低くなり、2 番目や 3 番目のような回帰式が得られる可能性がある。そして 3 番目の式のように係数の絶対値が大きくなればなるほど、2 つの説明変数のバランスがわずかに崩れただけでも予測値が大きく変わってしまい、予測が不安定になる 3 。

線形重回帰では (2.4.6) 式で表される誤差関数を最小にする係数を求めるのだが、誤差関数に w の大きさに応じたペナルティー項を加えることを考える。

$$E = \frac{1}{2} \sum_{n=1}^{N} \left( y_n - \sum_{k=0}^{K} x_{nk} w_k \right)^2 + \frac{\lambda}{\beta} \sum_{k=0}^{K} |w_k|^{\beta} \quad (2.4.28)$$

ここで  $\lambda$  は正の定数で、  $\beta$  は 1, 2 などがよく用いられる。このように誤差関数にペナルティー項を加えることを正則化といい、  $\beta=1$  の場合を L1 正則化、  $\beta=2$  の場合を L2 正則化という  $^4$ 。線形重回帰の場合には、 L1 正則化を Lasso (Tibshirani 1996)、 L2 正則化をリッジ回帰 (Horel and Kennard 1970) という。いずれの場合も、平均二乗誤差が同程度である場合には係数の大きさの和が小さいほど良いモデルとして採用されるた

め、(2.4.27) 式の 2 番目や 3 番目のような回帰式は得られにくくなる。

正則化は説明変数の数が多い場合に特に有効である。第 2.3.11 項で述べたように、説明変数が多いとモデルの表現力(複雑な関係性を表現できる能力)が高くなるため、学習データに対して過剰に最適化した係数が得られてしまう。この場合、未知のデータに対する誤差が大きくなり、ガイダンスの精度は低下する。多すぎる説明変数を減らす方法としては、説明変数間の相関係数に基づいて開発者が減らす方法や、次項で述べる手法がよく用いられるが、正則化法を用いることで強い多重共線性を持つ説明変数が選ばれにくくなるため、説明変数の選択を効率的に行うことができるようになる。

# 2.4.7 説明変数の選択

線形重回帰において、説明変数を選択する方法としては主に以下の3つが挙げられる。

- 赤池情報量基準 (AIC) を用いる方法
- 交差検証 (CV) を用いる方法
- 主成分分析を用いる方法

ここでは線形重回帰における AIC と CV を用いた説明 変数の選択について述べる。主成分分析については統計手法によらない手法であるため、第 2.3.11 項 (3) を参照していただきたい。

## (1) 赤池情報量基準を用いる方法

線形重回帰での AIC を求める。第 2.4.3 項と同様に、ここでも観測ノイズの分布として正規分布  $\epsilon_n \sim N(0,\sigma^2)$  を考える。このとき、確率密度関数と対数尤度は、

$$p(\epsilon_n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon_n^2}{2\sigma^2}\right)$$

$$\ln L = \sum_{n=1}^N \ln\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon_n^2}{2\sigma^2}\right)\right]$$

$$= -\frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \sigma^2 - \sum_{n=1}^N \frac{\epsilon_n^2}{2\sigma^2} \quad (2.4.29)$$

 $<sup>^3</sup>$  このため、例えば  $\hat{y}_n=0.5x_{n1}+0.5x_{n2}$  というような関係が得られた場合には、予測結果の解釈が難しくなるという問題はあるが、予測の安定性に関しては大きな問題は生じない。

 $<sup>^4</sup>$  L1, L2 のほかにも、0 ではない係数の数に比例したペナルティー項を与える L0 正則化や、絶対値が最大となる係数に比例したペナルティー項を与える L $\infty$  正則化、L1 と L2 を組み合わせた Elastic net など、様々な正則化法が提案されている。

となる。ここで観測ノイズの分散  $\sigma^2$  は通常は未知のパラメータであるため、学習データから  $\sigma^2$  の推定値  $\hat{\sigma}^2$  を最尤法で求めることにする。すなわち、上記の対数 尤度を  $\sigma^2$  の関数と見なし、 $\ln L\left(\sigma^2\right)$  を  $\sigma^2$  で微分して 0 になるときの  $\sigma^2$  を分散の推定値  $\hat{\sigma}^2$  とする。

$$\frac{\partial \ln L}{\partial \sigma^2} \bigg|_{\sigma^2 = \hat{\sigma}^2} = -\frac{N}{2\hat{\sigma}^2} + \sum_{n=1}^N \frac{\epsilon_n^2}{2(\hat{\sigma}^2)^2} = 0 \quad (2.4.30)$$

よって、

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{n=1}^{N} \epsilon_n^2 = \frac{\text{RSS}}{N}$$
 (2.4.31)

となる。ここで RSS は残差の二乗和 (Residual Sum of Square) である。この  $\hat{\sigma}^2$  を用いて対数尤度を書くと、

$$\ln L(\hat{\sigma}^2) = -\frac{N}{2} \ln 2\pi - \frac{N}{2} \ln \frac{RSS}{N} - \frac{N}{2}$$
 (2.4.32)

であるので、AIC は、

AIC = 
$$N \ln 2\pi + N \ln \frac{RSS}{N} + N + 2K$$
 (2.4.33)

となる。AIC を用いて説明変数を選択する場合、予測式を変える前と後でのAIC の差に着目することになるため、定数部分は無視してよい。学習データのサンプル数N はここでは定数であるから、右辺の第1 項と第3 項はAIC の計算には不要となる。よって線形重回帰の場合のAIC は、

$$AIC = N \ln \frac{RSS}{N} + 2K \tag{2.4.34}$$

と書くことができる。学習データに対して、AIC が最も小さくなる説明変数の組み合わせを総当たり法や変数増減法などで選択する。

# (2) 交差検証を用いる方法

CV は繰り返し計算を行うため計算に時間が掛かるという問題があるが、線形重回帰に用いる場合には計算を効率的に行う方法がある。LOOCV で平均二乗誤差 MSE を求めると、

$$MSE_{LOOCV} = \frac{1}{N} \sum_{n=1}^{N} \left( y_n - \hat{f}^{-n}(\boldsymbol{x}_n) \right)^2 \quad (2.4.35)$$

と書ける ((2.3.60) 式を再掲)。ここで、 $\hat{f}^{-n}$  は n 番目 の学習データを除いて作成した回帰式である。 煩雑なため証明は省略するが、(2.4.35) 式を変形すると、

$$MSE_{LOOCV} = \frac{1}{N} \sum_{n=1}^{N} \frac{\left(y_n - \hat{f}(\boldsymbol{x}_n)\right)^2}{\left(1 - [H]_{nn}\right)^2}$$
 (2.4.36)

と近似ではなく厳密に書ける。ここで  $[H]_{nn}$  は (2.4.11) 式で示したハット行列の n 行 n 列成分、 $\hat{f}$  は N 個の学習データを全て用いて作成した回帰式である。(2.4.36)

式では、 $\hat{f}^{-n}$  ではなく  $\hat{f}$  となっている所がポイントで、(2.4.35) 式ではデータを一つ抜いて回帰式を作成するということを N 回繰り返す必要があったのに対し、(2.4.36) 式では N 個のデータを全て用いた通常の回帰式の作成を 1 回だけ実行すれば良いため高速に計算できる。この方法は線形重回帰の場合に限られるが、学習データの数が多い場合には特に有効である。この計算を総当たり法や変数減少法等を用いて説明変数の組み合わせを変えながら実行し、 $MSE_{LOOCV}$  が最も小さくなる説明変数の組み合わせを選ぶ。

### 2.4.8 逐次学習

線形重回帰では (2.4.10) 式を用いた一括学習により係数を決定する場合が多いが、第 2.3.10 項で述べた確率的勾配降下法を用いることで逐次学習することも可能である。線形重回帰の誤差関数 E は残差の二乗和であり、これを係数 w の各成分で微分すると (2.4.7) 式になることから、時刻 t における係数  $w^{(t)}$  を逐次更新する式は (2.3.56) 式より、

$$w_k^{(t+1)} = w_k^{(t)} + \eta \ x_{tk} \left( y_t - \sum_{l=0}^K x_{tl} \ w_l^{(t)} \right) \quad (2.4.37)$$

となる。ここで  $y_t$ と  $x_t$  は時刻 t における観測値と説明変数である。第 2.3.10 項でも述べたように、逐次学習する場合には、学習率  $\eta$  の値を適切に調整し、説明変数を規格化して説明変数のオーダーを揃えておく必要がある。

## 2.4.9 線形重回帰の利用上の注意点

線形重回帰は最もシンプルな統計手法であり広く利用されているが、利用に当たっては注意すべき点が多い。ここでは例を示しながら線形重回帰を利用する上での注意点を述べる。

# (1) 線形重回帰の仮定に反していないか

第 2.4.2 項で示したように、線形重回帰では以下の 6 つの仮定の下で回帰式が決定される。

仮定 1. モデルの線形性

仮定 2. 説明変数は非確率変数

仮定3. 観測ノイズの不偏性

仮定 4. 観測ノイズの等分散性

仮定 5. 観測ノイズの独立性

仮定 6. 説明変数に完全な多重共線性はない

ガイダンスでは、数値予報モデルの予測値などの確定した値を用いて説明変数を作成するため、通常は仮定2は満たされていると考えて問題ないだろう。また、説明変数間にはある程度の相関は必ずあるものの、例えばある説明変数を2倍にした変数を説明変数に加える等、開発者が意図的に操作しない限りは説明変数が完全な多重共線性を持つことはないので、通常は仮定6も満たされていると考えて問題はない。残りの4つの

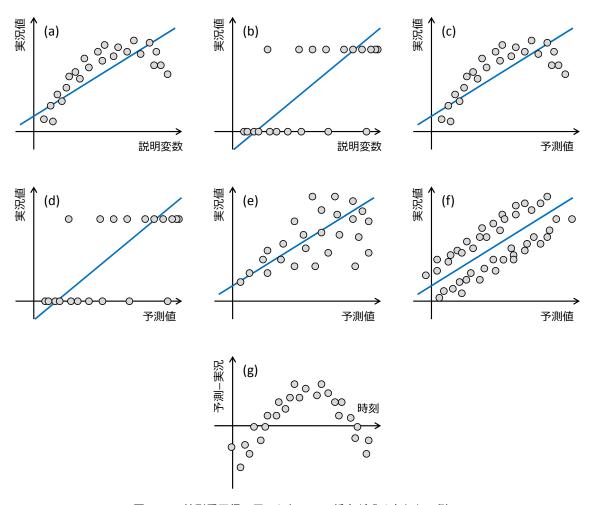


図 2.4.3 線形重回帰で用いられている仮定が成り立たない例

仮定に関しては、用いているデータが仮定を満たしているか否かは自明ではないため、開発を行うに当たってデータの性質を確かめておく必要がある。

データの性質を簡単に確かめる方法としてはデータ を実際にプロットしてみると良い。まず、仮定1のモデ ルの線形性については、実際にデータをプロットする ことで、目的変数(実況値)と説明変数に線形関係があ るか否かを確かめる。このとき、個別の説明変数が実 況値と線形関係を持たなければならないことに注意が 必要である。各説明変数と実況値をプロットしたとき、 図 2.4.3(a) のような関係にある場合には線形関係を満 たさないため、このままでは線形重回帰の説明変数に 用いることはできない。このような場合には、説明変 数を何らかの関数で変換することで線形化できる場合 がある。図 2.4.3(b) のような場合には、ロジスティック 回帰等のほかの統計手法を用いることを検討する。ま た、例えばある変数 u を説明変数として用いる場合、 u の 2 乗や対数など u の何らかの関数 f(u) を説明変 数に用いてはならない。なぜならば、もしuが実況値 と線形関係にあるならば、u の線形変換 f(u) = au + b (a,b) は定数)を除いては f(u) は実況と非線形関係となるため仮定 1 に反することになり、f(u)=au+b とした場合には u と完全な多重共線性を持つため仮定 6 に反するからである。逆に、もし  $f(u) \neq u$  が実況値と線形関係にあるならば、u 自身は実況値と非線形関係を持つか f(u) と完全な多重共線性を持つため、u を説明変数に用いることはできない。よって線形重回帰では、もし f(u) を説明変数に用いた場合、u の別の関数g(u) を説明変数に用いることはできない。

仮定3の観測ノイズの不偏性の仮定が満たされているか否かは、ガイダンスの予測値と実況値の関係をプロットした時に、予測値に依らず予測と実況の差が一定と見なせるかどうかを確かめることになる。図2.4.3(c)の場合は、全体としてはバイアスが0だったとしても、予測値の大きさによってバイアスが変化しているため、観測ノイズの不偏性は満たされていない。このような場合には、予測と実況が線形になるように説明変数の見直しを行うか、季節や予測値などで層別化するか、別の統計手法を用いるなどの検討が必要になる。図2.4.3(d)は2値データを線形モデルで予測しようとした例であ

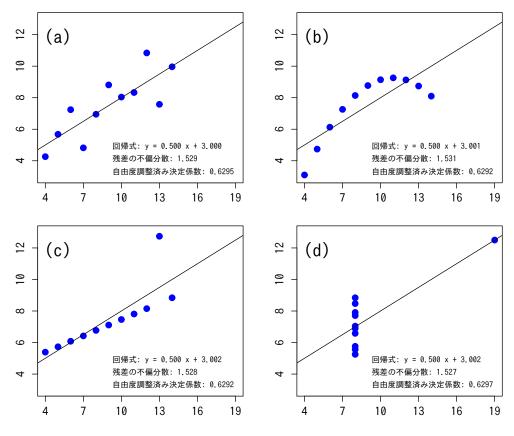


図 2.4.4 Anscombe (1973) の例

り、観測ノイズの不偏性は満たされていない。このような場合には、線形重回帰ではなくロジスティック回帰を用いるなどを検討する必要がある。

仮定 4 の観測ノイズの等分散性は、各データについて予測と実況の差の分散が一定であることを意味している。図  $2.4.3(c)\sim(e)$  はいずれもこの等分散性の仮定を満たしていないため、説明変数や統計手法の見直しが必要となる。

仮定 5 の観測ノイズの独立性が成り立たない場合、すなわち観測ノイズに相関があるのは、空間方向や時間方向に相関がある場合である。例えば、バイアス特性が異なる 2 つの地点の気温をまとめて予測しようとした場合には図 2.4.3(f) のようになるため、地点で層別化するか、地点特性を考慮した説明変数を導入するなどの検討が必要である。また、時間方向に誤差の相関がある場合には図 2.4.3(g) のようになるため、時刻や季節などで層別化したり、時刻や季節の特性を考慮した説明変数を導入するなどの検討が必要になる。

### (2) 回帰式の結果だけを見ていないか

回帰式を決定した後に、決定係数の値や回帰直線、 学習データや予測データに対する誤差といった回帰の 結果だけに注目するのではなく、予測値と実況値の関 係をプロットしてみることが重要である。

図 2.4.4 は Anscombe (1973) の例で、図中の点は回

帰式の作成に用いられたデータ、直線は回帰直線を表している。 $(a)\sim(d)$ のデータのうち、(a)以外のデータは線形重回帰を適用できないデータであることは図を見ればすぐに分かるが、図中に示した回帰式、残差の不偏分散、自由度調整済み決定係数は $(a)\sim(d)$ のいずれもほぼ同じ値になっており、これらの値を見ただけでは統計関係の正しさを知ることはできない。データと回帰直線の関係をプロットするほかにも、残差を予測値に対してプロットしたり、正規 Q-Q プロット (Wilks 1995 など)や Cook の距離 (Cook and Weisberg 1982)を利用したりすることで、線形重回帰に用いるデータの正しさを視覚的に確かめることができる。

### (3) 学習時、検証時の注意点

線形重回帰に限らず、多くの統計手法では学習に用いるデータは独立であることを仮定しているが、一定期間の連続した気象データを学習に用いる場合、学習データ間にはある程度の相関が含まれていることを考慮しなければならない。例えば時刻方向に回帰式を層別化していない場合には、同じ観測データに対して複数の時刻の予測値が得られるため、学習に使用する初期時刻や予報時間を限定するなどの工夫が必要となる。

検証では係数の学習に使用していないデータ(学習 データとは独立なデータ)を用いなければならない。 これは当然のことに思われるが、係数を決定すること

に加えて、説明変数の算出に利用するパラメータを調 整するような場合には注意が必要である。例えば3年 分のデータを用意して、初めの2年分のデータで係数 を学習し、残りの1年分のデータで検証することを考 える。これ自体はほぼ独立なデータを用いた検証とい えるのだが、検証した結果、期待していたような精度 が得られなかった場合、説明変数のパラメータを調整 して再び係数を学習して検証する、ということを繰り 返すことがある。この場合、回帰式と説明変数は用意 した3年分のデータにフィットするように調整されて しまうため、結果的に3年分のデータ全てを学習デー タとして扱ったことになってしまう。このような場合 にはさらに別途検証用のデータを用意する必要がある。 検証用のデータを別途用意できない場合には CV を用 いて精度を評価することになる。ただし第2.3.11項(2)でも述べたとおり、時間方向に相関を持つデータに CV を用いた場合には、実際の検証結果よりも誤差が小さ く見積もられることに注意が必要である。

### 2.4.10 まとめ

本節では、線形重回帰の手法とその周辺の技術及び利用上の注意点を述べた。線形重回帰は最もシンプルな手法であり様々に利用されているが、係数を決定する上で置かれている仮定が多く、これらが満たされていないデータに対しては、説明変数を見直したり、別の統計手法を利用するなどの検討が必要である。また、手法を正しく用いるためには、回帰の結果や検証結果だけに注目するのではなく、説明変数と実況値、予測値と実況値の関係などを実際にプロットしてみることが重要である。これは線形重回帰に限らず、全ての統計手法に共通していえることである。

付録 2.4.A 決定係数が (2.4.17) 式と書けることの 証明

ここでは、

$$\frac{\sum_{n=1}^{N} (\hat{y}_n - \bar{y})^2}{\sum_{n=1}^{N} (y_n - \bar{y})^2} = 1 - \frac{\sum_{n=1}^{N} (y_n - \hat{y}_n)^2}{\sum_{n=1}^{N} (y_n - \bar{y})^2}$$

と書けることを示す。左辺の分母は、

$$\sum_{n=1}^{N} (y_n - \bar{y})^2 = \sum_{n=1}^{N} (y_n - \hat{y}_n + \hat{y}_n - \bar{y})^2$$

$$= \sum_{n=1}^{N} (y_n - \hat{y}_n)^2 + \sum_{n=1}^{N} (\hat{y}_n - \bar{y})^2$$

$$+2\sum_{n=1}^{N} (y_n \hat{y}_n - \hat{y}_n \hat{y}_n) + 2 \bar{y} \sum_{n=1}^{N} (\hat{y}_n - y_n)$$
(2.4.38)

と書ける。ここで (2.4.38) 式の右辺第 3 項を行列形式 で書くと  $2\left( {{m y}^T}\hat{{m y}} - \hat{{m y}}^T\hat{{m y}} \right)$  であり、また、(2.4.11) 式より、 $\hat{{m y}} = H{m y}$  であることから、

$$\hat{\mathbf{y}}^{T}\hat{\mathbf{y}} = (H\mathbf{y})^{T}H\mathbf{y}$$

$$= \mathbf{y}^{T}HH\mathbf{y}$$

$$= \mathbf{y}^{T}X(X^{T}X)^{-1}X^{T}X(X^{T}X)^{-1}X^{T}\mathbf{y}$$

$$= \mathbf{y}^{T}H\mathbf{y}$$

$$= \mathbf{y}^{T}\hat{\mathbf{y}}$$
(2.4.39)

と書ける。ここで、ハット行列は対称行列であること (2.4.12) 式を用いた。これより (2.4.38) 式の右辺第 3 項は 0 になる。また (2.4.5) 式より  $\hat{y}_n = E(y_n)$  であるから、右辺第 4 項も 0 となる。以上より、

$$\sum_{n=1}^{N} (\hat{y}_n - \bar{y})^2 = \sum_{n=1}^{N} (y_n - \bar{y})^2 - \sum_{n=1}^{N} (y_n - \hat{y}_n)^2$$
(2.4.40)

であるから、これを (2.4.16) 式に代入すると (2.4.17) 式が得られる。

#### 参考文献

Anscombe, F. J., 1973: Graphs in statistical analysis. The American Statistician, 27(1), 17–21.

Cook, R. D. and S. Weisberg, 1982: Residuals and influence in regression. New York: Chapman and Hall, 229 pp.

廣津千尋, 1992: 線形モデルと最小二乗法. 自然科学の 統計学 第2章, 東京大学出版会, 26-78.

Horel, A. E. and R. W. Kennard, 1970: Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 12, 55–67.

Tibshirani, R., 1996: Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B*, **58(1)**, 267–288.

Wilks, D. S., 1995: Statistical methods in the atmospheric sciences. Academic press, 467 pp.

### 2.5 ロジスティック回帰 1

### 2.5.1 はじめに

前節で述べた線形重回帰は、気温のような連続値の予測に用いられる統計手法の一つである。これに対して本節で述べるロジスティック回帰は、例えば雷の有無などのように、現象を2つのクラスに分類する問題に用いられる統計手法の一つである。後述するように、ロジスティック回帰により得られる予測値は現象の発生確率と考えることができるため、ロジスティック回帰は確率型のガイダンスによく用いられている。気象庁でのロジスティック回帰の利用は発雷確率ガイダンスに始まり(高田 2007)、その後、乱気流指数や雲底確率ガイダンス、ガスト発生確率ガイダンスなどにも利用されるようになった。

本節では初めにロジスティック回帰の基礎的な理論について述べ、続いてロジスティック回帰を多クラス分類に適用した順序ロジスティック回帰と多項ロジスティック回帰について述べる。最後に利用上の注意点等を述べる。

### 2.5.2 ロジスティック回帰の基礎

ロジスティック回帰 (Cox 1958) は、目的変数が 2 値 データでベルヌーイ分布に従う場合に用いられる回帰 分析の一種  $^2$  である。2 値データとは、雷があった場合を 1、なかった場合を 0 としたような 2 つの値だけを取るデータのことである。ここでは初めに、確率がベルヌーイ分布で表される一般的な場合について考える。

ベルヌーイ分布とは、確率変数 Y が確率 p で現象あり (Y=1)、確率 1-p で現象なし (Y=0) を取る離散確率分布である。今、説明変数 x と係数 w が与えられているとする。このとき、現象ありとなる確率を  $P_r(Y=1|x,w)$  とし、

$$p \equiv P_r(Y = 1|\boldsymbol{x}, \boldsymbol{w}) \tag{2.5.1}$$

とすると、現象なしとなる確率は、

$$P_r(Y = 0|\mathbf{x}, \mathbf{w}) = 1 - p$$
 (2.5.2)

となる。ここで p は x と w の関数である。この 2 つをまとめると、ベルヌーイ分布の確率関数は以下のように書ける。

$$P_r(Y = y|x, w) = p^y (1-p)^{1-y}$$
 (2.5.3)

y を 2 値データの観測値とし、N 組の観測値と説明 変数のデータセット  $(y_1,x_1)$ ,  $\cdots$ ,  $(y_N,x_N)$  が与えられたとき、最適な係数 w を最尤法で求めることを考える。各観測が独立であると仮定すれば、尤度 L は係数

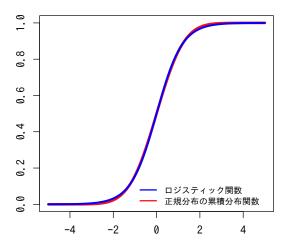


図 2.5.1 ロジスティック関数  $y=1/(1+\exp(-1.7x))$  と正規分布 N(0,1) の累積分布関数

ベクトルwの関数として、

$$L(\mathbf{w}) = \prod_{n=1}^{N} P_r(Y_n = y_n | \mathbf{x}_n, \mathbf{w})$$

$$= \prod_{n=1}^{N} p_n^{y_n} (1 - p_n)^{1 - y_n}$$
(2.5.4)

と書ける。よって対数尤度は次のように書ける。

$$\ln L(\boldsymbol{w}) = \sum_{n=1}^{N} [y_n \ln p_n + (1 - y_n) \ln(1 - p_n)]$$
 (2.5.5)

対数尤度が最大となる係数を求めたいので、対数尤度 w の各成分で微分する。

$$\frac{\partial \ln L}{\partial w_k} = \sum_{n=1}^{N} \left[ \frac{y_n}{p_n} \frac{\partial p_n}{\partial w_k} - \frac{1 - y_n}{1 - p_n} \frac{\partial p_n}{\partial w_k} \right] 
= \sum_{n=1}^{N} \frac{\partial p_n}{\partial w_k} \frac{y_n - p_n}{p_n (1 - p_n)}$$
(2.5.6)

ここまではベルヌーイ分布に従う統計モデルに対して一般的に成り立つ議論であり、確率  $p_n$  として具体的にどのような関数を用いるかによって統計手法が異なる。ロジスティック回帰では、 $p_n$  が次のロジスティック関数で表されるものと仮定する。

$$p_n = \frac{1}{1 + \exp\left(-\sum_{k=0}^K x_{nk} w_k\right)}$$
 (2.5.7)

ここで $w_0$  はバイアス項で、 $x_{n0}=1$  とする。ロジスティック関数は図 2.5.1 の青線のような関数である。(2.5.7) 式を用いると、

$$\frac{\partial p_n}{\partial w_k} = x_{nk} p_n \left( 1 - p_n \right) \tag{2.5.8}$$

<sup>1</sup> 丁藤 淳

<sup>&</sup>lt;sup>2</sup> ほかにも、プロビット回帰や complementary log-log 回帰などがある。

となることから、(2.5.6) 式は

$$\frac{\partial \ln L}{\partial w_k} = \sum_{n=1}^{N} x_{nk} (y_n - p_n)$$
 (2.5.9)

というシンプルな形で書くことができる。これが0になるときのwをニュートン・ラフソン法などを用いて数値的に求めれば、ロジスティック回帰における係数が求まる。求められた係数と予測データを(2.5.7)式に適用することで、ロジスティック回帰による予測値が得られる。ロジスティック回帰を2クラス分類問題に適用する場合には、pがある値(例えば0.5)以上であれば現象あり、それよりも小さければ現象なし、というように分類すれば良い。

ロジスティック関数は  $0\sim1$  の値を取る単調増加関数で、全域で微分可能であるという特性を持っていることから、確率を表現するのに適した関数であるといえる。このため、(2.5.7) 式では確率がロジスティック関数で書けると仮定したのだが、上記のような特性を持つ関数はロジスティック関数だけではない。例えば確率という意味でいえば、正規分布の累積分布関数(図2.5.1 の赤線)を用いても良さそうである。確率をこのように仮定した回帰をプロビット回帰といい、その確率密度関数は

$$p_n = \int_{-\infty}^{\sum_{k=0}^{K} x_{nk} w_k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (2.5.10)$$

である。しかし、(2.5.6) 式を求めるためには (2.5.10) 式を  $w_k$  で微分する必要があるため扱いが困難である。これに対してロジスティック関数は、(2.5.8) 式のように微分が容易にでき、対数尤度の微分もシンプルな形で記述できる。またパラメータを適当に調整すれば、図 2.5.1 のように正規分布の累積分布関数とほとんど同じ曲線になる。このような理由から、ロジスティック関数は 2 値データの回帰にしばしば用いられている。

# 2.5.3 正則化法、説明変数の選択、逐次学習

ロジスティック回帰の場合も線形重回帰と同様に、正 則化、説明変数の選択、逐次学習といった手法を用い ることができる。これらについて以下で述べる。

# (1) 正則化法

ロジスティック回帰における誤差関数 E は負の対数 尤度  $-\ln L$  である。これに係数の大きさに依存したペナルティー項を加えることで多重共線性の影響を減らすことができる。

$$E = -\ln L + \frac{\lambda}{\beta} \sum_{k=0}^{K} |w_k|^{\beta}$$
 (2.5.11)

 $\lambda$  は正の正則化定数で、線形重回帰と同様に、 $\beta$  が 1 の場合は L1 正則化、2 の場合は L2 正則化となる。

### (2) 説明変数の選択

ロジスティック回帰の場合も AIC や CV、主成分分析により説明変数を選択することができる。 AIC を用いる場合には対数尤度の式 (2.5.5) をそのまま用いれば良い。

AIC = 
$$-2 \ln L + 2K$$
  
=  $-2 \sum_{n=1}^{N} [y_n \ln p_n + (1 - y_n) \ln(1 - p_n)] + 2K$   
(2.5.12)

この式を用いて、AIC が最小になる説明変数の組み合わせを選択する。

CV を用いる場合は、例えばブライアスコアを基準 にすることになる。LOOCV を利用するならば、

$$BS_{LOOCV} = \frac{1}{N} \sum_{n=1}^{N} (y_n - \hat{p}^{-n}(\boldsymbol{x}_n))^2 \qquad (2.5.13)$$

となる。ここで  $\hat{p}^{-n}(x_n)$  は、n 番目のデータを除いて作成したロジスティック回帰に  $x_n$  を適用した場合の予測値である。

主成分分析は説明変数間の分散共分散または相関係数のみで決まり、目的変数には依らないため、統計手法に関わらず第 2.3.11 項 (3) で述べた手法で説明変数を選択できる。

# (3) 逐次学習

ロジスティック回帰における誤差関数 E は負の対数 尤度  $-\ln L$  であるため、誤差関数を w で微分した値は (2.5.9) 式に負号を付けたものになる。

$$\frac{\partial E}{\partial w_k} = -\sum_{n=1}^{N} x_{nk} (y_n - p_n)$$
 (2.5.14)

これを (2.3.56) 式に代入すれば、ロジスティック回帰における逐次学習の式が得られる。

$$w_k^{(t+1)} = w_k^{(t)} + \eta x_{tk} \left( y_t - p(\boldsymbol{x}_t, \boldsymbol{w}^{(t)}) \right)$$
 (2.5.15)

第 2.3.10 項で述べたように、係数を逐次学習する場合には、学習率  $\eta$  を適切に調整するとともに、説明変数を規格化して説明変数のオーダーを揃えておく必要がある。

### 2.5.4 順序ロジスティック回帰

順序ロジスティック回帰 (McCullagh 1980) は、目的変数 y が 0, 1, 2,  $\cdots$  のように順序のある階級に分かれている場合のロジスティック回帰である。例えば降水強度を弱、並、強に階級分けするような場合に用いることができる。

ここでは階級 m が  $0, 1, \dots, M-1$  の M 個のクラスに分かれているとする。順序ロジスティック回帰

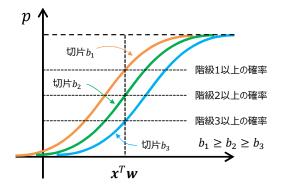


図 2.5.2 順序ロジスティック回帰の概念図

では、n 番目の実況に対応する確率変数  $Y_n$  が m 番目以上の階級である確率は

$$P_r(Y_n \ge m | \boldsymbol{x}_n, \boldsymbol{w}) = \frac{1}{1 + \exp\left(-b_m - \sum_{k=1}^K x_{nk} w_k\right)}$$
(2.5.16)

であると仮定する。ここで  $b_m$  は m 番目以上の階級 である確率に対するロジスティック関数の切片である。 (2.5.7) 式では切片を和に含めているが、上式では和の 外に出して記述している。(2.5.16) 式では、階級の違 いは切片のみであり、説明変数に掛かる係数は全ての 階級で等しいと仮定している。例えば階級が $0 \sim 3$ の 4階級で分けられる場合を図で表すと図 2.5.2 のように なる。1番目の曲線(左の曲線)は階級が1以上か1 未満か(すなわち階級0か)を分ける曲線で、切片は  $b_1$  である。2 番目の曲線(中央の曲線)、3 番目の曲線 (右の曲線)も同様であり、 $b_1 \geq b_2 \geq b_3$ である。説明 変数 x と係数 w が与えられたとき、階級が 1, 2, 3 以 上である確率は図に示したようになる。4 つの階級は 3本のロジスティック関数で分けられ、各曲線の傾きは 全て等しく、切片の大きさによって線が横軸方向に移 動している。傾きを等しくすることで、確率の逆転が 生じることがなくなる。

それぞれの階級になる確率は (2.5.16) 式の差で表すことができる。

$$P_r(Y_n = m | \boldsymbol{x}_n, \boldsymbol{b}, \boldsymbol{w})$$

$$= P_r(Y_n \ge m | \boldsymbol{x}_n, \boldsymbol{b}, \boldsymbol{w}) - P_r(Y_n \ge m + 1 | \boldsymbol{x}_n, \boldsymbol{b}, \boldsymbol{w})$$

$$= \frac{1}{1 + \exp\left(-b_m - \sum_{k=1}^K x_{nk} w_k\right)}$$

$$- \frac{1}{1 + \exp\left(-b_{m+1} - \sum_{k=1}^K x_{nk} w_k\right)}$$
(2.5.17)

ここで、 $b_0=+\infty, b_M=-\infty$  とする。N 組の学習データ  $({m y}_1,{m x}_1),\ \cdots,\ ({m y}_N,{m x}_N)$  が与えられたとき、各観測が独立であると仮定すれば、対数尤度は切片  ${m b}$  と係数

w の関数として、

$$\ln L(\boldsymbol{b}, \boldsymbol{w}) = \sum_{n=1}^{N} \ln P_r(Y_n = m | \boldsymbol{x}_n, \boldsymbol{b}, \boldsymbol{w}) \quad (2.5.18)$$

となる。(2.5.18) 式を b と w でそれぞれ微分して 0 になる解を数値的に求めることで、順序ロジスティック回帰の係数が求まる。このようにして得られた係数と説明変数を (2.5.17) 式に代入することで、各階級になる確率の予測値を求めることができる。

### 2.5.5 多項ロジスティック回帰

ロジスティック回帰を多クラス分類に拡張したものを多項ロジスティック回帰<sup>3</sup>(Engel 1988) という。多項ロジスティック回帰は、例えば天気を晴れ、曇り、雨、雪に分類するような問題に用いることができる。

目的変数 y は M クラスのカテゴリーに対応した M 次元ベクトルで、実況は m 番目のクラスが起き  $(y_m=1)$ 、残りは起きない  $(y_m=0)$  ものとする。このような、ベクトルの 1 つの成分だけが 1 で残りは全て 0 であるベクトルを one-hot ベクトルという。 y は one-hot ベクトルであるから、 $\sum_{m=1}^M y_m=1$  となる。説明変数 x と係数 y が与えられたとき、y 番目のクラスの確率変数 y が y を取る確率を

$$p_m \equiv P_r \left( Y_m = 1 | \boldsymbol{x}, W \right) \tag{2.5.19}$$

と書くことにする。説明変数が K 個ある場合、x は切片も含めた K+1 次元ベクトルで、W は K+1 行 M 列の行列である。1 番目のクラスの実況が  $y_1$ 、 $y_2$  番目のクラスの実況が  $y_2$ 、 $\cdots$  である確率は、

$$P_r(Y_1 = y_1, \dots, Y_M = y_M | \boldsymbol{x}, W) = \prod_{m=1}^{M} p_m^{y_m}$$
(2.5.20)

となる。ここで  $\sum_{m=1}^M p_m = 1$ 、 $\sum_{m=1}^M y_m = 1$  である。2 クラス分類の場合、つまり M=2 の場合は、 $p_1+p_2=1$ 、 $y_1+y_2=1$  であり、 $p_1\equiv p$ 、 $y_1\equiv y$  とすると、 $p_2=1-p$ 、 $y_2=1-y$  であることから、

$$P_r (Y_1 = y_1, Y_2 = y_2 | \mathbf{x}, W)$$
  
=  $p_1^{y_1} p_2^{y_2} = p^y (1 - p)^{1 - y}$  (2.5.21)

となり、ベルヌーイ分布の確率関数 (2.5.3) 式と一致することが確かめられる。

N 組の学習データ $(y_1,x_1)$ ,  $\cdots$ ,  $(y_N,x_N)$  が与えられたとする。各観測が独立であると仮定すれば、尤度 L は係数行列 W の関数として、

$$L(W) = \prod_{n=1}^{N} \prod_{m=1}^{M} p_{nm}^{y_{nm}}$$
 (2.5.22)

<sup>&</sup>lt;sup>3</sup> ソフトマックス回帰、多クラス回帰、多分岐ロジスティック回帰などと呼ばれることもある。

と書けるので、対数尤度は

$$\ln L(W) = \sum_{n=1}^{N} \sum_{m=1}^{M} y_{nm} \ln p_{nm}$$
 (2.5.23)

となる。ここで、確率  $p_{nm}$  が次のソフトマックス関数 (第 2.6.5 項 (2) も参照 ) で表されるものと仮定する。

$$p_{nm} = \frac{e^{\phi_{nm}}}{\sum_{i=1}^{M} e^{\phi_{nj}}}$$
 (2.5.24)

$$\phi_{nm} = \sum_{k=0}^{K} x_{nk} w_{km} \tag{2.5.25}$$

これを用いると、対数尤度は

$$\ln L(W) = \sum_{n=1}^{N} \sum_{m=1}^{M} y_{nm} \left( \phi_{nm} - \ln \sum_{j=1}^{M} e^{\phi_{nj}} \right)$$
(2.5.26)

となる。この対数尤度を係数行列 W の各成分で微分すると、 $\sum_{m=1}^{M}y_{m}=1$  であることを用いて、

$$\frac{\partial \ln L}{\partial w_{km}} = \sum_{n=1}^{N} \sum_{l=1}^{M} y_{nl} \left[ x_{nk} \delta_{lm} - x_{nk} \frac{e^{\phi_{nm}}}{\sum_{j=1}^{M} e^{\phi_{nj}}} \right] 
= \sum_{n=1}^{N} x_{nk} \left( y_{nm} - p_{nm} \right)$$
(2.5.27)

となり、(2.5.9) 式と同様なシンプルな形で書ける。ここで  $\delta_{lm}$  はクロネッカーのデルタである。(2.5.27) 式が 0 になる W を数値的に求めることで、多項ロジスティック回帰における係数が得られる。求められた係数を (2.5.24) 式に適用することで、m 番目のクラスに分類される確率が得られる。複数のクラスの中から 1 つのクラスを予測する場合には、確率が最も高いクラスを選択すれば良い。

### 2.5.6 利用上の注意点

ロジスティック回帰では以下の仮定を用いて回帰係数を導出しているため、用いるデータはこれらの仮定と矛盾してはいけない。

- 観測値がベルヌーイ分布に従うこと
- 確率がロジスティック関数で表されること
- 観測が独立であること

以下ではそれぞれについて注意点を述べ、最後に上記 以外の一般的な注意点を述べる。

## (1) 観測値がベルヌーイ分布に従う

ロジスティック回帰に用いる観測値は現象があり、なしの2値データ(0か1)でなければならない。もし観測値として何らかの確率的なもの(割合等)が得られるとしても、その割合を用いて(2.5.9)式の解を求めてはならない。ベルヌーイ分布は確率変数Yが確率pで

現象あり (Y=1)、確率 1-p で現象なし (Y=0) を取る離散確率分布であるから、期待値と分散はそれぞれ次のように書ける。

$$E(Y) = 1 \times p + 0 \times (1 - p) = p \tag{2.5.28}$$

$$V(Y) = E(Y^2) - E(Y)^2 = p (1 - p)$$
 (2.5.29)

ここで p はロジスティック回帰により求められた確率である。横軸に p、縦軸に V(Y) を取ると図 2.5.3 のようになる。ベルヌーイ分布の分散は、確率が 0 または 1 付近で 0 に近く(ばらつきが小さく) 0.5 で最大値 0.25 をとる。このため、例えば線形重回帰で仮定したような、等分散性を持つデータはロジスティック回帰には適用することはできない。

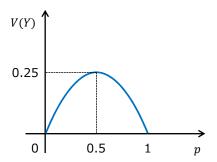


図 2.5.3 ベルヌーイ分布の分散

# (2) 確率がロジスティック関数で表される

ロジスティック回帰では、現象ありの確率が (2.5.7) 式のようなロジスティック関数で表されると仮定している。(2.5.7) 式を変形 (ロジット変換) すると、

$$logit(p) \equiv ln \frac{p}{1-p} = \sum_{k=0}^{K} x_k w_k$$
 (2.5.30)

となる。 $\ln(p/(1-p))$  を p のロジットという。今、ある説明変数  $x_k$  の値が  $x_k \to x_k + \Delta x_k$  と変化したとき、

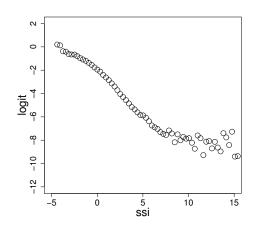


図 2.5.4 MSM の SSI と発雷の有無のロジットの関係。2012 年 3 月から 2013 年 2 月の 1 年間のデータを利用。

現象ありのロジットが  $\mathrm{logit}(p) \to \mathrm{logit}(p) + \Delta \mathrm{logit}(p)$  と変化したとすると、

$$\Delta \text{logit}(p) = \Delta x_k w_k \tag{2.5.31}$$

となる。これは、ある期間のデータに対して、横軸にある説明変数をとり、縦軸にその説明変数のある区間内に含まれる実況の有無(1 か 0 か)から算出したロジットをプロットした場合、説明変数がロジットと線形関係にあることを意味している。このことはまた、線形重回帰(第 2.4.9 項 (1))でも述べた通り、ある変数u の関数 f(u) (f(u)=u を含む)が既に説明変数に用いられている場合、別の関数 g(u) を説明変数に用いることはできないことも意味している。例として、MSM発雷確率ガイダンスの説明変数に用いられている SSIと発雷の有無のロジットの関係を図 2.5.4 に示す。これを見ると、多少の非線形性はあるものの、発雷の予測に重要な  $0\sim5$  付近では SSI とロジットは概ね線形関係にあることが分かる。

# (3) 観測の独立性

ロジスティック回帰では各観測データが独立であることを仮定している。すなわち観測データについて、空間方向や時間方向に相関がないことを仮定している。例えば図 2.5.5 のようなデータの場合、時間方向に相関があるため、時刻や季節で層別化したり、時間変化や季節変化を考慮した説明変数を加えたり、相関を考慮してデータを間引いたりする必要がある。また、同じ観測データを初期時刻の異なる説明変数に対応させた場合、同じような説明変数に対して同じ観測値を用いることになるため、相関の強い観測を用いたことにはるため、相関の強い観測を用いたことになるため、相関の強い観測を用いたこと同様の影響を与えてしまう。同じ観測データを複数回使用しないように、回帰式を初期時刻や予報時間で層別化する必要がある。

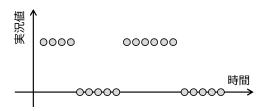


図 2.5.5 時間方向に相関のある観測データの例

# (4) その他の注意事項

ここまで述べた  $(1)\sim(3)$  の事項のほかに、線形重回帰と同様の注意点も挙げられる。説明変数間に多重共線性がある場合には、ロジスティック回帰の場合でも係数の推定精度が低下するため、説明変数を選択して使用する説明変数の数を減らすか、学習データの数を増やして係数推定の誤差を減らす必要がある。また、第2.4.9 項 (3) で述べたように、学習データやパラメータ

調整用のデータとは独立なデータを用いて性能を評価 する必要があることにも注意が必要である。

#### 2.5.7 まとめ

本節ではロジスティック回帰の基礎的な理論及び利用上の注意点を述べた。ロジスティック回帰は目的変数が2値データである場合に発生確率を予測することができ、シンプルな手法でありながら従来用いられていた手法と比べて高い予測精度を持つことから、気象庁のガイダンスにおいても確率予測に広く利用されている。開発においては、ロジスティック回帰を用いて予測した場合の検証結果(ブライアスコアや信頼度曲線など)だけに着目するのではなく、第2.5.6項で述べたように、実況データや説明変数をプロットしてみて、データの特性がロジスティック回帰に適しているか確認することも重要である。

# 参考文献

Cox, D. R., 1958: The regression analysis of binary sequences. *Journal of the Royal Statistical Society*. *Series B*, **20(2)**, 215–242.

Engel, J., 1988: Polytomous logistic regression. *Statistica neerlandica*, **42(4)**, 233–252.

McCullagh, P., 1980: Regression models for ordinal data. *Journal of the Royal Statistical Society. Series B*, **42(2)**, 109–142.

高田伸一, 2007: 航空気象予報ガイダンス. 平成 19 年度数値予報研修テキスト, 気象庁予報部, 87-93.

### 2.6 ニューラルネットワーク 1

### 2.6.1 はじめに

ニューラルネットワークは、神経細胞(ニューロン)の機能の一部をモデル化した機械学習アルゴリズムである。その研究は 1940 年代に始まり、Rosenblatt (1958, 1962) によるパーセプトロンと誤り訂正学習法を機に第 1 次ブームが、Rumelhart et al. (1986) による誤差逆伝播法の再発見を機に第 2 次ブームが起き、盛んに研究が行われた。その後 2012 年に ILSVRC <sup>2</sup> において Hinton らのグループ (Krizhevsky et al. 2012) がニューラルネットワークの階層を深くしたディープニューラルネットワーク (Hinton et al. 2006; LeCun et al. 2015)を用いて圧勝したことで第 3 次ブームが始まり、2018年現在も大きな注目を集めている。

ニューラルネットワークの特徴の一つは、第 2.6.3 項でも示すように、線形分離不可能な (任意の)関係も扱うことができることである。これに対して線形重回帰やロジスティック回帰では線形分離可能な関係しか扱うことができない。またニューラルネットワークでは、線形重回帰やロジスティック回帰で課されていた仮定がなく、目的変数や説明変数の関係を詳しく吟味する必要がないという特徴もある。このためネットワークの構築に関する自由度が非常に高い一方で、調整すべきパラメータが多くなり、計算を効率的に行うことが重要になっている。

気象庁のガイダンスにニューラルネットワークが利用されるようになったのは 1996 年から (柳野 1995) であり、基本的には 1980 年代から 1990 年代に開発された、いわゆる第 2 世代ニューラルネットワークの技術に基づいている。2018 年現在では最大降水量ガイダンス、降雪量地点ガイダンス、日照率ガイダンス、最小湿度ガイダンス、雲ガイダンスにニューラルネットワークが利用されている。

本節では、第2世代ニューラルネットワークの技術を中心に解説するが、近年開発され、ディープニューラルネットワークに用いられている技術についてもいくつか述べる。これらの新しい技術は、第2世代のニューラルネットワークにもそのまま適用可能であり、ガイダンスの予測精度の向上に資すると考えられる。より詳しい解説は、例えば Kermanshahi (1999)、岡谷 (2015)、斎藤 (2016) などを参照していただきたい。以下では、単にニューラルネットワークと書いた場合には、第2世代のニューラルネットワークまたは、ディープニューラルネットワークまで含めた一般的な意味でのニューラルネットワークを指すものとする。

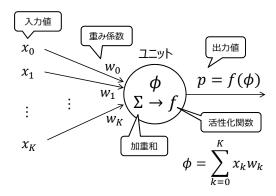


図 2.6.1 1 つのユニットへの入出力の模式図

本節では、初めに第2.6.2 項から第2.6.4 項でニュー ラルネットワークの構成と簡単な例を示し、第2.6.5項 と第2.6.6 項でニューラルネットワークによる予測値の 算出と係数の学習について述べる。続いて第2.6.7項 と第2.6.8 項でニューラルネットワークの学習に必要 な様々な技術を述べ、最後に利用上の注意点を述べる。 ディープニューラルネットワーク (またはディープラー ニング)の概要は第5.2節で述べる。ニューラルネット ワークでは、その性質上ほかの統計手法と比べて変数 や添字の数が多くなる。付録 2.3.A に、本章でも用い られる主な変数と添字をまとめてあるので、適宜参照 していただきたい。また、ネットワークの構成によっ て重み係数wなどの添字の付き方が変わるため、一般 的な議論をする場合にはいくつかの添字を省略する場 合がある。具体的な表現を知りたい場合には適宜添字 を補っていただきたい。

### 2.6.2 ニューラルネットワークを構成するパーツ

ニューラルネットワークは、図 2.6.1 で示したユニット  $^3$  を組み合わせることで構成される。各ユニットに対して K 個の入力データ x があり、それぞれに対応する K 個の重み係数 w を掛けて加重和  $\phi$  を求める。ここで 0 番目の係数  $w_0$  はバイアス項であり、対応する入力値  $x_0$  は常に 1 であるとする。その後  $\phi$  を活性化関数 f で変換した結果を出力値 p として出力する。活性化関数としては、下記の関数など  $^4$  が用いられている(図 2.6.2)。

ステップ関数: 
$$f(\phi) = \begin{cases} 1 & (\phi > 0) \\ 0 & (\phi \le 0) \end{cases}$$
 (2.6.1)

ロジスティック関数: 
$$f(\phi) = \frac{1}{1 + e^{-\phi}}$$
 (2.6.2)

双曲線正接関数: 
$$f(\phi) = \tanh \phi$$
 (2.6.3)

<sup>1</sup> 工藤 淳

<sup>&</sup>lt;sup>2</sup> ImageNet Large Scale Visual Recognition Challenge (http://www.image-net.org/challenges/LSVRC/)。米国の複数の大学が共同で開催している大規模画像識別コンペティション。1000 クラスの物体判別などを競っている。

<sup>&</sup>lt;sup>3</sup> ノードやニューロン、パーセプトロンと呼ばれることもあ る。

 $<sup>^4</sup>$  ほかにも、ソフトサイン関数:  $\phi/(1+|\phi|)$ 、ソフトプラス 関数:  $\ln(1+e^\phi)$ 、動径基底関数:  $e^{-\beta\phi}$ 、多項式:  $\phi^n$ 、絶対 値:  $|\phi|$  などもある。

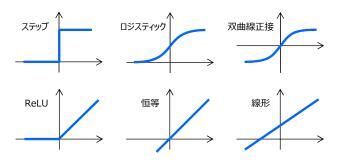


図 2.6.2 活性化関数の例

ReLU 関数: 
$$f(\phi) = \begin{cases} \phi & (\phi > 0) \\ 0 & (\phi \le 0) \end{cases}$$
 (2.6.4)

恒等関数: 
$$f(\phi) = \phi$$
 (2.6.5)

線形関数: 
$$f(\phi) = a\phi + b$$
 (2.6.6)

気象庁のガイダンスでは、ロジスティック関数、恒等 関数、線形関数が用いられている。双曲線正接関数は口 ジスティック関数と同様な形をしているが、原点で回転 対称性があるという特徴がある。ReLU 関数 (Rectified Linear Unit (正規化線形関数)またはランプ関数)は シンプルでありながら非線形な関数で、無限大でも勾 配が0にならないという特徴を持つ。従来用いられて いたロジスティック関数や双曲線正接関数と比べて精 度が向上する (Glorot et al. 2011; LeCun et al. 2015) ことから、ディープニューラルネットワークの多くで は ReLU 関数が用いられている。ここで示した例のう ち、ステップ関数、ロジスティック関数、双曲線正接関 数、ReLU 関数は非線形関数で、残りの2つは線形関 数である。次項で示すように、ニューラルネットワー クでは活性化関数が非線形関数であるか否かが重要な 意味を持つ。なお、ニューラルネットワークに関する 多くのテキストやこれまでのガイダンスの解説ではロ ジスティック関数のことをシグモイド関数と呼んでい るが、本稿ではロジスティック関数と呼ぶことにする。 シグモイド関数は狭義にはロジスティック関数を指す が、一般には双曲線正接関数や累積正規分布関数など のS字をした関数全般を指す。ここでは関数型を明確 にするためと、第2.5節のロジスティック回帰の解説と 用語を統一するために、ロジスティック関数と呼ぶこ とにする。

#### 2.6.3 簡単な例

ここでは簡単な例を用いて、ニューラルネットワークが線形分離不可能な関係も表現可能であることと、活性化関数にロジスティック関数や ReLU 関数などのステップ関数以外の非線形関数が用いられている理由を述べる。

図 2.6.3 のようにユニットに 2 つのデータ  $x_1$ と  $x_2$  が入力され、p が出力されるものとする。それぞれの値は

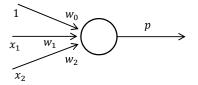


図 2.6.3 論理積を表現するユニット

表 2.6.1 論理積			
	$x_1$	$x_2$	出力
1	1	1	1
2	1	0	0
3	0	1	0
4	0	0	0

表 2.6.2		排他的論理和	
	$x_1$	$x_2$	出力
1	1	1	0
2	1	0	1
3	0	1	1
4	0	0	0

0か1である。このとき表 2.6.1 で示した論理積 (AND) をユニットを用いて表現することを考える。活性化関数にはステップ関数を用いることにする。すなわち、

$$\phi = w_0 + x_1 w_1 + x_2 w_2$$

$$p = \begin{cases} 1 & (\phi > 0) \\ 0 & (\phi \le 0) \end{cases}$$
(2.6.7)

とする。ここで例えば、 $(w_0,w_1,w_2)=(-5,3,3)$  とすると、 $x_1=1,x_2=1$  の場合には  $\phi=1>0$  で p=1 などとなることから、論理積が表現できていることを確かめられる。この w を用いて (2.6.7) 式を書くと、

$$p = \begin{cases} 1 & (3x_1 + 3x_2 - 5 > 0) \\ 0 & (3x_1 + 3x_2 - 5 \le 0) \end{cases}$$
 (2.6.8)

となる。これは、 $3x_1+3x_2-5=0$  を境に点  $(x_1,x_2)$  がこの線よりも上側にあれば 1 を下側にあれば 0 を返す、と解釈できる(図 2.6.4 左)。このように、一本の直線でデータを分類できることを線形分離可能という。逆の見方をすれば、w としてどのような値を用いたとしても、1 本の直線でしかクラスを分類できないとい

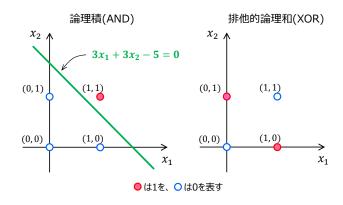


図 2.6.4 論理積(左)と排他的論理和(右)の図。赤丸は1 を、青丸は0を表す。

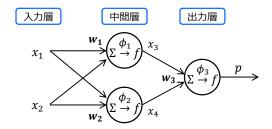


図 2.6.5 ユニットを 2 層に組み合わせたネットワーク。図を 簡略化するために、ここではそれぞれのユニットへのバイ アス項を省略している。

える。注意したいのは、活性化関数に用いたステップ 関数は非線形関数であるにも関わらず、線形分離可能 な問題しか解くことができないということである。これは活性化関数にロジスティック関数など他の非線形 関数を用いても同じである。ロジスティック関数でクラス分類をする場合、例えば p=0.5 を閾値として 1 か 0 かを分類することになる。(2.6.2) 式より、p=0.5 となるのは  $\phi=0$  のときであり、やはり  $\phi=0$  という直線を境に 2 つのクラスに分類されることになる。

線形分離不可能な問題の例として、排他的論理和 (XOR) がある (表 2.6.2)。排他的論理和の入出力の関係を図に示すと図 2.6.4 右のようになる。この図を見れば分かるように、排他的論理和の場合には 1 本の直線では 1 と 0 (赤丸と青丸)を分離することはできない。すなわち線形分離不可能である。そこで、ユニットを複数組み合わせて図 2.6.5 のようなネットワークを組むことにする。重み係数 w は、

$$\mathbf{w}_{1} = (w_{10}, w_{11}, w_{12}) = (5, -3, -3)$$

$$\mathbf{w}_{2} = (w_{20}, w_{21}, w_{22}) = (-3, 5, 5)$$

$$\mathbf{w}_{3} = (w_{30}, w_{31}, w_{32}) = (-5, 3, 3)$$
(2.6.9)

であるとする。ここで、 $w_{10}, w_{11}, w_{12}$  などはそれぞれ、1 番目のユニットでのバイアス項と  $x_1, x_2$  の重みを表す。活性化関数は全てステップ関数を用いる。この係数を用いて  $x_1, x_2$  に 0 または 1 を与えて具体的に計算すると表 2.6.3 のようになり、排他的論理和を表現できていることがわかる。図 2.6.5 のように、複数のユニットを層に並べたとき、入力データが与えられる層を入力層、計算結果が出力される層を出力層と呼び、その中間の層を中間層(または隠れ層)と呼ぶ。中間層は 1 層だけでも良いし 2 層以上あっても良い。

論理積と排他的論理和を解くにあたり、上記では重み係数はパラメータとして与えられたが、実際の分類問題に適用しようとした場合には何らかの方法で自動的に最適な係数を求めなくてはならない。論理積については次の誤り訂正学習法と呼ばれる手法で係数を自

表 2.6.3 図 2.6.5 のネットワークに (2.6.9) 式の係数を与え て計算した結果

	$x_1$	$x_2$	$x_3$	$x_4$	出力
1	1	1	0	1	0
2	1	0	1	1	1
3	0	1	1	1	1
4	0	0	1	0	0

動的に求めることができる。

$$w_k^{(s+1)} = w_k^{(s)} + \eta \sum_n (y_n - p_n^{(s)}) x_{nk}$$
 (2.6.10)

ここで s は学習のステップ、n は学習データの番号(表 2.6.1 の左列の番号  $1{\sim}4$  に相当)、 $y_n$  は n 番目の教師 データ(表 2.6.1 では右列の値に相当)、 $p_n^{(s)}$  は s ステップ目の係数を用いて計算した n 番目の出力値、 $x_{nk}$  は n 番目のデータにおける  $x_k$  の値、 $\eta$  は正の学習率である。  $\eta=0.1$  などとし、ある適当な初期値  $w_k^{(0)}$  から始めて上記ステップを複数回繰り返すことで正しい係数が得られる。

排他的論理和の場合は誤り訂正学習法では正しい係数を求めることはできないのだが、活性化関数に非線形関数を採用し、第 2.6.6 項で述べる誤差逆伝播法を用いることで最適な係数を学習することができる。ただし、活性化関数はステップ関数以外の非線形関数である必要がある。なぜならば、誤差逆伝播法では活性化関数の微分を用いて係数を更新するため、微分がデルタ関数になるステップ関数では係数を適切に更新できないからである。このため、第二世代以降のニューラルネットワークでは、活性化関数にロジスティック関数や ReLU 関数など、ステップ関数以外の非線形関数が用いられている。

1つのユニットでは表現できなかった排他的論理和を 表現できたのは、ユニットを2層に重ねたためである。 中間層の活性化関数に非線形関数を用いて、ユニット を2層以上重ねることで、理論上は任意の非線形クラ ス分類や回帰を近似できる (Cybenko 1989)。 では中間 層の活性化関数が線形関数(2.6.6)式だったらどうか。 これは実際に計算してみれば容易に確認できるが、中 間層の活性化関数を線形関数にすると、出力層の加重 和(図2.6.5の場合は $\phi_3$ )は入力値xの線形結合で表 される。これはすなわち中間層がないネットワークと 等価であることを意味しており、非線形分類問題を解 くことはできない。よって、ニューラルネットワーク の特性を活かすためには中間層に非線形関数を用いる 必要がある。中間層の活性化関数が非線形関数であれ ば、出力層の活性化関数は線形関数や恒等関数を用い ても良い(非線形関数でも良い)。

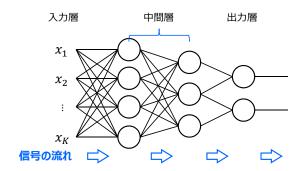


図 2.6.6 順伝播型ニューラルネットワークの構成

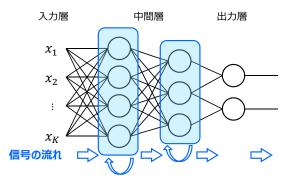


図 2.6.7 再帰型ニューラルネットワークの構成

### 2.6.4 ニューラルネットワークの構成

ニューラルネットワークは、ネットワークの構成に よって様々な種類がある。主なネットワークには以下 のようなものがある。

- 順伝播型ニューラルネットワーク
- 再帰型ニューラルネットワーク
- 相互結合型ニューラルネットワーク

順伝播型(フィードフォワード)ニューラルネットワークは、図 2.6.6 のようにユニットを複数の階層に並べたニューラルネットワークで、信号が入力層から出力層に向かって一方向のみに伝播していく。中間層は1層以上あり、それぞれの層のユニットは複数ある。画像認識などの分類・識別問題に用いられる最も基本的なニューラルネットワークである。

再帰型(リカレント)ニューラルネットワーク(Elman 1990)は、図 2.6.7 のように順伝播型ニューラルネットワークの中間層に有向閉路 5 を持たせたネットワークである。中間層では前回の出力値が入力値として扱われるため、短期の記憶として作用する。これによって前後のデータとの関連性を扱うことができ、音声認識や言語処理などの時系列データの処理に用いられる。短期の記憶だけでなく、長期の記憶も持たせた LSTM (Long Short-Term Memory, Hochreiter and Schmidhuber 1997) と呼ばれるネットワークもある。

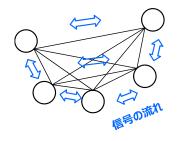


図 2.6.8 相互結合型ニューラルネットワークの構成

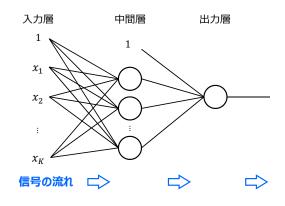


図 2.6.9 ガイダンスに用いられているニューラルネットワークの構成。入力層と中間層の 1 はバイアス項に対応。

相互結合型ニューラルネットワークは、図 2.6.8 のように自分自身を除く全てのユニットと結合したネットワークである。連想記憶や組み合わせ最適化問題に利用されるホップフィールドネットワーク (Hopfield 1982) や、ホップフィールドネットワークでのユニットの状態変化を確率的に行うボルツマンマシン (Ackley et al. 1985) などがある。

気象庁のガイダンスで用いられているニューラルネットワークは、図 2.6.9 のように、中間層が 1 層で出力層のユニット数が 1 の順伝播型ニューラルネットワークの構造を持っている。説明変数を入力層に与えることで、出力層からガイダンスの予測値が得られる。図 2.6.9 のように中間層が 1 層のニューラルネットワークを、本稿では(入力層 1 層と出力層 1 層を合わせて) 3 層ニューラルネットワーク 6 と呼ぶ。

ここまで見てきたように、ニューラルネットワークには主に3つの構成があり、中間層の層数やユニットの数、出力ユニットの数などに自由度があるが、本節では気象庁のガイダンスに用いられている、中間層が1層の順伝播型ニューラルネットワークを中心に解説する。

# 2.6.5 順方向の計算

前節で見たように、順伝播型ニューラルネットワークでは、入力層に与えられた信号は出力層に向かって

<sup>5</sup> 方向を持つ(双方向ではない)閉じた経路のこと。

 $<sup>^6</sup>$  ユニットとしての重なりは 2 層であるため、2 層ニューラルネットワークと呼ぶ文献もある。

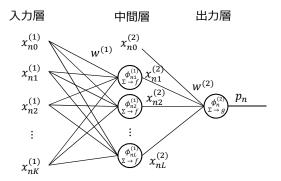


図 2.6.10 出力が 1 ユニットの 3 層順伝播型ニューラルネットワーク

一方向 (順方向)に伝播して何らかの値を出力する。以下では 3 層ニューラルネットワークについて、出力層のユニット数が 1 個の場合と M 個の場合の順方向の計算を行う。本節と次節では、3 層の順伝播型ニューラルネットワークについて述べるため、記述を簡略化するためにこれを単にニューラルネットワークと呼ぶ。

# (1) 出力が 1 ユニットの場合

図 2.6.10 に、気象庁のガイダンスで用いられているものと同じ、出力が 1 ユニットのニューラルネットワークの構成を示す。今、N 個の学習データが与えられており、その中の n 番目の入力データに対する出力値を求めることを考える。ここでは係数 w は何らかの方法で既に与えられているものとする。入力層から中間層への計算は以下の通りである。

$$\phi_{nl}^{(1)} = \sum_{k=0}^{K} x_{nk}^{(1)} w_{kl}^{(1)}$$
(2.6.11)

$$x_{nl}^{(2)} = f\left(\phi_{nl}^{(1)}\right) \tag{2.6.12}$$

ここで、 $x_{nk}^{(1)}$  は図に示したように 1 番目の層、すなわち入力層から中間層への入力値で、 $x_{n0}^{(1)}=1$  である。 $w_{kl}^{(1)}$  は入力層からの入力データに対応する中間層の l 番目のユニットの係数、 $x_{nl}^{(2)}$  は中間層の l 番目のユニットからの出力値で、f は中間層の活性化関数を表す。中間層の活性化関数は、気象庁のガイダンスではロジスティック関数が用いられているが、ディープニューラルネットワークでは  $\mathrm{ReLU}$  関数がよく用いられている。

中間層から出力層への計算は以下の通りである。

$$\phi_n^{(2)} = \sum_{l=0}^{L} x_{nl}^{(2)} w_l^{(2)}$$
 (2.6.13)

$$p_n = g\left(\phi_n^{(2)}\right) \tag{2.6.14}$$

出力層の活性化関数 g は、目的変数に応じて決めることになる。基本的には、目的変数が連続値(回帰)の場合には恒等関数を、2 値分類(確率予測)の場合にはロジスティック関数が用いられる。

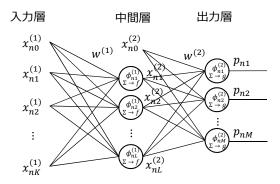


図 2.6.11 出力が M ユニットの 3 層順伝播型ニューラルネットワーク

## (2) 出力が複数ユニットの場合

図 2.6.11 に、出力層のユニットが M 個ある場合のニューラルネットワークの構成を示す。出力が複数ユニットのネットワークは 2018 年現在の気象庁のガイダンスでは用いられていないが、手書き文字判別などの多クラス分類問題に適用できるため、ニューラルネットワークでは頻繁に用いられている。多クラス分類は、天気カテゴリーの判別や降水量の階級判別などガイダンスにも応用できるため、ここで出力層のユニットが多数ある場合の順方向の計算についても書いておく。

入力層から中間層への計算は (2.6.11) 式および (2.6.12) 式と全く同じである。中間層から出力層への計算は以下の通りである。

$$\phi_{nm}^{(2)} = \sum_{l=0}^{L} x_{nl}^{(2)} w_{lm}^{(2)}$$
 (2.6.15)

$$p_{nm} = g\left(\phi_{nm}^{(2)}\right) \tag{2.6.16}$$

ここで *m* は出力層のユニット番号である。出力層の活性化関数には次のソフトマックス関数が用いられる。

$$g\left(\phi_{nm}^{(2)}\right) = \frac{e^{\phi_{nm}^{(2)}}}{\sum_{i=1}^{M} e^{\phi_{ni}^{(2)}}}$$
(2.6.17)

ソフトマックス関数は  $0\sim 1$  の値をとり、全てのクラスについて和をとると 1 になるため、それぞれのクラスに分類される確率を表すと解釈できる。複数のクラスから一つのクラスを選ぶ場合には最も確率の高いクラスを選択すれば良い。ただし単に確率の高いクラスを選ぶだけならば (2.6.17) 式を計算する必要はなく、最も大きい  $\phi_{nm}^{(2)}$  を選択するだけで良い。ソフトマックス関数によって求められた確率値は、次節の係数学習時に誤差関数を求める際に利用される。

実用上の問題として、 $\phi_{nm}^{(2)}$  が大きくなると  $e^{\phi_{nm}^{(2)}}$  の計算がオーバーフローする場合がある。このような場合の対策として、 $\phi_{ni}^{(2)}$  の中の最大値を  $\alpha_n$  としたとき、(2.6.17) 式の分母分子に予め  $e^{-\alpha_n}$  を掛けておけばよい。

$$g\left(\phi_{nm}^{(2)}\right) = \frac{e^{\phi_{nm}^{(2)} - \alpha_n}}{\sum_{i=1}^{M} e^{\phi_{ni}^{(2)} - \alpha_n}}$$
(2.6.18)

中間層の層数が2層以上ある場合も、層の数に応じてここまでに述べた計算を繰り返していくだけで順方向の計算を行うことができる。

### 2.6.6 係数の学習

第 2.6.3 項で述べた誤り訂正学習法では、線形分離可能な場合しか適切に学習することはできなかったが、第 2.3.10 項で述べた最急降下法や確率的勾配降下法を用いることで、線形分離不可能な場合でも学習によって適切な係数を求めることが可能になる。

ニューラルネットワークでは (2.3.53) 式  $\sim (2.3.55)$  式 を用いることで一括学習、ミニバッチ学習、逐次学習 のいずれも可能である。一括学習の場合は全ての学習 データを用いて学習し、ミニバッチ学習の場合はラン ダムに選んだ1つまたは複数のサンプルに対して学習 する。逐次学習では新しいデータを入手する度にその データを用いて学習する。いずれの学習方法でも、係 数の初期値にはランダムな値を与えることが多い(第 2.6.8 項(3) も参照)。一括学習とミニバッチ学習の場合 は、学習データに対して多数回の係数更新ステップ<sup>7</sup>を 繰り返して係数を決定する。ガイダンスに用いられて いるニューラルネットワークの逐次学習の場合は、運 用時には前回の係数を1ステップだけ更新するが、運 用開始前にミニバッチ学習などを用いて多数回の繰り 返し学習を行い、学習期間のデータに対して事前に係 数を最適化している。

一括学習を用いた場合、1 ステップの係数更新を行うためには全ての学習データで算出した誤差関数  $E_n$  に対して  $w_k$  での微分を求める必要があるため計算に時間が掛かるが、ミニバッチ学習では、ランダムに選択された一部の学習データに対して微分を計算すれば良いため、計算量が少なくなるというメリットがある。またミニバッチ学習を用いた場合には、選ばれたサンプルに応じて誤差関数の勾配の方向が変わるため、大域的な極小解が得られやすい (Keskar et al. 2017) という特徴がある。これにより一括学習を用いた場合と比べて過学習が抑制され予測精度(汎化能力)が高くなる傾向がある。このような理由により、近年のニューラルネットワークでは一括学習に代わってミニバッチ学習が用いられている。

(2.3.54) 式で示したように、ミニバッチ学習では係数を以下のようにして決定する。

$$w_k^{(s+1)} = w_k^{(s)} - \eta \sum_{n \in \mathcal{D}} \left. \frac{\partial E_n}{\partial w_k} \right|_{\boldsymbol{w} = \boldsymbol{w}^{(s)}}$$
(2.6.19)

ここで、 $\eta$  は学習率、 $E_n$  は n 番目の学習データに対する誤差関数、 $\mathcal{D}$  は N 個の学習データの中からランダムに選ばれたサンプルを表す。選ばれたサンプルの数をバッチサイズと呼ぶ。ミニバッチ学習ではサンプルの

選び方に自由度があるが、通常はバッチサイズを固定して、ステップごとに選択するサンプルを変える。適切なバッチサイズを事前に知ることはできないが、多すぎるとサンプルによる勾配のばらつきが小さくなってしまうため、ミニバッチ学習の良さが失われてしまう。バッチサイズとしては  $10\sim100$  程度にすることが多い (岡谷 2015)。降雪量地点ガイダンスの場合は数100 程度の学習データに対してバッチサイズを 10 としている。

ミニバッチ学習を用いる場合、バッチサイズはパラ メータであり、適切な値に設定するために、バッチサ イズを変えた実験を何度か繰り返すことになる。この とき、トータルの誤差関数  $E = \sum_{n \in \mathcal{D}} E_n$  はバッチサ イズに比例して大きくなる。なぜならば $E_n$ は常に正 なので、サンプルごとの誤差が同程度だとすれば、バッ チサイズに比例してトータルの誤差関数は大きくなる からである。このため $\eta$ を固定した場合、(2.6.19)式 での係数の1回の更新幅が設定したサンプル数に比例 して変化してしまう。そこで、各誤差関数をバッチサ イズ  $N_D$  で割った値として定義しておけば、バッチサ イズを変える度に $\eta$ を調整する必要がなくなる。これ に加えて、開発の過程の中で、一括学習、ミニバッチ 学習、逐次学習を使い分けながら最もよい手法を選択 する、ということも考慮すれば、学習方法によらず 1 回の学習に利用するデータ数(以下、本項では学習方 法に関わらずこれを N と表記する) で割った値を  $E_n$ と定義しておけば、その都度誤差関数の定義を変える 必要がないため便利である。

 $E_n$  の具体的な表現は出力層の活性化関数に応じて設定する。活性化関数が線形関数や恒等関数の場合(回帰の場合)は二乗誤差が用いられる。

$$E_n = \frac{1}{2N} (p_n - y_n)^2 \tag{2.6.20}$$

活性化関数がロジスティック関数の場合(確率予測の場合)は上記の二乗誤差または負の対数尤度が用いられる。

$$E_n = -\frac{1}{N} \left[ y_n \ln p_n + (1 - y_n) \ln(1 - p_n) \right] (2.6.21)$$

2018 年現在の気象庁のガイダンスでは二乗誤差を用いているが、近年のニューラルネットワークでは負の対数尤度がよく用いられている。活性化関数がソフトマックス関数の場合(多クラス分類の場合)は交差エントロピーが用いられる。

$$E_n = -\frac{1}{N} \sum_{m=1}^{M} y_{nm} \ln p_{nm}$$
 (2.6.22)

ここで  $y_{nm}$  の m 列成分は M 個の中の 1 つだけが 1 で残りは全て 0 のベクトル (one-hot ベクトル) で、  $\sum_{m=1}^{M} y_{nm} = 1$  が成り立つ。また、各クラスに分類される確率の和は 1 であることから、 $\sum_{m=1}^{M} p_{nm} = 1$  が

<sup>&</sup>lt;sup>7</sup> 繰り返し回数のことをエポック数とも呼ぶ

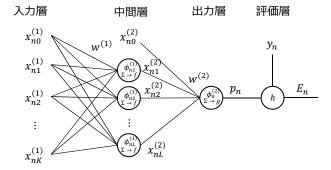


図 2.6.12 評価層を加えた出力が 1 ユニットの 3 層順伝播型 ニューラルネットワーク

成り立つ。M=2、すなわち 2 クラス分類の場合には、交差エントロピーは (2.6.21) 式と一致する。

一括学習、ミニバッチ学習、逐次学習のいずれを利用する場合でも、 $\partial E_n/\partial w_k$  を求める必要がある。以下では数値微分と誤差逆伝播法という  $\partial E_n/\partial w_k$  を求める 2 つの手法について述べる。

### (1) 数値微分による勾配計算

誤差関数の  $w_k$  方向の勾配を求める方法として、最も単純には、 $w_k$  を  $\Delta w_k$  だけ変化させて順方向の計算を行い、 $E_n$  の変化量  $\Delta E_n$  を求めれば良いだろう。ただしニューラルネットワークの出力は  $w_k$  に対して非線形なので中心差分を取ることにする。 $w_k \to w_k + \Delta w_k$ としたとき、 $E_n \to E_n + \Delta E_n^+$ 、 $w_k \to w_k - \Delta w_k$ としたとき、 $E_n \to E_n + \Delta E_n^-$  であるとすれば、

$$\frac{\partial E_n}{\partial w_k} \simeq \frac{\Delta E_n^+ - \Delta E_n^-}{2\Delta w_k} \tag{2.6.23}$$

となる。これを全ての $w_k$ に対して行うことで係数を 更新できる。この方法を数値微分と呼ぶ。

数値微分による学習は単純だが、1 つの係数を 1 ステップ更新する度に順方向の計算を 2 回行う必要があるため、全体の学習に掛かる時間は係数の数に比例して長くなってしまう。ニューラルネットワークは他の統計手法と比べて学習に掛かる時間が長い上に、第 2.6.1項でも述べたように調整すべきパラメータが多く、最適なパラメータを得るために何度も繰り返して学習する必要がある。このため 1 ステップの係数更新に掛かる時間を短くすることは開発を効率的に行う上で非常に重要である。勾配計算を高速に行う方法が次に述べる誤差逆伝播法 (Rumelhart et al. 1986)である。

# (2) 誤差逆伝播法による勾配計算

ここでは気象庁のガイダンスに用いられている、中間1層、出力1ユニットの場合の誤差逆伝播法について述べる。出力が多ユニットの場合や中間層が2層以上ある場合もここで述べる方法を拡張することで容易に計算できる。誤差逆伝播法の流れを把握するために、図2.6.12のように図2.6.10のネットワークに評価層を

表 2.6.4 出力層の活性化関数と用いられる誤差関数

 誤差関数	出力層の活性化関数
二乗誤差	線形関数(恒等関数)
二乗誤差	ロジスティック関数
負の対数尤度	ロジスティック関数

追加し、評価層ではニューラルネットワークの出力と 教師データから関数 h により誤差関数  $E_n$  を出力する と考える。

係数更新のために最終的に求めたい値は  $E_n$  の w 方向の勾配  $\partial E_n/\partial w_{kl}^{(1)}$  と  $\partial E_n/\partial w_l^{(2)}$  だが、まずは  $p_n$  方向の勾配  $\partial E_n/\partial p_n$  を計算する。出力層が 1 ユニットの場合、活性化関数は線形関数(恒等関数を含む)かロジスティック関数である。この場合の誤差関数は (2.6.20)式か (2.6.21)式である。よって  $\partial E_n/\partial p_n$  は、誤差関数が二乗誤差の場合は

$$\frac{\partial E_n}{\partial p_n} = \frac{1}{N} (p_n - y_n) \tag{2.6.24}$$

誤差関数が負の対数尤度の場合は

$$\frac{\partial E_n}{\partial p_n} = \frac{1}{N} \frac{p_n - y_n}{p_n (1 - p_n)} \tag{2.6.25}$$

となる。このことは、教師データ  $y_n$  が与えられれば、順方向の計算時に得られた出力値(すなわち既知の値)  $p_n$  を用いることで容易に  $\partial E_n/\partial p_n$  が算出できることを表している

続いて $w_l^{(2)}$ 方向の $E_n$ の勾配を計算する。 $p_n$ は $w_l^{(2)}$ の関数であるから、合成関数の微分の連鎖公式により、

$$\frac{\partial E_n}{\partial w_l^{(2)}} = \frac{\partial E_n}{\partial p_n} \frac{\partial p_n}{\partial \phi_n^{(2)}} \frac{\partial \phi_n^{(2)}}{\partial w_l^{(2)}}$$
(2.6.26)

と書ける。 $\partial E_n/\partial p_n$  は既に求めてあるので残りの部分を求める。 $\partial p_n/\partial \phi_n^{(2)}$  は、出力層の活性化関数が線形関数  $p_n=a\phi_n^{(2)}+b$  の場合は

$$\frac{\partial p_n}{\partial \phi_n^{(2)}} = a \tag{2.6.27}$$

ロジスティック関数の場合は

$$\frac{\partial p_n}{\partial \phi_n^{(2)}} = p_n (1 - p_n) \tag{2.6.28}$$

となる。また  $\partial\phi_n^{(2)}/\partial w_l^{(2)}$  は、 $\phi_n^{(2)}=\sum_{l=0}^L x_{nl}^{(2)}w_l^{(2)}$ より、

$$\frac{\partial \phi_n^{(2)}}{\partial w_l^{(2)}} = x_{nl}^{(2)} \tag{2.6.29}$$

となる。以上を組み合わせることで  $\partial E_n/\partial w_l^{(2)}$  が計算できる。以後の計算でも出力層の活性化関数と誤差関数の組み合わせにより計算式が異なるので、組み合わ

せを表 2.6.4 にまとめておく。それぞれの組み合わせに ついて誤差関数の勾配を計算すると、 の場合は、

$$\frac{\partial E_n}{\partial w_l^{(2)}} = \frac{a}{N} (p_n - y_n) x_{nl}^{(2)} \tag{2.6.30}$$

の場合は、

$$\frac{\partial E_n}{\partial w_l^{(2)}} = \frac{1}{N} (p_n - y_n) p_n (1 - p_n) x_{nl}^{(2)}$$
 (2.6.31)

の場合は、

$$\frac{\partial E_n}{\partial w_l^{(2)}} = \frac{1}{N} (p_n - y_n) x_{nl}^{(2)} \tag{2.6.32}$$

となる。出力層の活性化関数がロジスティック関数の場合、誤差関数に負の対数尤度を用いると、(2.6.32) 式のように (2.6.30) 式と同型になってシンプルに書けるため、近年のニューラルネットワークでは の組み合わせが用いられている。w 方向の誤差関数の微分を求める際に、 $\partial E_n/\partial p_n$  を求め、続いて  $\partial p_n/\partial \phi_n^{(2)}$  を求め、 ...、というように、図 2.6.12 で見た場合に順方向の計算方向とは逆向きに誤差の微分を計算することから、この方法を誤差逆伝播法という。

(2.6.30) 式  $\sim$  (2.6.32) 式の計算に必要な変数は  $p_n, y_n, x_{nl}^{(2)}$  であるが、 $y_n$  は教師データとして、 $x_{nl}^{(2)}$  は説明変数として与えられている値であり、 $p_n$  は順方向の計算を 1 回行えば求められる値である。数値微分で勾配を求める場合には係数の数に比例して順方向の計算を行う必要があったが、誤差逆伝播法の場合には順方向の計算を 1 回行うだけで全ての係数について勾配が求まる。さらにいえば、(2.6.26) 式のうち  $\partial E_n/\partial p_n \cdot \partial p_n/\partial \phi_n^{(2)}$  は全ての  $w_l^{(2)}$  について共通であるから、1 つの  $w_l^{(2)}$  について計算した結果を保持しておけばこの部分を毎回計算する必要もなくなる。このため誤差逆伝播法を用いれば勾配計算を高速に行うことができる。

次に  $\partial E_n/\partial w_{kl}^{(1)}$  を求める。合成関数の微分より、

$$\frac{\partial E_n}{\partial w_{kl}^{(1)}} = \frac{\partial E_n}{\partial p_n} \frac{\partial p_n}{\partial \phi_n^{(2)}} \frac{\partial \phi_n^{(2)}}{\partial x_{nl}^{(2)}} \frac{\partial x_{nl}^{(2)}}{\partial \phi_{nl}^{(1)}} \frac{\partial \phi_{nl}^{(1)}}{\partial w_{kl}^{(1)}} \qquad (2.6.33)$$

と書ける。ここでは出力1ユニットの場合のみ扱うが、 出力層のユニットが複数ある場合には、合成関数の微 分は

$$\frac{\partial E_n}{\partial w_{kl}^{(1)}} = \sum_{m=1}^{M} \frac{\partial E_n}{\partial p_{nm}} \frac{\partial p_{nm}}{\partial \phi_{nm}^{(2)}} \frac{\partial \phi_{nm}^{(2)}}{\partial x_{nl}^{(2)}} \frac{\partial x_{nl}^{(2)}}{\partial \phi_{nl}^{(1)}} \frac{\partial \phi_{nl}^{(1)}}{\partial w_{kl}^{(1)}}$$

$$(2.6.34)$$

のように和をとる必要があることに注意する。(2.6.33)式の右辺の初めの 2 つは (2.6.24) 式  $\sim (2.6.28)$  式で既に求めてあるので残りの 3 つを計算する。 $\phi_n^{(2)}=\sum_{l=0}^L x_{nl}^{(2)} w_l^{(2)}$  および  $\phi_{nl}^{(1)}=\sum_{k=0}^K x_{nk}^{(1)} w_{kl}^{(1)}$  である

ことと、中間層の活性化関数がロジスティック関数  $x_{nl}^{(2)}=(1+e^{-\phi_{nl}^{(1)}})^{-1}$  であるならば、

$$\frac{\partial \phi_n^{(2)}}{\partial x_{nl}^{(2)}} = w_l^{(2)} \tag{2.6.35}$$

$$\frac{\partial x_{nl}^{(2)}}{\partial \phi_{nl}^{(1)}} = x_{nl}^{(2)} \left( 1 - x_{nl}^{(2)} \right) \tag{2.6.36}$$

$$\frac{\partial \phi_{nl}^{(1)}}{\partial w_{bl}^{(1)}} = x_{nk}^{(1)} \tag{2.6.37}$$

となる。これらを合わせると、 と の場合( の場合は a=1 と見る)は、

$$\frac{\partial E_n}{\partial w_{kl}^{(1)}} = \frac{a}{N} (p_n - y_n) w_l^{(2)} x_{nl}^{(2)} \left( 1 - x_{nl}^{(2)} \right) x_{nk}^{(1)}$$
(2.6.38)

の場合は、

$$\frac{\partial E_n}{\partial w_{kl}^{(1)}} = \frac{1}{N} (p_n - y_n) p_n (1 - p_n) w_l^{(2)} x_{nl}^{(2)} \left( 1 - x_{nl}^{(2)} \right) x_{nk}^{(1)}$$
(2.6.39)

となる。以上で中間1層、出力1ユニットの場合の勾配を誤差逆伝播法で求めることができた。

誤差逆伝播法を利用することの最大の利点は、前述 のとおり勾配計算を高速で行うことができることであ る。一方で、誤差逆伝播法では各ユニットでの入出力 に対する局所的な勾配を掛け合わせることで誤差関数 の勾配を求めているため、数値微分が順方向の計算に 基づく非線形演算であるのに対して、誤差逆伝播法は 線形演算8となっている。このことは、誤差関数の情 報が逆向きに伝播していく過程で、どこかのユニット で勾配が 0 または ∞ に近づいたとすると、そのユニッ トにつながる全ての係数の勾配が0または $\infty$ になって しまうことを意味する。活性化関数にロジスティック 関数などのシグモイド型の関数を用いている場合、入 力値が0から遠ざかると勾配が急速に0に近くなるこ とから、このようなことは容易に起こりうる。勾配が 0に近づくと係数の更新速度が非常に遅くなってしま い計算効率が悪くなり、∞に近づくと係数の更新が非 常に不安定になってしまう。これを勾配消失問題とい う。ニューラルネットワークの階層を深くすると、そ の分だけ微分を掛け合わせる数が多くなるため、勾配 消失問題が起きやすくなる。このことが第2次ニュー ラルネットワークブームを終焉させた理由の一つ<sup>9</sup>で

 $<sup>\</sup>frac{8}{8}$  数値微分では w の変化量を 2 倍にしても誤差関数は 2 倍にはならないが、誤差逆伝播法では誤差関数が 2 倍になれば w 方向の勾配も 2 倍になる。

<sup>&</sup>lt;sup>9</sup> ニューラルネットワークでは係数以外のパラメータ (学習率やユニットの数、層の数など)が多く、これらを適切に決定する方法が見つかっていなかったことや、これらのパラメータを設定するコストの割に第 2.10 節で述べるサポートベクターマシンやランダムフォレストなどの手法と比べて精度が高くなかったことも理由の一つ。

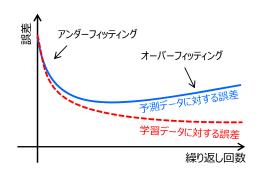


図 2.6.13 学習回数を変化させたときの学習データと予測 データそれぞれに対する誤差の変化

あり、勾配消失問題を回避する様々な手法が開発されてネットワークの階層を深くすることが可能になったことが第3次ニューラルネットワークブームを生み出した一つの要因となっている(岡谷(2015)など)。

### 2.6.7 過学習とその対策

ニューラルネットワークでの係数の学習は最急降下法や確率的勾配降下法を用いて行うため、誤差関数の最小値を探索するために係数更新ステップを多数回繰り返して行うことになる。このとき、繰り返し回数が少ないと学習が十分に進まず誤差が大きくなる(アンダーフィッティング)ためある程度以上の繰り返し計算が必要になるが、繰り返し回数が多すぎると学習データに対する誤差は小さくなる一方で、未知データに対する誤差は大きくなってしまう。これを過学習(オーバーフィッティング)という(図 2.6.13 )。

学習データの数に対してネットワークの自由度(ユニット数など)が多すぎる場合、ネットワークの表現能力が高過ぎるために繰り返し回数を多くすると学習データに適合しすぎて過学習が生じる。このため過学習対策としては、過学習の状態になる前に繰り返し学習を止める方法と、ネットワークの自由度を減らす方法の2つが考えられる。前者の方法として早期終了(Early Stopping)が、後者の方法として正則化とドロップアウトがある。これらは組み合わせて用いることもできる。

早期終了は、学習データを係数学習用と検証用に分け、検証用データに対して誤差が増加し始めたところで学習を終了する手法である。早期終了は簡単な方法で予測データに対する誤差を評価できる一方、学習データの中から検証用のデータを用意する必要があるため、学習用のサンプルが減るという問題がある。学習データが少ない場合には交差検証を用いる方法もある。

正則化では誤差関数に係数の大きさに応じたペナル ティー項を加えることで過学習を抑制する。すなわち、

$$E_n \to E_n + \frac{\lambda}{N\beta} \sum_{j,k,l} \left| w_{kl}^{(j)} \right|^{\beta} \tag{2.6.40}$$

とする。線形重回帰(第2.4.6項)などと同様に、 $\beta=1$ 

#### 確率pでユニットを残す

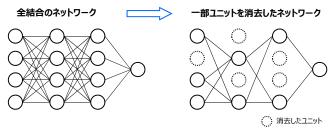


図 2.6.14 ドロップアウトの概念図

ならば L1 正則化、 $\beta=2$  ならば L2 正則化などとなる。ネットワークの自由度が高すぎて過学習を起こすというのであれば、自由度 (ユニット数) を減らせば良いだろうというのがドロップアウト (Hinton et al. 2012; Srivastava et al. 2014) の考え方である。ドロップアウトでは、学習時の係数更新ステップ毎にユニットの一部をランダムに消去する  $^{10}$  (図  $^{2.6.14}$ )。消去されたユニットからは、順方向・逆方向とも情報は伝達されない。ユニットを残す割合  $^{p}$  は入力層と中間層で異なる値を用いても良い。ドロップアウトを用いて学習した場合でも、予測時には全てのユニットを用いる。ただし、学習時に確率  $^{p}$  でユニットを残したことにより各ユニットの寄与量(または係数の大きさ)が平均的に  $^{1/p}$  倍になっていることから、これを戻すために各係数には  $^{p}$  を掛けた値を用いる。

ドロップアウトにおいて、ユニットをランダムに消去することは、ネットワークの自由度を強制的に減らしているという意味のほかに、様々なネットワークで学習した結果をアンサンブル平均しているという意味もある。複数の手法で予測した結果を平均すると予測精度が高くなる場合が多いが、複数の手法やネットワークで学習・予測することは計算コストや維持コストが掛かってしまう。ドロップアウトはこれと比べると十分に低コストで実装できる。

#### 2.6.8 学習のテクニック

線形重回帰やロジスティック回帰、カルマンフィルタなどの統計手法と比べ、ニューラルネットワークにはユニット数・学習率・正則化係数などの開発者が設定しなければならないパラメータが多く、適切な値を見つけるためには学習と検証を何度も繰り返す必要がある。また、係数の数が多く、係数空間上の誤差関数が複雑であるため、パラメータの設定によっては学習に時間が掛かったり最小値探索が失敗したりする場合もある。よってニューラルネットワークを用いた開発を効率的に進めるためには学習を効率的に行う必要があり、そのための様々なテクニックが開発されている。ここではその中からよく用いられている手法を紹介する。

<sup>10</sup> どのユニットを消去するかは確率的に決まるため、場合によってはある層の全てのユニットが消去されることもある。

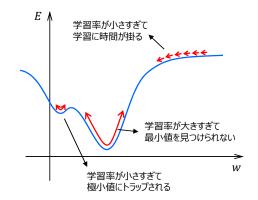


図 2.6.15 学習率の大きさと最小値探索の概念図

## (1) 学習率に関する手法

ニューラルネットワークの係数の学習に最急降下法 や確率的勾配降下法を用いる場合、学習率 $\eta$ が小さす ぎると学習に時間が掛かりすぎたり、極小値にトラッ プされたり、鞍部から抜け出すのに時間が掛かったり し、逆に大きすぎると最小値付近で振動して最小値を 見つけられなかったりする(図2.6.15)。このため学習 率を適切に設定することは重要であるが、ニューラル ネットワークの誤差関数は一般に複雑な形をしている ため、固定の学習率を用いることには限界がある。特 に図 2.6.16 のように、w のある成分に対して誤差関数 が緩やかに変化する場合、単純な最急降下法ではそれ に直行する方向に振動してしまい学習効率が悪くなる。 ただし最小値付近で振動しているのであれば、最小値 から多少離れた場所で計算を終了したとしても最小値 と比べて誤差は同程度であるだろうから、予測精度に は大きな違いはないかもしれない。しかし誤差関数が 緩やかな鞍部になっている場合には、誤差の変化が小 さいことから最小値に達したと判断し学習を止めてし まうと、本来の最小値から離れた係数を採用すること になり、予測精度を低下させる原因となる。ネットワー クのサイズが大きくなるほどこのような鞍部は多くな り、鞍部からいかにして抜け出すかということの方が 重要になってくる (Dauphin et al. 2014)。

固定値ではない学習率を用いる方法として、初めは 大きな学習率を使用し、学習が進むとともに小さくす るという方法がある。例えば次のようにする。

$$\eta = \eta_0 - \alpha s \tag{2.6.41}$$

ここで  $\eta_0$  は  $\eta$  の初期値、s は学習のステップで、 $\alpha$  は正のパラメータである。このほかにも、初めは  $\eta=\eta_0$  としておいて、ある程度学習が進んだと判断した時点で  $\eta=0.1\eta_0$  などとする方法もある。

より高度な方法として、モメンタム (Rumelhart et al. 1986) や AdaGrad (Duchi et al. 2011) などの方法もよく用いられる。モメンタムでは、前ステップとの係数の差を  $\Delta w^{(s)} \equiv w^{(s)} - w^{(s-1)}$  として、各ステップの

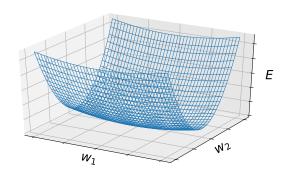


図 2.6.16 誤差関数が一つの方向に緩やかに変化する例

更新に慣性項 $\alpha \Delta w^{(s)}$ を加える。

$$w^{(s+1)} = w^{(s)} - \eta \left. \frac{\partial E}{\partial w} \right|_{w=w^{(s)}} + \alpha \Delta w^{(s)}$$
 (2.6.42)

ここで  $\alpha$  は正のパラメータで、 $0.5\sim0.9$  程度の値を設定する。慣性項は、w がある方向に振動している場合には振動を抑制し、一定方向に進んでいる場合には速度を増加させるように働くことで収束を早める。

AdaGrad では係数更新を次のように行う。

$$w_k^{(s+1)} = w_k^{(s)} - \frac{\eta}{\sqrt{h_k^{(s)}}} \left. \frac{\partial E}{\partial w_k} \right|_{w=w^{(s)}}$$
 (2.6.43)

$$h_k^{(s)} = \sum_{t=0}^{s} \left( \frac{\partial E}{\partial w_k} \right)^2 \bigg|_{w=w^{(t)}}$$
 (2.6.44)

ここで  $h_k^{(s)}$  はステップ 0 から s までの  $w_k$  方向の勾配 の二乗和であり、 $h_k^{(s)}$  が大きくなるほど  $w_k$  の更新速度 が遅くなっていく。例えば勾配が  $w_k$  方向には急であった場合、 $h_k^{(s)}$  は急速に大きくなるため、 $w_k$  の更新速度 はすぐに遅くなる。逆に勾配が  $w_k$  方向には緩やかで あったとすると、 $h_k^{(s)}$  は小さな値のままになり、更新速度が大きい状態が維持される。AdaGrad は w のそれぞれの方向について個別に更新速度が調整される。

ほかにも AdaGrad を改良した AdaDelta (Zeiler 2012)や、AdaGrad とモメンタムを組み合わせた Adam (Kingma and Ba 2015)など様々な手法が提案されている。AdaGrad、AdaDelta、Adam は最急降下法やモメンタムと比べて鞍部から効率的に抜け出すことができるため、特に近年のディープニューラルネットワークで良く利用されている。基本的には単純な最急降下法を用いるよりもモメンタムや AdaGrad などを用いた方が収束が早くなるが、常に最適な手法は存在せず、状況に応じて使い分けることになる。2018年現在の気象庁のガイダンスではモメンタムを用いているものが多い。

# (2) 説明変数の標準化

第 2.3.10 項でも述べたが、ニューラルネットワーク においても説明変数のオーダーに大きな差があると係

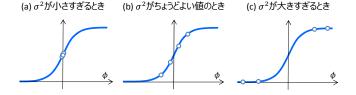


図 2.6.17 係数の初期値を用いて計算された中間層の各ユニットへの加重和(図中の丸)とそのユニットからの出力値の模式図(活性化関数がロジスティック関数の場合)。

数の学習効率が悪くなる。ニューラルネットワークでは係数の初期値は乱数で与えることが多いのだが、同じオーダーの乱数を初期値に与えた場合、説明変数のオーダーに大きな差があると加重和  $\phi = \sum_{k=0}^K x_k w_k$ を取ったときにオーダーの大きな説明変数だけが強調されてしまうため、大きなオーダーの説明変数を打ち消すような適切な係数に収束するまでに時間が掛かってしまう。このような理由から説明変数のオーダーに差がある場合には、学習期間の平均と標準偏差を用いて記明変数を標準化するか、最大・最小値を用いて $0\sim1$ の値に変換しておく。

### (3) 係数の初期値

ニューラルネットワークでは、係数の初期値は  $N(0,\sigma^2)$  に従う乱数などで与える。これは初期値に揺 らぎを与えずに完全に同じ値から学習を開始すると、 中間層の各ユニットから次の層への出力は常に全て同 じ値となり、中間層に複数のユニットを配置する意味 がなくなるからである。よって係数の初期値にはある 程度の揺らぎ (分散  $\sigma^2$ ) を与える必要があるのだが、  $\sigma^2$  の与え方によって学習効率が変わってしまう。 $\sigma^2$  が 小さすぎると全ての係数の初期値が  $w_0 \simeq 0$  となり、ど のような組み合わせで説明変数が与えられたとしても、 加重和はほとんど 0 になってしまう(図 2.6.17(a))。 これは係数の初期値に揺らぎを全く与えない場合に近 く、学習がほとんど進まなくなってしまう。一方で $\sigma^2$ が大きすぎると、中間層への入力の加重和が 0 から大 きく離れた値を取る可能性が高くなる(図 2.6.17(c))。 この時、図のように活性化関数にロジスティック関数 などのシグモイド型の関数を用いていると、中間層か らの出力が0または1に近い値に固定されてしまい、 活性化関数の傾きが小さくなるため、誤差逆伝播法で の学習が進まなくなってしまう(勾配消失問題が起き る)。以上のことから、図 2.6.17(b) のように、中間層 への入力の加重和がちょうど良い分散を持つようにwの初期値の分散  $\sigma^2$  を決めれば良いといえる。

Glorot and Bengio (2010) は、活性化関数 f が対称でかつ 0 での微分係数が 1、すなわち f'(0)=1 の場合に、係数の初期値の分散を  $\sigma^2=1/M$  にする手法を示した。ここで M は前の層のユニット数である。この初期化方法は Xavier の初期値と呼ばれ、ディープニュー

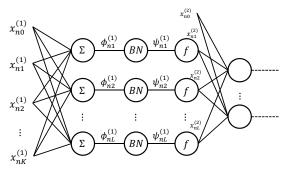


図 2.6.18 バッチ標準化を活性化関数の前に行う場合の模式 図。バッチ標準化の演算を BN で表す。

ラルネットワークで広く利用されている。

He et al. (2015) は、活性化関数が非対称な ReLU 関数の場合には係数の初期値の分散を  $\sigma^2=2/M$  とすることを提案し、ReLU 関数の場合に Xavier の初期値を用いた場合と比べて学習が速く進むことと、ネットワークを深くした場合に Xavier の初期値と比べて係数の収束性が良いことを示した  $^{11}$ 。He et al. (2015) による初期化の手法は He の初期値と呼ばれている。

Kumar (2017) は Xavier の初期値を 0 で微分可能な活性化関数に拡張し、中間層のユニット数が多い場合、活性化関数を 1 次の微分までで近似することで、

$$\sigma^2 = \frac{1}{Mf'(0)^2 (1 + f(0)^2)}$$
 (2.6.45)

とする手法を示した(この式の導出は付録 2.6.A に示す)。これを用いると、例えば 0 での微分係数が 1 になる双曲線正接関数では  $\sigma^2=1/M$  となり Xavier の初期値と一致することが確かめられる。また、ロジスティック関数の場合には f(0)=0.5, f'(0)=0.25 より、 $\sigma^2=64/(5M)$  となる。Kumar (2017) は活性化関数にロジスティック関数を用いた 10 層のニューラルネットワークで画像識別を行い、(2.6.45) 式を用いることでXavier の初期値を利用した場合と比べて非常に速く学習が進むことを示している。

# (4) バッチ標準化

中間層からの出力値の分散が大きすぎたり小さすぎたりすることが問題、というのであれば、いつでも適切な分散になるように強制的に変換してやればよいだろう、というアイディアがバッチ標準化(Batch Normalization, Ioffe and Szegedy 2015)である。バッチ標準化の演算は中間層の活性化関数の前か後に行う。中間層のユニットを演算の役割ごとに分割すると図 2.6.18 のようになる。図で BN はバッチ標準化の演算を表す。

バッチ標準化ではミニバッチ学習を行うことを前提 としている。ミニバッチ学習による1ステップの係数

<sup>1 22</sup> 層と 30 層のネットワークで比較し、Xavier の初期値では 22 層では収束するが、30 層では収束しない一方、彼らの初期化を用いた場合には 30 層のネットワークでも収束することを示した。

更新を考えた場合、ミニバッチの中には  $N_{\mathcal{D}}$  個のサンプルが存在する。それぞれのサンプルに対して順方向の計算を行うことで、 $N_{\mathcal{D}}$  個の中間層への入力の加重和  $\phi^{(1)}$  が得られることになる。これらの加重和の標本平均  $\bar{\phi}_i^{(1)}$  と標本分散  $s_i^{(1)}$  はそれぞれ

$$\bar{\phi}_l^{(1)} = \frac{1}{N_D} \sum_{n=1}^{N_D} \phi_{nl}^{(1)}$$
 (2.6.46)

$$s_l^{(1)^2} = \frac{1}{N_D} \sum_{n=1}^{N_D} \left( \phi_{nl}^{(1)} - \bar{\phi}_l^{(1)} \right)^2$$
 (2.6.47)

である。バッチ標準化では各加重和を

$$\hat{\phi}_{nl}^{(1)} = \frac{\phi_{nl}^{(1)} - \bar{\phi}_l^{(1)}}{\sqrt{s_l^{(1)^2} + \epsilon}}$$
(2.6.48)

として標準化した後に、

$$\psi_{nl}^{(1)} = \gamma^{(1)}\hat{\phi}_{nl}^{(1)} + \beta^{(1)} \tag{2.6.49}$$

と変換する。ここで  $\epsilon$  はゼロ割を防ぐための小さな数で、 $\beta^{(1)}$  と $\gamma^{(1)}$  は  $\beta^{(1)}=0$ ,  $\gamma^{(1)}=1$  を初期値としておいて、誤差逆伝播法で学習させる係数である。同様の演算を全ての中間層に対して行う。バッチ標準化を用いることで、学習率を大きくできて学習に掛かる時間を短縮できる、係数の初期値のことを細かく考えなくてもよい、過学習が抑制できる、という優れた効果があり、精度向上に大きく寄与することから 2018 年現在では広く利用されている。

## 2.6.9 まとめと利用上の注意点

本節では気象庁のガイダンスに利用されている出力 1ユニットの3層順伝播型ニューラルネットワークを中心に、ニューラルネットワークでの予測値の計算、係数の学習方法、過学習への対策や学習のテクニックについて述べた。

ニューラルネットワークと線形重回帰やロジスティック回帰の大きな違いは、ニューラルネットワークは目的変数と説明変数が非線形関係を持つ場合にも適用できることと、係数を決定する上での仮定(制約)がないことが挙げられる  $^{12}$ 。ニューラルネットワークでは、線形重回帰やロジスティック回帰では注意しなければならなかった目的変数や説明変数の多重共線性や線形性、等分散性などを考慮する必要はない。ただし学習効率の観点から、第  $^{2.6.8}$  項の  $^{(2)}$  で述べたように説明変数を標準化する必要はある。また、予測対象が連続値(回帰)か  $^{2}$  クラス分類(確率予測)か多クラス分類かなど、問題に応じてネットワークや活性化関数を設定する必要がある。

ニューラルネットワークは適用範囲が広い反面、他の統計手法と比べて開発者が設定しなければならないパラメータが多く、かつ適切に設定しないと学習が多まく進まないことや、学習に時間が掛かるという問題がある。このため、学習を効率的に行うことがガイダンスの精度を向上させる上で重要なポイントとなう。また、線形重回帰やロジスティック回帰、および次(あまた、線形重回帰やロジスティック回帰、および次(がの説明変数がどれくらい寄与しているか、異常な値が予測された場合に何が原因だったのかなど)を理解することが困難であるという問題がある。このため、ガイダンスを開発・運用するに当たっては、中間層からの出力値およびその分布、各ユニットの重み係数、説明変数を変化させた場合の予測値などをモニターすることも重要である。

中間層のユニット数や学習率、正則化の係数など、パラメータを調整する際には、学習データを係数学習用のデータとパラメータ調整の評価用データに分けて、評価用データに対して精度が高くなるようにパラメータを調整する(パラメータ調整の評価に検証用データを用いてはならない)。よって、パラメータの調整も含めてニューラルネットワークを用いたガイダンスを開発する場合には、用意したデータを係数学習用データ、パラメータ調整の評価用データ、予測精度検証用のデータに分割して使用することになる。分割することでサンプル数が少なくなる場合には交差検証の利用を検討する。

近年のニューラルネットワークブームにより、有効な 手法が次々と生み出されている。これらの手法はディー プニューラルネットワークでの利用を目的としたもの ではあるが、気象庁のガイダンスに利用されている 3 層のニューラルネットワークにも適用可能であるため、 最新の研究の動向を踏まえつつ、有効な手法を適宜導 入するよう検討していきたい。

# 付録 2.6.A Kumar による初期値

ここでは Kumar (2017) に基づいて (2.6.45) 式を導出する。全部で L 層(入力層を含めると L+1 層)の深いネットワークを考える。第 l 層のユニット数を M、第 l+1 層のユニット数を K とすると、第 l+1 層からの出力値  $x_k^{(l+1)}$  は、

$$x_k^{(l+1)} = f\left(\phi_k^{(l)}\right)$$
 (2.6.50)

$$\phi_k^{(l)} = \sum_{m=1}^M \psi_{mk}^{(l)} \tag{2.6.51}$$

$$\psi_{mk}^{(l)} = x_m^{(l)} w_{mk}^{(l)} \tag{2.6.52}$$

と書ける。ここで、係数 w の初期値の分布は  $N(0,v^2)$  であるとする。また、入力データは N(0,1) に標準化されているものとする。記述の定義として、 $\phi_m^{(l)}$  の標本平均を  $\bar{\phi}^{(l)}$ 、標本分散を  $s^{(l)^2}$  と書く。

<sup>12</sup> ディープニューラルネットワークでは特徴量 ( 顔の輪郭や目の形など ) を自動的に学習することも特徴の一つに挙げられる。

1回目の更新ステップを考えた場合、 $w_{mk}^{(l)}$ は  $x_m^{(l)}$ と独立である。また、1 ステップ目の順方向の計算では、第 l 層への入力データ  $x_m^{(l)}$  の期待値と分散はユニットによらず等しいとする。すなわち、

$$E\left(x_m^{(l)}\right) = \mu^{(l)} \tag{2.6.53}$$

$$V\left(x_m^{(l)}\right) = \sigma^{(l)^2} \tag{2.6.54}$$

と書けるものとする。ここで知りたいのは1回目の更新ステップで各層への入力データの分散が全て等しい値1、つまり、

$$\sigma^{(1)^2} \simeq \sigma^{(2)^2} \simeq \cdots \sigma^{(L)^2} = 1$$
 (2.6.55)

を満たすようなwの初期値の分散 $v^2$ である。 $w_{mk}^{(l)}$ の期待値が0であることから、 $\bar{\phi}^{(l)}$ の期待値は、

$$E\left(\bar{\phi}^{(l)}\right) = E\left(\phi_k^{(l)}\right) = E\left(\sum_{m=1}^M x_m^{(l)} w_{mk}^{(l)}\right)$$
$$= \sum_{m=1}^M E\left(x_m^{(l)}\right) E\left(w_{mk}^{(l)}\right) = 0 \qquad (2.6.56)$$

となる。また $\phi_k^{(l)}$ の分散は、

$$V\left(\phi_{k}^{(l)}\right) = E\left(\phi_{k}^{(l)^{2}}\right) - E\left(\phi_{k}^{(l)}\right)^{2}$$

$$= \sum_{m,j} E\left(x_{m}^{(l)} w_{mk}^{(l)} x_{j}^{(l)} w_{jk}^{(l)}\right)$$

$$= \sum_{m=j} E\left(x_{m}^{(l)} w_{mk}^{(l)} x_{j}^{(l)} w_{jk}^{(l)}\right)$$

$$= \sum_{m=1}^{M} E\left(x_{m}^{(l)^{2}} w_{mk}^{(l)^{2}}\right)$$

$$= \sum_{m=1}^{M} E\left(x_{m}^{(l)^{2}}\right) E\left(w_{mk}^{(l)^{2}}\right) \qquad (2.6.57)$$

である。ここで、(2.6.53) 式と(2.6.54) 式より、

$$\sigma^{(l)^{2}} = V\left(x_{m}^{(l)}\right) = E\left(x_{m}^{(l)^{2}}\right) - E\left(x_{m}^{(l)}\right)^{2}$$
$$= E\left(x_{m}^{(l)^{2}}\right) - \mu^{(l)^{2}}$$
(2.6.58)

である。これと、w の初期値の分散が  $v^2$ 、すなわち  $V\left(w_{mk}^{(l)}\right)=E\left({w_{mk}^{(l)}}^2\right)=v^2$  を (2.6.57) 式に代入すると、

$$V\left(\phi_k^{(l)}\right) = Mv^2 \left(\sigma^{(l)^2} + \mu^{(l)^2}\right)$$
 (2.6.59)

と書ける。ここで、1 ステップ目では、 $w_{mk}^{(l)}$ は  $x_m^{(l)}$  と独立であることから、

$$E\left(\psi_{mk}^{(l)}\right) = E\left(x_m^{(l)}\right)E\left(w_{mk}^{(l)}\right) = 0 \tag{2.6.60}$$

かつ、 $i \neq j$  について、

$$Cov\left(\psi_{mi}^{(l)},\psi_{mj}^{(l)}\right)$$

$$= E\left(\psi_{mi}^{(l)}\psi_{mj}^{(l)}\right) - E\left(\psi_{mi}^{(l)}\right) E\left(\psi_{mj}^{(l)}\right)$$

$$= E\left(x_{m}^{(l)}w_{mi}^{(l)}x_{m}^{(l)}w_{mj}^{(l)}\right)$$

$$= E\left(x_{m}^{(l)}\right) E\left(w_{mi}^{(l)}\right) E\left(x_{m}^{(l)}\right) E\left(w_{mj}^{(l)}\right)$$

$$= 0 \tag{2.6.61}$$

であることから、 $\psi_{mi}^{(l)}$ と $\psi_{mj}^{(l)}$   $(i \neq j)$  は互いに独立かつ同じ確率分布を持つ。よって M が大きい場合には中心極限定理より、 $\psi_{mk}^{(l)}$  の和である  $\phi_k^{(l)} = \sum_{m=1}^M \psi_{mk}^{(l)}$  は正規分布で近似できる。このことは 1 ステップ目以外でも概ね成り立つ。

ここで、活性化関数  $f(\phi)$  が  $\phi=0$  で微分可能であると仮定する。(2.6.50) 式を 0 のまわりでテーラー展開して  $\phi$  の 1 次までで近似すると、

$$x_k^{(l+1)} \simeq f(0) + f'(0) \,\phi_k^{(l)}$$
 (2.6.62)

であるから、(2.6.56) 式より、

$$E\left(x_k^{(l+1)}\right) \simeq f(0) + f'(0)E\left(\phi_k^{(l)}\right) = f(0)$$
 (2.6.63)

となる。よって、 $\mu^{(l+1)} \simeq f(0)$  であり、 $\mu^{(l+1)}$  は l に依らないことが分かる。すなわち、 $l \geq 1$  について、

$$\mu^{(l)} \simeq f(0) \tag{2.6.64}$$

といえる。これと (2.6.54) 式、(2.6.59) 式および (2.6.62) 式より、

$$\sigma^{(l+1)^2} = V\left(x_k^{(l+1)}\right) \simeq f'(0)^2 V\left(\phi_k^{(l)}\right)$$
$$= Mv^2 f'(0)^2 \left(\sigma^{(l)^2} + f(0)^2\right) \tag{2.6.65}$$

となる。条件 (2.6.55) 式より、 $\sigma^{(l)^2}$  が l に依らず 1 であるから、

$$v^{2} = \frac{1}{Mf'(0)^{2}(1+f(0)^{2})}$$
 (2.6.66)

が成り立つ。

# 参考文献

Ackley, D. H., G. E. Hinton, and T. J. Sejnowski, 1985: A learning algorithm for boltzmann machines. *Cognitive Science*, **9(1)**, 147–169.

Cybenko, G., 1989: Approximation by superpositions of a sigmoidal function. *Mathmatics of Control, Signals and Systems*, **2(4)**, 303–314.

Dauphin, Y. N., R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, 2014: Identifying and attacking the saddle point problem in highdimensional non-convex optimization. Proceedings of the 27th International Conference on Neural Information Processing Systems, 2933–2941.

- Duchi, J., E. Hazan, and Y. Singer, 2011: Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12, 2121–2159.
- Elman, J. L., 1990: Finding structure in time. *Cognitive Science*, **14**, 179–211.
- Glorot, X. and Y. Bengio, 2010: Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, 249–256.
- Glorot, X., A. Bordes, and Y. Bengio, 2011: Deep Sparse Rectifier Neural Networks. Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, 15, 315–323.
- He, K., X. Zhang, S. Ren, and J. Sun, 2015: Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. Proceedings of the 2015 IEEE International Conference on Computer Vision, 1026–1034.
- Hinton, G. E., S. Osindero, and Y. Teh, 2006: A fast learning algorithm for deep belief nets. *Neural Computation*, 18, 1527–1554.
- Hinton, G. E., N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, 2012: Improving neural networks by preventing coadaptation of feature detectors. arXiv preprint, arXiv:1207.0580.
- Hochreiter, S. and J. Schmidhuber, 1997: Long shortterm memory. *Neural Computation*, **9(8)**, 1735–1780.
- Hopfield, J. J., 1982: Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, **79(8)**, 2554–2558.
- Ioffe, S. and C. Szegedy, 2015: Batch Normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning, 448– 456.
- Kermanshahi, B., 1999: ニューラルネットワークの設計と応用. 昭晃堂, 146 pp.
- Keskar, N. S., D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, 2017: On large-batch training for deep learning: Generalization gap and sharp minima. Conference paper at ICLR 2017.
- Kingma, D. P. and J. L. Ba, 2015: Adam: A method for stochastic optimization. Conference paper at ICLR 2015.

- Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012: ImageNet Classification with Deep Convolutional Neural Networks. Proceedings of the 25th International Conference on Neural Information Processing Systems, 1097–1105.
- Kumar, S. K., 2017: On weight initialization in deep neural networks. arXiv preprint, arXiv:1704.08863.
- LeCun, Y., Y. Bengio, and G. E. Hinton, 2015: Deep learning. *Nature*, **521**, 436–444.
- 岡谷貴之, 2015: 機械学習プロフェッショナルシリーズ 深層学習. 講談社, 165 pp.
- Rosenblatt, F., 1958: The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, **65(6)**, 386–408.
- Rosenblatt, F., 1962: Principles of neurodynamics; perceptrons and the theory of brain mechanisms. Spartan Books, 616 pp.
- Rumelhart, D. E., G. E. Hinton, and R. J. Williams, 1986: Learning representations by backpropagating errors. *Nature*, **323**, 533–536.
- 斎藤康毅, 2016: ゼロから作る Deep Learning. オライリー・ジャパン, 298 pp.
- Srivastava, N., G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, 2014: Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15, 1929–1958.
- 柳野健, 1995: ニューラルネットによるガイダンス. 平成7年度量的予報研修テキスト, 気象庁予報部, 54-69.
- Zeiler, M. D., 2012: ADADELTA: An adaptive learning rate method. arXiv preprint, arXiv:1212.5701.

## 2.7.1 はじめに

カルマンフィルタは、ノイズを持つ観測の時系列デー タを元に、時々刻々と変化するシステムの現在の状態 を推定する時系列解析の手法の一つである。カルマン フィルタの応用範囲は広く、人工衛星やロボットの制 御、カーナビ、物体追跡、経済学、統計学、気象予測 など様々な分野で利用されている。気象予測において は、ガイダンスのほかにもデータ同化にカルマンフィ ルタが利用されている。データ同化で求めたいものは 現在の大気の状態であり、これを各種観測データと第 一推定値(数値予報モデルの予測値)を用いて推定す る。ガイダンスで求めたいものは目的変数と説明変数 を結びつける係数であり、これを観測値と第一推定値 (前回の係数)を用いて推定することになる。データ同 化の場合には大気の状態を表す変数の次元が多すぎる ため、カルマンフィルタではなく、その近似である4 次元変分法やアンサンブルカルマンフィルタなどが用 いられるが(堀田・太田 2011)、ガイダンスの場合には 変数の次元 (係数の数)が少ないため、カルマンフィ ルタを直接用いて係数を求めることが可能である。

ここまで見てきた線形重回帰やロジスティック回帰、ニューラルネットワークは、一定期間のデータに対して最適な係数を決定する、という統計手法であった。特に一括学習を用いた場合、これらの手法ではデータの並び順に意味はなく、どの順番でデータが並んでいたとしても同じ結果が得られる。これに対してカルマンフィルタは時系列データを扱う手法であるという点で上記の手法と大きく異なる。カルマンフィルタではデータの並び順に意味があり、順番を替えると異なる係数が得られる。

ガイダンスでカルマンフィルタを利用することの最 大の利点は、係数が最適な値になるように逐次更新さ れることにある。確率的勾配降下法を用いることでも 係数を逐次更新することは可能だが、ノイズを考慮し た時系列データを扱うという意味で、カルマンフィル タの方がより洗練された手法である。ガイダンスへの カルマンフィルタの利用に関する研究は1980年代に始 まっている (Persson 1989, 1991; Simonsen 1991)。 当 時のガイダンスは主に線形重回帰による一括学習が用 いられており、数値予報モデルの更新時に係数を再学 習しなければガイダンスの予測精度が低下することが 問題となっていた (Simonsen 1991)。 気象庁のガイダ ンスにも同様の問題があり(第1.3節を参照) 1996年 からカルマンフィルタが利用されるようになった (瀬上 ほか 1995)。2018 年現在では気温、風、平均降水量、 降水確率、時系列湿度、視程の各ガイダンスにカルマ ンフィルタが利用されている。

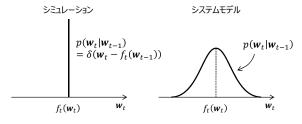


図 2.7.1 シミュレーションとシステムモデルの確率分布

本節では、初めにカルマンフィルタによる係数更新の関係式を導出し、続いてカルマンフィルタでのパラメータ設定や整合性の確認、利用上の注意点を述べる。カルマンフィルタの関係式の導出にはいくつかの方法が考案されており、最小分散推定値や線形最小分散推定値に基づいて解説される場合も多いが、本節では北川(1993)、樋口(2011)に基づき、係数の条件付き期待値から関係式を導出する。条件付き期待値は最小分散推定値と一致することが示されている(片山 2000)。

### 2.7.2 シミュレーションとシステムモデル

時刻 t の状態を表す変数 w (これを状態変数と呼ぶ)の時間発展を表す方程式を  $w_t = f_t(w_{t-1})$  と書くと、状態変数の初期値  $w_0$  が与えられれば、任意の時刻の状態が厳密に求まる。これをシミュレーションという。シミュレーションでは結果は一意に決まるが、あえて確率密度関数で表すと、デルタ関数を用いて以下のように書ける。

$$p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1}) = \delta(\boldsymbol{w}_t - f_t(\boldsymbol{w}_{t-1}))$$
 (2.7.1)

ここでシミュレーションに"遊び"を許し、観測データを取り込める自由度を持った式  $w_t \simeq f_t(w_{t-1})$  へと拡張する。近似記号を用いる代わりに新たな項  $u_t$  を導入し、以下のように書く。

$$\boldsymbol{w}_t = f_t(\boldsymbol{w}_{t-1}) + \boldsymbol{u}_t \tag{2.7.2}$$

これをシステムモデルという。 $u_t$  は  $w_t$  と独立なノイズで、システムノイズと呼ぶ。ガイダンスの場合には  $w_t$  は係数で、上式は時刻 t-1 の係数を用いて時刻 t の係数を求める式となる。シミュレーションでは結果 は一つであるのに対し、システムモデルでは、 $w_{t-1}$  が与えられたときの  $w_t$  の確率密度関数は何らかの確率分布になる(図 2.7.1)。今、 $u_t$  が与えられたとして、(2.7.1) 式と同様に確率密度関数をデルタ関数を用いて書くと、

$$p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1}) = \delta(\boldsymbol{w}_t - f_t(\boldsymbol{w}_{t-1}) - \boldsymbol{u}_t)$$
 (2.7.3)

となる。

# 2.7.3 観測モデル

例えば日々の最高気温の時系列データなど、観測の時系列データを $y_t$ とする。ガイダンスの場合、 $y_t$ はス

<sup>1</sup> 工藤 淳

カラー量であるが、カルマンフィルタではベクトルとして扱われる。本節ではガイダンスに限らない一般的な場合として、初めは  $y_t$  をベクトルとして扱うことにする。

観測値  $y_t$  と状態変数  $w_t$  の関係を表す演算を  $h_t$ (これを観測演算子と呼ぶ)としたとき、シミュレーション的な考え方でいえば  $y_t=h_t(w_t)$  であり、 $h_t$ と  $w_t$  が与えられれば  $y_t$  は一意に決まるが、通常は両者は一致せず、 $y_t \simeq h_t(w_t)$  である。近似記号を用いる代わりにシステムモデルと同様に

$$\mathbf{y}_t = h_t(\mathbf{w}_t) + \mathbf{v}_t \tag{2.7.4}$$

と書き、この式を観測モデルという。 $v_t$  は  $w_t$  と独立な ノイズで観測ノイズという。ガイダンスの場合には、 $y_t$  は目的変数、 $h_t(w_t)$  は時刻 t の予測値となる。(2.7.4) 式からもわかるように、観測ノイズは観測と予測の差を表す量である。観測ノイズという言葉から測定誤差のことを想像するかもしれないが、観測ノイズには測定誤差のほかに観測モデルが不完全であることによる誤差の両方が寄与していることに注意する。

### 2.7.4 一般状態空間モデル

システムモデルと観測モデルの連立モデルを状態空間モデルという。(2.7.2) 式と(2.7.4) 式をもっと一般的に、

$$\boldsymbol{w}_t \sim p(\boldsymbol{w}_t | \boldsymbol{w}_{t-1}) \tag{2.7.5}$$

$$\mathbf{y}_t \sim p(\mathbf{y}_t | \mathbf{w}_t) \tag{2.7.6}$$

と書いたとき、これを一般状態空間モデルという。後で述べるように、カルマンフィルタは最もシンプルな一般状態空間モデルである線形・ガウス状態空間モデルから導出される。

 $w_t$  や  $y_t$  の分布は、(2.7.5) 式や (2.7.6) 式のように  $w_{t-1}$  や  $w_t$  だけで決まるものではなく、それ以前の状態である  $w_1, w_2, \cdots, w_{t-1}$  や  $y_1, y_2, \cdots, y_t$  にも依存していると思うかもしれない。しかしここでは、時刻 t の条件付確率が時刻 t-1 の状態のみに依存し、それ以前の状態には依存しない、というマルコフ性の 仮定を用いている。すなわち、

$$p(\mathbf{w}_t|\mathbf{w}_{1:t-1},\mathbf{y}_{1:t}) = p(\mathbf{w}_t|\mathbf{w}_{t-1})$$
 (2.7.7)

$$p(y_t|w_{1:t}, y_{1:t-1}) = p(y_t|w_t)$$
 (2.7.8)

であることを仮定している。ここで、 $x_{1:t}$  という表記は  $x_1, x_2, \dots, x_t$  を意味する。

ここではまず、一般状態空間モデルに対して $w_t$ の確率分布を求め、その結果を線形・ガウス状態空間モデルに適用することで、カルマンフィルタでの係数更新式を導出する。

# 2.7.5 逐次ベイズフィルタ

状態空間モデルが与えられた時に、時刻 s (s < t)までの観測データから時刻 t における状態変数の推定

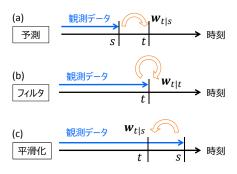


図 2.7.2 予測、フィルタ、平滑化

値を求めることを予測という(図 2.7.2(a) 》。 具体的には、 $w_t|y_{1:s}$  の分布の期待値を求めることに相当する。特に、s=t-1 の場合を一期先予測という。同様に、時刻 t までの観測データから時刻 t の状態変数を推定することをフィルタ(図 2.7.2(b))、時刻 s (s>t) までの観測データから時刻 t の状態変数を推定することを平滑化という(図 2.7.2(c))。 ここで図にも示したように、時刻 s までの観測データから求められた時刻 t における状態変数の推定値を  $w_{t|s}$  と書く。このように書いた場合、 $w_{t|s}$  は確率変数ではなく何らかの確定した値であることに注意する。

ある時刻の状態変数の分布が分かっている場合、予測、フィルタ、平滑化の手続きを繰り返すことで、あらゆる時刻の確率分布を得ることができる。これを逐次ベイズフィルタという。カルマンフィルタは、一期先予測とフィルタを繰り返すことで現在の時刻の状態変数を推定する。

以下では一般状態空間モデルに対して、一期先予測の確率密度関数  $p(\boldsymbol{w}_t|\boldsymbol{y}_{1:t-1})$  とフィルタの確率密度関数  $p(\boldsymbol{w}_t|\boldsymbol{y}_{1:t})$  を書き下す。一期先予測の確率密度関数 は、周辺化 (2.3.40) 式とマルコフ性 (2.7.7) 式を用いることで、次のように書ける。

$$p(\mathbf{w}_{t}|\mathbf{y}_{1:t-1})$$

$$= \int p(\mathbf{w}_{t}|\mathbf{w}_{t-1}, \mathbf{y}_{1:t-1}) p(\mathbf{w}_{t-1}|\mathbf{y}_{1:t-1}) d\mathbf{w}_{t-1}$$

$$= \int p(\mathbf{w}_{t}|\mathbf{w}_{t-1}) p(\mathbf{w}_{t-1}|\mathbf{y}_{1:t-1}) d\mathbf{w}_{t-1} \qquad (2.7.9)$$

またフィルタの確率密度関数は、ベイズの定理 (2.3.38)式、周辺化 (2.3.40) 式とマルコフ性 (2.7.8) 式を用いることで、次のように書ける。

$$p(\boldsymbol{w}_{t}|\boldsymbol{y}_{1:t})$$

$$= \frac{p(\boldsymbol{y}_{t}|\boldsymbol{w}_{t},\boldsymbol{y}_{1:t-1})p(\boldsymbol{w}_{t}|\boldsymbol{y}_{1:t-1})}{\int p(\boldsymbol{y}_{t}|\boldsymbol{w}_{t},\boldsymbol{y}_{1:t-1})p(\boldsymbol{w}_{t}|\boldsymbol{y}_{1:t-1})d\boldsymbol{w}_{t}}$$

$$= \frac{p(\boldsymbol{y}_{t}|\boldsymbol{w}_{t})p(\boldsymbol{w}_{t}|\boldsymbol{y}_{1:t-1})}{\int p(\boldsymbol{y}_{t}|\boldsymbol{w}_{t})p(\boldsymbol{w}_{t}|\boldsymbol{y}_{1:t-1})d\boldsymbol{w}_{t}}$$
(2.7.10)

一期先予測の式に含まれる  $p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1})$  はシス テムモデルから与えられる確率密度関数を表し、  $p(\boldsymbol{w}_{t-1}|\boldsymbol{y}_{1:t-1})$  は時刻 t-1 のフィルタ分布の確率密 度関数を表している。このことは、時刻t-1のフィル タ分布が与えられれば、システムモデルを用いて時刻 t の一期先予測の確率密度関数  $p(oldsymbol{w}_t|oldsymbol{y}_{1:t-1})$  が得られ ることを表している。また、フィルタの式に含まれる  $p(y_t|w_t)$  は観測モデルから与えられる確率密度関数を 表し、 $p(\mathbf{w}_t|\mathbf{y}_{1:t-1})$  は時刻 t-1 の一期先予測の確率密 度関数を表している。このことは、時刻 t-1 の一期 先予測の分布が与えられれば、観測モデルを用いて時 刻 t のフィルタの確率密度関数  $p(\boldsymbol{w}_t|\boldsymbol{y}_{1:t})$  が得られる ことを表している。すなわち、ある時刻の一期先予測 またはフィルタの分布が与えられれば、一期先予測と フィルタを繰り返すことで、それ以降の任意の時刻の 一期先予測とフィルタの分布が得られることを意味し ている。

## 2.7.6 カルマンフィルタ

ここではカルマンフィルタに用いられる線形・ガウス 状態空間モデルに対して一期先予測とフィルタを具体 的に求めることで、ガイダンスにおけるカルマンフィ ルタの係数更新の関係式を導く。

## (1) 線形・ガウス状態空間モデル

線形・ガウス状態空間モデルは以下のように表される状態空間モデルである。

$$\boldsymbol{w}_t = F_t \boldsymbol{w}_{t-1} + G_t \boldsymbol{u}_t \tag{2.7.11}$$

$$\mathbf{y}_t = H_t \mathbf{w}_t + \mathbf{v}_t \tag{2.7.12}$$

ここで、 $w_t$  は K 次元ベクトル、 $F_t$  は  $K \times K$  行列、 $u_t$  は L 次元ベクトル、 $G_t$  は  $K \times L$  行列、 $y_t$  と  $v_t$  は M 次元ベクトル、 $H_t$  は  $M \times K$  行列である。また、 $u_t$  と  $v_t$  は平均 0 の正規分布に従うノイズである。線形・ガウス状態空間モデルは、 $w_t$  と  $w_{t-1}$ 、 $u_t$  の関係、および、 $y_t$  と  $w_t$  の関係が線形であり、システムノイズ と観測ノイズが正規分布で表される状態空間モデルである。

 $F_t,G_t,H_t$  は現象を支配する物理法則や、統計や経験、実験などに基づく関係式や仮定により決められる何らかの既知の行列である。ガイダンスの場合、係数やシステムノイズを時間変化させる法則はない、すなわち、 $F_t \to I$  (単位行列)  $G_tu_t \to u_t$  と考えて差し支えないだろう。また、目的変数はスカラーであるから、 $y_t$ と $v_t$  はスカラーになり、 $H_t$  は K 次元ベクトルになる。よってガイダンスに適用する場合には、線形・ガウス状態空間モデルは以下のようにシンプルな形で書ける。

$$\boldsymbol{w}_t = \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \tag{2.7.13}$$

$$y_t = \boldsymbol{x}_t^T \boldsymbol{w}_t + v_t \tag{2.7.14}$$

ここで、 $w_t$ ,  $x_t$ ,  $u_t$  はいずれも K 次元ベクトル、 $u_t \sim N(0,U_t)$ 、 $v_t \sim N(0,D_t)$  で、 $U_t$  は  $K \times K$  の対称行列、 $D_t$  はスカラーである。

カルマンフィルタの係数更新式を導く前に、カルマンフィルタにおける仮定をまとめておく。最後の仮定については、次の(2)と(3)で述べる。

- 線形・ガウス状態空間モデル: (2.7.13), (2.7.14) 式
- $\boldsymbol{u}_t \sim N(0, U_t)$
- $v_t \sim N(0, D_t)$
- $u_t$ と $w_t$  は独立:  $p(u_t|w_t) = p(u_t)$
- $v_t$ と $\mathbf{w}_t$ は独立:  $p(v_t|\mathbf{w}_t) = p(v_t)$
- マルコフ性: (2.7.7), (2.7.8) 式
- 時刻 t-1 の w のフィルタの分布は正規分布に従う

#### (2) 一期先予測

ー期先予測 (2.7.9) 式をガイダンスのカルマンフィルタについて具体的に求める。まず  $p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1})$  については、システムノイズ  $\boldsymbol{u}_t$  での周辺化 (2.3.40) 式と  $\boldsymbol{u}_t$  と  $\boldsymbol{w}_t$  が独立である仮定より、

$$p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1}) = \int p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1}, \boldsymbol{u}_t) p(\boldsymbol{u}_t) d\boldsymbol{u}_t \quad (2.7.15)$$

となる。 $p(\mathbf{w}_t|\mathbf{w}_{t-1},\mathbf{u}_t)$  はシステムモデル (2.7.13) 式から一意に決まるため、(2.7.3) 式と同様に、

$$p(\mathbf{w}_t | \mathbf{w}_{t-1}, \mathbf{u}_t) = \delta(\mathbf{w}_t - \mathbf{w}_{t-1} - \mathbf{u}_t)$$
 (2.7.16)

と書ける。また、 $m{u}_t \sim N(0,U_t)$  との仮定から、確率密度関数を具体的に書くと

$$p(\boldsymbol{u}_t) = (2\pi)^{-\frac{K}{2}} |U_t|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \boldsymbol{u}_t^T U_t^{-1} \boldsymbol{u}_t\right] (2.7.17)$$

であり、これらを(2.7.15)式に代入すると $w_t|w_{t-1}$ の確率密度関数は以下のようになる。

 $p(\boldsymbol{w}_t|\boldsymbol{w}_{t-1})$ 

$$= (2\pi)^{-\frac{K}{2}} |U_t|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} (\boldsymbol{w}_t - \boldsymbol{w}_{t-1})^T U_t^{-1} (\boldsymbol{w}_t - \boldsymbol{w}_{t-1}) \right]$$
(2.7.18)

すなわち、

$$\mathbf{w}_t | \mathbf{w}_{t-1} \sim N(\mathbf{w}_{t-1}, U_t)$$
 (2.7.19)

であり、 $w_t | w_{t-1}$  は正規分布になる。

次に  $p(\boldsymbol{w}_{t-1}|y_{1:t-1})$  を求めるのだが、これは時刻 t-1 におけるフィルタの分布であり、時刻 t-1 の時点で何らかの方法によって導かれているはずの確率密度関数である。そこで、これについては平均が  $\boldsymbol{w}_{t-1|t-1}$  の K 次元ベクトル、分散共分散行列が  $Q_{t-1|t-1}$  の  $K\times K$ 行列の正規分布に従うと仮定する(第 2.7.6 項 (1) の最後の仮定)。 つまり、

$$|w_{t-1}| y_{1:t-1} \sim N(w_{t-1|t-1}, Q_{t-1|t-1})$$
 (2.7.20)

であると仮定する。この仮定を用いる理由は第 2.7.6 項 (3) で述べることにする。

(2.7.19) 式と (2.7.20) 式が成り立つ場合、 $m{w}_t$  の一期 先予測の分布は、

$$|\boldsymbol{w}_t| y_{1:t-1} \sim N(\boldsymbol{w}_{t|t-1}, Q_{t|t-1})$$
 (2.7.21)

$$\mathbf{w}_{t|t-1} = \mathbf{w}_{t-1|t-1} \tag{2.7.22}$$

$$Q_{t|t-1} = Q_{t-1|t-1} + U_t (2.7.23)$$

となり(証明は付録  $2.7.\mathrm{A}$ )、 $w_t$  の一期先予測の分布は平均が  $w_{t|t-1}$  で分散共分散行列が  $Q_{t|t-1}$  の正規分布になることがわかる。このことを式で書くと、

$$E(\mathbf{w}_{t}|y_{1:t-1}) = \mathbf{w}_{t|t-1}$$

$$V(\mathbf{w}_{t}|y_{1:t-1}) = Q_{t|t-1}$$

$$= E\left[\left(\mathbf{w}_{t} - \mathbf{w}_{t|t-1}\right)\left(\mathbf{w}_{t} - \mathbf{w}_{t|t-1}\right)^{T} \middle| y_{1:t-1}\right]$$
(2.7.25)

となる。 $(2.7.21)\sim(2.7.23)$  式の結果は、時刻 t-1 におけるフィルタの分布が正規分布に従うという仮定を用いて導かれている。すなわち、t-1 のフィルタの分布が正規分布ならば、t-1 の一期先予測の分布もまた正規分布になる、といえる。

### (3) フィルタ

フィルタの確率分布をガイダンスのカルマンフィルタについて具体的に求める。フィルタの確率分布を求めるためには、フィルタの式 (2.7.10) に含まれる  $y_t|w_t$  と  $w_t|y_{1:t-1}$  の分布が分かればよいのだが、 $w_t|y_{1:t-1}$  は (2.7.21) 式で既に求めているので、 $y_t|w_t$  の分布を求める。観測モデル (2.7.14) 式より

$$p(y_t - \boldsymbol{x}_t^T \boldsymbol{w}_t) = p(v_t) \tag{2.7.26}$$

となるので、この両辺に $w_t$ の条件を付け、 $v_t$ と $w_t$ は独立であるという仮定を用いると、

$$p(y_t - \boldsymbol{x}_t^T \boldsymbol{w}_t | \boldsymbol{w}_t) = p(v_t) \tag{2.7.27}$$

となる。観測ノイズ $v_t$  の確率分布は $N(0,D_t)$  に従うと仮定しているので、

$$y_t | \boldsymbol{w}_t \sim N(\boldsymbol{x}_t^T \boldsymbol{w}_t, D_t) \tag{2.7.28}$$

となる。これと (2.7.21) 式が成り立つ場合、時刻 t におけるフィルタ分布 (2.7.10) 式は、

$$|w_t| y_{1:t} \sim N(w_{t|t}, Q_{t|t})$$
 (2.7.29)

$$\boldsymbol{w}_{t|t} = \boldsymbol{w}_{t|t-1} + \boldsymbol{K}_t \nu_t \tag{2.7.30}$$

$$\nu_t = y_t - \boldsymbol{x}_t^T \boldsymbol{w}_{t|t-1} \tag{2.7.31}$$

$$Q_{t|t} = Q_{t|t-1} - \mathbf{K}_t \mathbf{x}_t^T Q_{t|t-1}$$
 (2.7.32)

$$\mathbf{K}_{t} = Q_{t|t-1} \mathbf{x}_{t} \left( \mathbf{x}_{t}^{T} Q_{t|t-1} \mathbf{x}_{t} + D_{t} \right)^{-1}$$
 (2.7.33)

となり(証明は付録 2.7.B)、 $w_t$  のフィルタ分布は平均が  $w_{t|t}$  で分散共分散行列が  $Q_{t|t}$  の正規分布になることがわかる。これは時刻 t の観測値  $y_t$  が得られたときの  $w_t$  の最適な推定値が  $w_{t|t}$  であることを表している。 $\nu_t$  はイノベーションと呼ばれ、予測値と観測値の差を表す。 $K_t$  はカルマンゲインと呼ばれる K 次元ベクトルで、(2.7.30) 式から分かるように、カルマンフィルタでの状態変数更新の手続きにおいて  $w_t$  の変化率を表す量である。(2.7.24) 式,(2.7.25) 式と同様にフィルタ分布の期待値と分散共分散行列を示しておく。

$$E\left(\boldsymbol{w}_{t}|y_{1:t}\right) = \boldsymbol{w}_{t|t}$$

$$V\left(\boldsymbol{w}_{t}|y_{1:t}\right) = Q_{t|t}$$

$$= E\left[\left(\boldsymbol{w}_{t} - \boldsymbol{w}_{t|t}\right)\left(\boldsymbol{w}_{t} - \boldsymbol{w}_{t|t}\right)^{T}|y_{1:t}\right]$$

$$(2.7.34)$$

時刻 t のフィルタ分布は正規分布で表されることを示したが、これは時刻 t-1 の一期先予測の分布が正規分布であることから導かれた結果である。そして時刻 t-1 の一期先予測の分布が正規分布であることは、時刻 t-1 のフィルタ分布が正規分布で表されることから導かれている。すなわち、ある時刻のフィルタ分布が正規分布をしているのであれば、それ以降の全ての時刻において、一期先予測とフィルタの分布は正規分布になる。特定の時刻にこだわらないのであればフィルタ分布の初期値が正規分布に従うと仮定すればよく、これは一般的な(最小分散推定値に基づく)カルマンフィルタの導出に用いられている仮定と一致する。

一般に、分散共分散行列は半正定値である。すなわち、分散共分散行列 Q を  $K \times K$  行列、任意の K 次元ベクトルを x としたとき、 $x^TQx \geq 0$  であり、Q の全ての固有値は 0 以上となる。よって、(2.7.32) 式の両辺に左から  $x_t^T$ 、右から  $x_t$ を掛けると、

$$\mathbf{x}_{t}^{T} Q_{t|t} \mathbf{x}_{t} = \left(\mathbf{x}_{t}^{T} - \mathbf{x}_{t}^{T} \mathbf{K}_{t} \mathbf{x}_{t}^{T}\right) Q_{t|t-1} \mathbf{x}_{t}$$

$$= \left(1 - \frac{\mathbf{x}_{t}^{T} Q_{t|t-1} \mathbf{x}_{t}}{\mathbf{x}_{t}^{T} Q_{t|t-1} \mathbf{x}_{t} + D_{t}}\right) \mathbf{x}_{t}^{T} Q_{t|t-1} \mathbf{x}_{t}$$

$$< \mathbf{x}_{t}^{T} Q_{t|t-1} \mathbf{x}_{t}$$

$$(2.7.36)$$

が成り立つことから、 $Q_{t|t} \preceq Q_{t|t-1}$  である。

一期先予測時の係数の誤差分散  $Q_{t|t-1}$  は前の時刻のフィルタ値  $Q_{t-1|t-1}$  にシステムノイズの分散  $U_t$  を加えた値となっている ((2.7.23) 式 )。つまり、予測によって係数の誤差分散が拡大することを示している。そこに観測値  $y_t$  が得られることにより、フィルタ後の誤差分散  $Q_{t|t}$  は予測時  $Q_{t|t-1}$  よりも小さくなる ((2.7.36) 式 )。このように、カルマンフィルタでは観測値を取り込むことで誤差分散を小さい状態に保ちながら係数を更新していることが分かる。

(4) ガイダンスにおけるカルマンフィルタのまとめ ガイダンスにおけるカルマンフィルタの係数更新の 関係式は、一期先予測の式 (2.7.22), (2.7.23) と、フィ

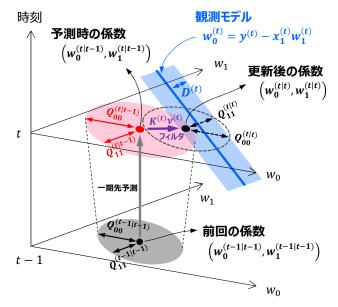


図 2.7.3 カルマンフィルタでの係数更新のイメージ。係数 が 2 個の場合。時刻の添字は ( ) を付けて上付きで示して いる。

ルタの式 (2.7.30)~(2.7.33) になる。また、一期先予測とフィルタの分布はともに正規分布 ( (2.7.21), (2.7.29)式)になる。

実装時の参考とするために、ガイダンスにおけるカルマンフィルタの関係式を行列の成分で書いておく。添字が多くなるため、時刻は()を付けて上付で示す。

$$w_i^{(t|t-1)} = w_i^{(t-1|t-1)} (2.7.37)$$

$$Q_{ij}^{(t|t-1)} = Q_{ij}^{(t-1|t-1)} + U_{ij}^{(t)}$$
 (2.7.38)

$$w_i^{(t|t)} = w_i^{(t|t-1)} + K_i^{(t)} \nu^{(t)}$$
 (2.7.39)

$$\nu^{(t)} = y^{(t)} - \sum_{k=0}^{K} x_k^{(t)} w_k^{(t|t-1)}$$
(2.7.40)

$$Q_{ij}^{(t|t)} = Q_{ij}^{(t|t-1)} - K_i^{(t)} \sum_{k=0}^K x_k^{(t)} Q_{kj}^{(t|t-1)}$$
 (2.7.41)

$$K_i^{(t)} = \frac{\sum_{k=0}^K Q_{ik}^{(t|t-1)} x_k^{(t)}}{\sum_{m,n=0}^K Q_{mn}^{(t|t-1)} x_m^{(t)} x_n^{(t)} + D^{(t)}}$$
(2.7.42)

$$x_0^{(t)} = 1, \quad i, j = 0, \cdots, K$$

係数が 2 個  $(w_0,w_1)$  の場合のカルマンフィルタでの係数更新のイメージを図 2.7.3 に示す。ここでも添字が多くなるため、時刻は ( )を付けて上付で示している。時刻 t-1 でのフィルタ分布から、係数の期待値は  $\left(w_0^{(t-1|t-1)},w_1^{(t-1|t-1)}\right)$ 、分散共分散行列は  $Q^{(t-1|t-1)}$  であり、 $w_0,w_1$  方向の分散は  $Q_{00}^{(t-1|t-1)},Q_{11}^{(t-1|t-1)}$  となる。(2.7.22) 式より、時刻 t での予測時には時刻 t-1 のフィルタの係数をそのまま利用することになるため、係数の値はフィルタの係数と同じだが、(2.7.23) 式より、その分散は時刻 t-1 のフィルタの値と比べてシ

ステムノイズの分だけ大きくなる。このことは図では 赤の楕円に対応しており、フィルタ分布(灰色の楕円) と比べて中心位置は変わらないが広がりは大きくなっ ている。ここで時刻 t の観測  $y^{(t)}$  が得られると、既に 与えられている説明変数  $x_1^{(t)}$  と観測モデル (2.7.14) 式 より、未知の変数である時刻tの係数の真値 $w_0^{(t)}, w_1^{(t)}$ が  $w_0^{(t)} = y^{(t)} - x_1^{(t)} w_1^{(t)}$  を中心として観測ノイズの分 散  $D^{(t)}$  を持った領域 (図中の青の直線および長方形) に分布していることがわかる。フィルタの式 (2.7.10) の右辺の分子は観測モデルの確率密度関数と時刻 t-1の一期先予測の確率密度関数の積になっており、この 2 つの確率密度の積が最も大きくなる係数が時刻 t の フィルタ分布の期待値 $\left(w_0^{(t|t)},w_1^{(t|t)}
ight)$ であり、その分 散共分散行列は $Q^{(t|t)}$ となる。このとき $D^{(t)}$ を大きな 値に設定すると図中の青の長方形で示した領域が広が り、更新後の係数は予測時の係数に近い値になる。逆 にシステムノイズ $U^{(t)}$ を大きな値に設定すると図中の 赤の楕円で示した領域が広がり、更新後の係数は観測 値に近い値になる。相対的に $D^{(t)}$ が小さい場合には、 新たな観測が得られる度に観測に寄せるように係数が 決まるため、係数の変動幅が大きくなる。逆に $D^{(t)}$ が 大きい場合には係数の変動幅は小さくなり、予測と実 況の差が大きかったとしても係数はあまり変化しなく なる。予測時の係数を更新後の係数に変化させるベク トル ( 図中の紫矢印 ) は、イノベーション  $\nu^{(t)}$  にカル マンゲイン  $K^{(t)}$  を掛けたものになる。図で破線の楕円 や第 2.7.6 項 (3) で示したように、観測値の情報を得る ことで、フィルタ分布の分散は一期先予測の分散より も小さくなる。以上のことを繰り返すことで、カルマ ンフィルタでは最新の観測値の情報を取り込みながら 係数を逐次更新する。

## 2.7.7 整合性の確認

我々は通常、カルマンフィルタで推定される状態変数の真値を知ることはできない。よって推定値の正しさを直接的に見積もることはできないが、イノベーションを用いることでカルマンフィルタによる状態推定が正しく行われているか確認することができる。

カルマンフィルタにおいて、イノベーションは平均が 0 の正規分布に従うホワイトノイズであるという性質を持つ (例えば Brown and Rutan 1985; Jwo and Cho 2007; Bulut 2011)。 すなわち、

$$E(\nu_t|y_{1:t-1}) = 0 (2.7.43)$$

$$E\left(\nu_{t}\nu_{s}^{T}|y_{1:t-1}\right) = 0 \quad (t > s) \tag{2.7.44}$$

### であり、その分散は

$$S_{t} \equiv V\left(\nu_{t} | y_{1:t-1}\right) = E\left(\nu_{t} \nu_{t}^{T} | y_{1:t-1}\right)$$

$$= E\left[\left(y_{t} - \boldsymbol{x}_{t}^{T} \boldsymbol{w}_{t|t-1}\right) \left(y_{t} - \boldsymbol{x}_{t}^{T} \boldsymbol{w}_{t|t-1}\right)^{T} | y_{1:t-1}\right]$$

$$= \boldsymbol{x}_{t}^{T} Q_{t|t-1} \boldsymbol{x}_{t} + D_{t}$$

$$(2.7.45)$$

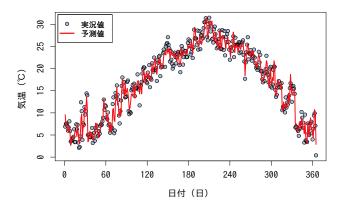


図 2.7.4 カルマンフィルタによる福岡の午前 9 時の気温予測。 横軸は左端が 2014 年 1 月 1 日 00UTC 初期値、右端が 2014 年 12 月 31 日 00UTC 初期値。

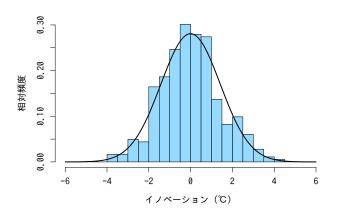


図 2.7.6 図 2.7.4 の気温予測におけるイノベーションの相対頻度のヒストグラム。実線は期間内のイノベーションから求めた平均と標準偏差をもつ正規分布。

と書ける。ここで (2.7.14) 式、(2.7.25) 式などを用いた。 $S_t$  はイノベーションの分散の理論値といえる。もし第 2.7.6 項 (1) で述べた仮定が満たされているならば、イノベーションの分布は (2.7.43) 式と (2.7.45) 式を満たす正規分布であり、 $\pm \sqrt{S_t}$  の範囲に約 68%、 $\pm 2\sqrt{S_t}$  の範囲に約 95%のイノベーションが含まれることになる。イノベーションを時系列などでプロットすることでこれらが満たされているか確認できる  $^2$ 。

図 2.7.4 は、カルマンフィルタによる福岡の午前 9 時の気温予測の例である。この例ではルーチンの気温ガイダンスと同様の説明変数と実況を用いているが、初期値や  $D_t$ ,  $U_t$  の設定などは簡易的な値を用いている。学習期間は 2013 年の 1 年間とし、00UTC 初期値の 24 時間予測について、2014 年 1 月 1 日から 12 月 31 日までの 1 年間 ( 365 日分 ) の予測結果を表示している。この期間の数値予報モデルの地上気温の予測は ME が -2.44 °C、RMSE が 2.93 °C であるのに対し、カルマン

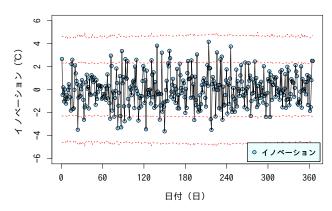


図 2.7.5 図 2.7.4 の気温予測におけるイノベーションの時系列。横軸は図 2.7.4 と同じ。赤破線は  $\pm \sqrt{S_t}$  および  $\pm 2\sqrt{S_t}$  を表す。

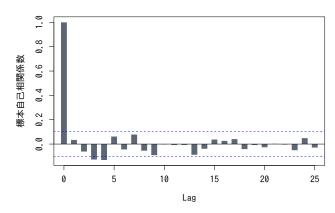


図 2.7.7 図 2.7.4 の気温予測におけるイノベーションの標本 自己相関係数。横軸はラグで、青破線はホワイトノイズを 仮定した場合の標本自己相関係数の 95%信頼区間 (例えば Shumway and Stoffer 2000)。

フィルタの予測は ME が -0.00 °C、RMSE が 1.42 °C であり、数値予報モデルを大幅に改善している。図を見る限りでも実況との対応は概ね良さそうであり、カルマンフィルタによる予測は適切に行われているように思われる。

図 2.7.4 と同じ気温予測の時系列について、イノベーションの時系列、相対頻度のヒストグラム、標本自己相関係数を図 2.7.5  $\sim$  図 2.7.7 に示す。イノベーションの時系列を見ると、日付によらず 0 を中心として同程度のばらつきを持っていることがわかる。またヒストグラムより、平均が 0 の正規分布に近い分布をしており、標本自己相関係数からは時間方向の相関が弱いことがわかる。これらのことから、イノベーションの性質のうち、平均 0 の正規分布に従うホワイトノイズであるということについては満たされていそうである。一方イノベーションのばらつきを  $S_t$  と比較すると、期待されるよりも狭い範囲にイノベーションが集中しており、イノベーションの分散が過小であるといえる。

イノベーションの分散は直接的には  $D_t$ と  $U_t$  の大き さによって変化するため、分散が過小または過大であ

 $<sup>^2</sup>$  より厳密には、標準化 2 乗イノベーション  $u_t^T S_t^{-1} 
u_t$  の移動平均が自由度 m (ガイダンスの場合は目的変数がスカラーであるため m=1 )の  $\chi^2$  分布に従うことを確かめればよい。

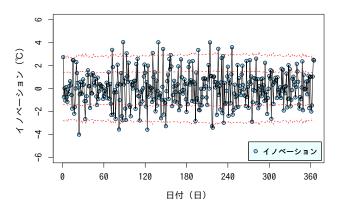


図 2.7.8 図 2.7.5 と同じ。ただし、 $D_t$  を調整してイノベーションのばらつきを調整した結果。

る場合には  $D_t$ や  $U_t$  を調整すれば適切なばらつきを持つように修正することはできる。しかし、目的変数と説明変数の関係が線形になっていない場合やノイズが正規分布ではない場合など、カルマンフィルタの仮定がそもそも満たされていないならば、 $D_t$ や  $U_t$  だけを調整することは予測精度の低下につながる場合がある。図 2.7.8 は上記の気温予測について、 $D_t$  を調整してイノベーションのばらつきが適切に近くなるようにした場合の実験結果である。この図を見ると、 $\pm \sqrt{S_t}$  の範囲に適切に近い数のイノベーションが含まれていることがわかる。また図には示さないが、予測値の時系列、イノベーションの相対頻度と標本自己相関には不適切な点は見られない。しかし予測精度は、ME が 0.001 °C、RMSE が 1.46 °C となっており、調整前と比べてやや低下している。

最後に不適切なイノベーションの分布の例を図 2.7.9 に示す。この図は図 2.7.4 と同じ気温ガイダンスに用 いられている説明変数のうち、ガイダンスへの寄与量 が小さい(すなわちあまり重要ではない)説明変数を 1 つだけ用いて気温を予測した場合のイノベーション の時系列である。このイノベーションは0を中心とし た分布をしておらず、ばらつきの大きさは理論値より も大きくなっている。また、時間方向に相関を持つこ とからホワイトノイズではなく、ばらつきの大きさと しても時間によって差が見られる。よってこの例では、 イノベーションの性質をいずれも満たしていない。し かしながら図 2.7.4 で示したように、このような説明 変数を用いているにもかかわらず、カルマンフィルタ としては一見すると正常な動作をしているように見え、 また、この説明変数を除いた場合には予測精度として は低下してしまう。このことは、この説明変数の見直 しや、他の説明変数も含めた全体的な見直しが必要で あることを示唆している。

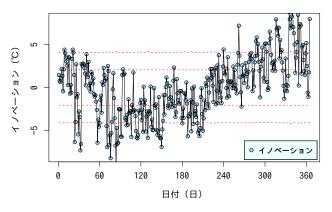


図 2.7.9 不適切なイノベーションの例。図の見方は図 2.7.5 や 図 2.7.8 と同じ。

### 2.7.8 パラメータの調整

ガイダンスのカルマンフィルタにおいて、 $D_t, U_t$  お よび係数と  $Q_{t|t}$  の初期値は既知のものとして扱われて おり、開発者が事前に設定しなければならないパラメー タである。これらのうち係数の初期値は、予測精度に 大きな影響は与えないため3乱数や0を与えても問題 ない。また  $Q_{t|t}$  の初期値についても、乱数を与えても 経験的には適切な値に収束するため大きな問題はない。 一方、 $D_t$ と $U_t$  は係数の変動幅を決める重要なパラメー タであるため、乱数などを与えることは不適切であり、 何らかの方法で適切な値を与える必要がある。 $D_t$  は観 測ノイズの分散であるから、過去の一定期間のデータ から算出した予測の平均二乗誤差を元に大まかな値を 見積もることができる。また $U_t$ はシステムノイズの分 散共分散行列であるから、(2.7.13) 式より、今回と前回 の係数の差の二乗和から大まかな値を見積もることが できる。気象庁のガイダンスでは、 $D_t$  と  $U_t$  の非対角 成分は0と仮定し、上記の見積りを元に、学習データ に対して繰り返し計算しながら  $D_t$  と  $U_t$  (時間変化さ せる場合にはその初期値)を決定している。

より効率的に  $D_t$ と  $U_t$  を見積もる方法として、最尤法に基づく手法 (Shumway and Stoffer 2000) と EM アルゴリズム (Dempster et al. 1977; Shumway and Stoffer 1982; Ghahramani and Hinton 1996; Shumway and Stoffer 2000) が提案されている。これらの手法はいずれも、繰り返し計算によってパラメータを見積もっている。ここでは  $D_t, U_t$  などのパラメータをまとめて  $\theta$  と置く。以下では  $\theta$  は時間変化しないものとして、最尤法に基づいたパラメータの推定方法を示す。

イノベーション  $\nu_t$  は平均 0、分散  $S_t$  の正規分布に 従うことから、イノベーションの確率密度関数は

$$p(\nu_t) = \frac{1}{\sqrt{2\pi S_t}} \exp\left(-\frac{\nu_t^2}{2S_t}\right)$$
 (2.7.46)

 $<sup>^3</sup>$  例えば (瀬上ほか 1995) の第 5.1.3 項での理想実験を参照。係数を推定する手法がカルマンフィルタなのであるから当然のことともいえる。

である。今、時刻  $t=1\sim N$  のデータが与えられているとする。イノベーションの負の対数尤度を誤差関数 E とすると、E は定数項を除いて

$$E \equiv -\ln L = \frac{1}{2} \sum_{t=1}^{N} \left( \ln S_t + \frac{\nu_t^2}{S_t} \right)$$
 (2.7.47)

となる。ここで  $S_t$ や $\nu_t$ は $\theta$  の関数である。 $\theta \to \pm \Delta \theta$  とした場合の誤差関数の値を  $E_\pm$  とすると、 $\theta$  方向の E の微分は

$$\frac{\partial E}{\partial \theta} \simeq \frac{E_{+} - E_{-}}{2\Delta \theta} \tag{2.7.48}$$

となる。同様にして  $\theta$  の 2 階微分も得られる。 $\theta$  の更新は以下の手順で行う。

- 1.  $\theta$  の初期値を与える。
- 2. 学習期間のデータ (  $t=1\sim N$  ) に対してカルマンフィルタによる係数更新を行い、 $\nu_t$ と  $S_t$  の時系列を求める。
- $3.~\nu_t$ と  $S_t$  および E の微分を用いてニュートン・ラフソン法などの繰り返し計算を 1 回分だけ行い、 $\theta$  を更新する。
- 4. 上記の  $2 \sim 3$  の手順を繰り返し、誤差関数が変化しなくなったとみなせるところで計算を終了する。

見積もるパラメータが多いことから、実際に計算する場合には予測への影響が小さいパラメータ (係数や  $Q_{t|t}$  の初期値など)には乱数や 0 を与えることとし、その他のパラメータについて上記の方法で見積もることになるだろう。

上記の手法では $U_t$ と $D_t$ を固定値として扱っている が、 $D_t$  は観測モデルが不完全であることも含めた誤差 を表しており、予測が難しい場合には $D_t$ を大きく、簡 単な場合には小さくすることが望ましい。しかし予測 の難しさを事前に知ることは困難である4ため、本稿 を執筆している 2018 年現在のガイダンスでは  $D_t$  は固 定値か、時間変化が緩やかな固定値に近い値として扱 われている。例えば平均降水量ガイダンスでは、前回 までの  $D_t$  と今回の予測の二乗誤差を重み付け平均す ることで  $D_t$  を更新しているが、前回までの  $D_t$  の重 みを非常に大きくしており、場の状況に応じて変化さ せているというよりは長期的な予測誤差の変化に対応 させることを目的としている。このような事情により、 ガイダンスの予測が大きく外れた場合には、実況に合 わせるために係数を大きく変化させてしまうことがあ る。しかし係数の大きな変化は、その後の予測で極端 な値を予測してしまうなど、一般に予測精度を低下さ せる要因となる。そこで気温ガイダンスでは、 $D_t$  は固 定値としているが、予測が大きく外れた場合には、係 数更新時に一時的に  $D_t$  を大きくして係数を実況に寄 せすぎない (大きく変化させない)ようにしている。

### 2.7.9 利用上の注意点

カルマンフィルタは線形・ガウス状態空間モデルに基づいて係数を更新しており、第 2.7.6 項 (1) で述べた仮定が満たされている必要がある。カルマンフィルタでは係数が変動するため、特にシステムモデルについては線形性やノイズの正規性を直接的に確認することは難しいが、観測モデルについては各目的変数と説明変数の関係をプロットすることで線形性やノイズの正規性をある程度確認することができる。新規に説明変数を追加する場合やガイダンスの改良を行う場合には、目的変数と説明変数の関係をプロットして線形性や正規性を確認するとよい。

第 2.7.7 項で述べたように、イノベーションの分布を時系列で見ることによって、カルマンフィルタによる係数更新が適切に行われているか否かを調べることができる。イノベーションの性質が満たされていない場合には、 $D_t, U_t$  および係数と  $Q_{t|t}$  の初期値の設定や、目的変数と説明変数の関係が適切であるかを確認する。第 2.7.7 項で示したように、上記のどれか一つを調整しただけでは全体としてはバランスが崩れて予測精度が低下することもある。このような場合には全体的な調整をすることが精度向上につながるだろう。

カルマンフィルタは係数の最適な推定値を逐次更新するため、数値予報モデルが更新されて予測特性が大きく変化した場合でも一定期間学習すれば適切な係数が自動的に得られると期待される。しかし、一定期間がどの程度かは対象としている現象の頻度に依存する。気温ガイダンスの場合には 2,3 週間程度の学習で適切に近い値まで学習が進むが、強風や大雨など頻度が少ない現象を対象としたガイダンスでは数か月以上掛かる。このためカルマンフィルタを用いたガイダンスであっても、数値予報モデルの特性が大きく変化する場合には、長期間の過去データを用いた事前学習が必要になる場合がある。

カルマンフィルタの仮定が満たされていないような 目的変数と説明変数に対しても、カルマンフィルタは ある程度適切な予測結果を与える。これは線形重回帰 やロジスティック回帰でも同様で、各手法で仮定して いる条件が成り立たない場合であっても、これらの統 計手法はある程度適切な(精度を持つ)予測結果を与 える。一括学習型の統計手法を用いた場合には、悪い 部分も含めて予測特性は変化しないため、値が異常で あることや不適切であることは判別しやすい。これに 対してカルマンフィルタのような逐次学習型の統計手 法の場合、手法の仮定が満たされていない状況では、 ある時に突然異常な係数を学習してしまう可能性があ る。そして予測特性が常に変化するという特性は、そ の変化が正常の範囲内であるのか異常であるのかを判 断することを困難にしてしまう。カルマンフィルタを 用いたガイダンスを開発・運用するに当たっては、カ ルマンフィルタの仮定との整合性や係数の日々の変化

 $<sup>^4</sup>$  アンサンブル予報のスプレッドが適切である場合、スプレッドを用いれば  $D_t$  を事前に見積もることができるかもしれない。

に注意する必要がある。

### 2.7.10 まとめ

本節では、カルマンフィルタを用いてガイダンスを 算出する場合の関係式をベイズ推定に基づく係数の条 件付き期待値から導出した。また、イノベーションの 時系列を用いた整合性の確認手法と、ガイダンスを開 発する上で必要となるパラメータの調整について述べ た。カルマンフィルタを用いることで、係数の特性が時 間変化する場合にも係数が逐次調整されていくことか ら、2018 年現在では様々なガイダンスにカルマンフィ ルタが利用されている。一方で、ガイダンスの開発・ 運用においては注意すべき点も多い。目的変数と説明 変数の特性や関係性を把握しながら開発を進めること が重要である。

### 参考文献

- Brown, S. D. and S. C. Rutan, 1985: Adaptive Kalman filtering. *Journal of Research of the National Bureau of Standards*, **90**, 403–407.
- Bulut, Y., 2011: Applied kalman filter theory. *Civil Engineering Dissertations*, **13**.
- Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B*, **39(1)**, 1–38.
- Ghahramani, Z. and G. E. Hinton, 1996: Parameter estimation for linear dynamical systems. *Technical Report CRG-TR-96-2*, 1–6.
- 樋口知之, 2011: データ同化入門. 朝倉書店, 240 pp.
- 堀田大介, 太田洋一郎, 2011: ローレンツモデルによる 学習. 数値予報課報告・別冊第 57 号, 気象庁予報部, 144-158.
- Jwo, D. and T. Cho, 2007: A practical note on evaluating Kalman filter performance optimality and degradation. Applied Mathematics and Computation, 193, 482–505.
- 片山徹, 2000: 新版 応用カルマンフィルタ. 朝倉書店, 255 pp.
- 北川源四郎, 1993: FORTRAN77 時系列解析プログラミング. 岩波書店, 390 pp.
- Persson, A., 1989: Kalman filtering a new approach to adaptive statistical interpretation of numerical meteorological forecasts. *ECMWF Newsletter*, **46**, 16–20.
- Persson, A., 1991: Kalman filtering a new approach to adaptive statistical interpretation of numerical meteorological forecasts. WMO/TD, 421, XX27–XX32.
- 瀬上哲秀, 大林正典, 國次雅司, 藤田司, 1995: カルマン フィルター. 平成7年度数値予報研修テキスト, 気象

**庁予報部**, 66-78.

- Shumway, R. H. and D. S. Stoffer, 1982: An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, 3, 253–264.
- Shumway, R. H. and D. S. Stoffer, 2000: *Time series analysis and its applications*. Springer, 549 pp.
- Simonsen, C., 1991: Self adaptive model output statistics based on kalman filtering. WMO/TD, 421, XX33–XX37.

### 付録 2.7.A 一期先予測の関係式の導出

ここでは、 $w_{t-1}|y_{1:t-1}\sim N(w_{t-1|t-1},Q_{t-1|t-1})$  および  $w_t|w_{t-1}\sim N(w_{t-1},U_t)$  のとき、ガイダンスのカルマンフィルタにおいて、下記の一期先予測の関係式が成り立つことを示す。

$$\mathbf{w}_t | y_{1:t-1} \sim N(\mathbf{w}_{t|t-1}, Q_{t|t-1})$$
 (2.7.49)

$$\mathbf{w}_{t|t-1} = \mathbf{w}_{t-1|t-1} \tag{2.7.50}$$

$$Q_{t|t-1} = Q_{t-1|t-1} + U_t (2.7.51)$$

以下では表記を簡単にするため、一時的に  $w_{t-1} \to v$ ,  $w_t \to w$ ,  $w_{t-1|t-1} \to \hat{v}$ ,  $y_{1:t-1} \to y$ ,  $Q_{t-1|t-1} \to Q$ ,  $U_t \to U$  と記述する。y はスカラー、v と w は K 次元ベクトル、Q と U は  $K \times K$  対称行列である。 与えられた条件より、

$$p(\mathbf{v}|y) = (2\pi)^{-\frac{K}{2}} |Q|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\mathbf{v} - \hat{\mathbf{v}})^T Q^{-1} (\mathbf{v} - \hat{\mathbf{v}})\right]$$
(2.7.52)

$$p(\boldsymbol{w}|\boldsymbol{v}) = (2\pi)^{-\frac{K}{2}} |U|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{w} - \boldsymbol{v})^T U^{-1} (\boldsymbol{w} - \boldsymbol{v})\right]$$
(2.7.53)

である。w|y の確率密度関数は、周辺化 (2.3.40) 式とマルコフ性 (2.7.7) 式より、

$$p(\boldsymbol{w}|y) = \int p(\boldsymbol{w}|\boldsymbol{v}, y)p(\boldsymbol{v}|y)d\boldsymbol{v} = \int p(\boldsymbol{w}|\boldsymbol{v})p(\boldsymbol{v}|y)d\boldsymbol{v}$$
(2.7.54)

と書けるので、右辺の被積分関数は以下のようになる。

$$p(\boldsymbol{w}|\boldsymbol{v})p(\boldsymbol{v}|y) = (2\pi)^{-K}|Q|^{-\frac{1}{2}}|U|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\left\{(\boldsymbol{v}-\hat{\boldsymbol{v}})^{T}Q^{-1}(\boldsymbol{v}-\hat{\boldsymbol{v}}) + (\boldsymbol{w}-\boldsymbol{v})^{T}U^{-1}(\boldsymbol{w}-\boldsymbol{v})\right\}\right]$$
(2.7.55)

これを v で積分することから、上式の右辺を v の頃とそれ以外の頃に分けることを考える。(2.7.55) 式の右辺の  $\{\ \}$  内を展開して整理すると、

$$\{ \} = \boldsymbol{v}^{T} Q^{-1} \boldsymbol{v} - \boldsymbol{v}^{T} Q^{-1} \hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}^{T} Q^{-1} \boldsymbol{v} + \hat{\boldsymbol{v}}^{T} Q^{-1} \hat{\boldsymbol{v}} + \boldsymbol{w}^{T} U^{-1} \boldsymbol{w} - \boldsymbol{w}^{T} U^{-1} \boldsymbol{v} - \boldsymbol{v}^{T} U^{-1} \boldsymbol{w} + \boldsymbol{v}^{T} U^{-1} \boldsymbol{v}$$

$$= \boldsymbol{v}^{T} (Q^{-1} + U^{-1}) \boldsymbol{v} - \boldsymbol{v}^{T} (Q^{-1} \hat{\boldsymbol{v}} + U^{-1} \boldsymbol{w}) - (\hat{\boldsymbol{v}}^{T} Q^{-1} + \boldsymbol{w}^{T} U^{-1}) \boldsymbol{v} + \hat{\boldsymbol{v}}^{T} Q^{-1} \hat{\boldsymbol{v}} + \boldsymbol{w}^{T} U^{-1} \boldsymbol{w}$$

$$(2.7.56)$$

ここで、ある K 次元ベクトル  $\xi$  を用いて、

$$(\boldsymbol{v} - \boldsymbol{\xi})^{T} (Q^{-1} + U^{-1}) (\boldsymbol{v} - \boldsymbol{\xi}) = \boldsymbol{v}^{T} (Q^{-1} + U^{-1}) \boldsymbol{v} - \boldsymbol{v}^{T} (Q^{-1} + U^{-1}) \boldsymbol{\xi} - \boldsymbol{\xi}^{T} (Q^{-1} + U^{-1}) \boldsymbol{v} + \boldsymbol{\xi}^{T} (Q^{-1} + U^{-1}) \boldsymbol{\xi}$$

$$(2.7.57)$$

と書けることから、

$$\boldsymbol{\xi} \equiv (Q^{-1} + U^{-1})^{-1} (Q^{-1}\hat{\boldsymbol{v}} + U^{-1}\boldsymbol{w})$$
(2.7.58)

とすると、Q と U が対称行列であることから、 $oldsymbol{\xi}^T$  は以下のようになる。

$$\boldsymbol{\xi}^{T} = (Q^{-1}\hat{\boldsymbol{v}} + U^{-1}\boldsymbol{w})^{T} \left\{ (Q^{-1} + U^{-1})^{T} \right\}^{-1} = (\hat{\boldsymbol{v}}^{T}Q^{-1} + \boldsymbol{w}^{T}U^{-1}) (Q^{-1} + U^{-1})^{-1}$$
(2.7.59)

この **ξ** を用いると、(2.7.56) 式は、

$$\{ \} = (\boldsymbol{v} - \boldsymbol{\xi})^T (Q^{-1} + U^{-1}) (\boldsymbol{v} - \boldsymbol{\xi}) - \boldsymbol{\xi}^T (Q^{-1} + U^{-1}) \boldsymbol{\xi} + \hat{\boldsymbol{v}}^T Q^{-1} \hat{\boldsymbol{v}} + \boldsymbol{w}^T U^{-1} \boldsymbol{w}$$
(2.7.60)

となり、被積分関数を v に関する項(第 1 項)とそれ以外の項に分けることができた。ここで (2.7.60) 式の右辺第 2 項を展開すると、

$$-\boldsymbol{\xi}^{T} \left( Q^{-1} + U^{-1} \right) \boldsymbol{\xi} = -\left( \hat{\boldsymbol{v}}^{T} Q^{-1} + \boldsymbol{w}^{T} U^{-1} \right) \left( Q^{-1} + U^{-1} \right)^{-1} \left( Q^{-1} \hat{\boldsymbol{v}} + U^{-1} \boldsymbol{w} \right)$$

$$= -\hat{\boldsymbol{v}}^{T} Q^{-1} \left( Q^{-1} + U^{-1} \right)^{-1} Q^{-1} \hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}^{T} Q^{-1} \left( Q^{-1} + U^{-1} \right)^{-1} U^{-1} \boldsymbol{w}$$

$$- \boldsymbol{w}^{T} U^{-1} \left( Q^{-1} + U^{-1} \right)^{-1} Q^{-1} \hat{\boldsymbol{v}} - \boldsymbol{w}^{T} U^{-1} \left( Q^{-1} + U^{-1} \right)^{-1} U^{-1} \boldsymbol{w} \quad (2.7.61)$$

であり、

$$U^{-1} (Q^{-1} + U^{-1})^{-1} Q^{-1} = [Q (Q^{-1} + U^{-1}) U]^{-1} = (Q + U)^{-1}$$
(2.7.62)

$$Q^{-1} (Q^{-1} + U^{-1})^{-1} U^{-1} = [U (Q^{-1} + U^{-1}) Q]^{-1} = (Q + U)^{-1}$$
(2.7.63)

$$Q^{-1} (Q^{-1} + U^{-1})^{-1} Q^{-1} = (Q^{-1} + U^{-1}) (Q^{-1} + U^{-1})^{-1} Q^{-1} - U^{-1} (Q^{-1} + U^{-1})^{-1} Q^{-1}$$

$$= Q^{-1} - (Q + U)^{-1}$$
(2.7.64)

$$U^{-1} (Q^{-1} + U^{-1})^{-1} U^{-1} = U^{-1} - (Q + U)^{-1}$$
(2.7.65)

であることを用いると、(2.7.60) 式の右辺第2項は、

$$-\boldsymbol{\xi}^{T} (Q^{-1} + U^{-1}) \boldsymbol{\xi}$$

$$= -\hat{\boldsymbol{v}}^{T} Q^{-1} \hat{\boldsymbol{v}} + \hat{\boldsymbol{v}}^{T} (Q + U)^{-1} \hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}^{T} (Q + U)^{-1} \boldsymbol{w} - \boldsymbol{w}^{T} (Q + U)^{-1} \hat{\boldsymbol{v}} - \boldsymbol{w}^{T} U^{-1} \boldsymbol{w} + \boldsymbol{w}^{T} (Q + U)^{-1} \boldsymbol{w}$$
(2.7.66)

となる。よって (2.7.60) 式の右辺第 2,3,4 項を  $\Delta$  と置くと、(2.7.60) 式の右辺第 3,4 項と (2.7.66) 式の右辺第 1,5 項が打ち消し合うため、

$$\Delta \equiv \hat{\boldsymbol{v}}^{T} (Q+U)^{-1} \hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}^{T} (Q+U)^{-1} \boldsymbol{w} - \boldsymbol{w}^{T} (Q+U)^{-1} \hat{\boldsymbol{v}} + \boldsymbol{w}^{T} (Q+U)^{-1} \boldsymbol{w}$$

$$= (\boldsymbol{w} - \hat{\boldsymbol{v}})^{T} (Q+U)^{-1} (\boldsymbol{w} - \hat{\boldsymbol{v}})$$
(2.7.67)

となる。これで準備ができたので (2.7.54) 式の計算に戻る。

$$p(\boldsymbol{w}|y) = (2\pi)^{-K} |Q|^{-\frac{1}{2}} |U|^{-\frac{1}{2}} \int \exp\left[-\frac{1}{2} \left\{ (\boldsymbol{v} - \hat{\boldsymbol{v}})^T Q^{-1} (\boldsymbol{v} - \hat{\boldsymbol{v}}) + (\boldsymbol{w} - \boldsymbol{v})^T U^{-1} (\boldsymbol{w} - \boldsymbol{v}) \right\} \right] d\boldsymbol{v}$$

$$= (2\pi)^{-K} |Q|^{-\frac{1}{2}} |U|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\boldsymbol{w} - \hat{\boldsymbol{v}})^T (Q + U)^{-1} (\boldsymbol{w} - \hat{\boldsymbol{v}})\right] \int \exp\left[-\frac{1}{2} (\boldsymbol{v} - \boldsymbol{\xi})^T (Q^{-1} + U^{-1}) (\boldsymbol{v} - \boldsymbol{\xi})\right] d\boldsymbol{v}$$

$$= (2\pi)^{-\frac{K}{2}} |Q|^{-\frac{1}{2}} |Q^{-1} + U^{-1}|^{-\frac{1}{2}} |U|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\boldsymbol{w} - \hat{\boldsymbol{v}})^T (Q + U)^{-1} (\boldsymbol{w} - \hat{\boldsymbol{v}})\right]$$

$$(2.7.68)$$

ここで正規分布の積分 (2.3.43) 式を用いた。Q と U はともに  $K \times K$  対称行列であり、|QU| = |Q||U| となることを用いると、

$$|Q|^{-\frac{1}{2}}|Q^{-1} + U^{-1}|^{-\frac{1}{2}}|U|^{-\frac{1}{2}} = |Q + U|^{-\frac{1}{2}}$$
(2.7.69)

であるから、

$$p(\boldsymbol{w}|y) = (2\pi)^{-\frac{K}{2}}|Q + U|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{w} - \hat{\boldsymbol{v}})^T (Q + U)^{-1}(\boldsymbol{w} - \hat{\boldsymbol{v}})\right]$$
(2.7.70)

となり、 $w|y\sim N(\hat{v},Q+U)$  となる。元の表記で書けば、 $w_t|y_{1:t-1}\sim N(w_{t-1|t-1},Q_{t-1|t-1}+U_t)$  である。ここで  $w_t|y_{1:t-1}$  は時刻 t-1 の一期先予測の分布であるから、その期待値と分散共分散行列は  $w_{t|t-1}$ と  $Q_{t|t-1}$  であるので、

$$w_{t|t-1} \equiv w_{t-1|t-1}$$
 (2.7.71)

$$Q_{t|t-1} \equiv Q_{t-1|t-1} + U_t \tag{2.7.72}$$

とすると、

$$\mathbf{w}_t | y_{1:t-1} \sim N(\mathbf{w}_{t|t-1}, Q_{t|t-1})$$
 (2.7.73)

が得られる。

### 付録 2.7.B フィルタの関係式の導出

ここでは、 $y_t|w_t \sim N(x_t^Tw_t, D_t)$  および  $w_t|y_{1:t-1} \sim N(w_{t|t-1}, Q_{t|t-1})$  のとき、フィルタ分布の式 (2.7.10) を解くことで、ガイダンスのカルマンフィルタにおいて、下記のフィルタの関係式が成り立つことを示す。

$$\mathbf{w}_{t}|y_{1:t} \sim N(\mathbf{w}_{t|t}, Q_{t|t})$$
 (2.7.74)

$$\mathbf{w}_{t|t} = \mathbf{w}_{t|t-1} + \mathbf{K}_t \left( y_t - \mathbf{x}_t^T \mathbf{w}_{t|t-1} \right)$$
(2.7.75)

$$Q_{t|t} = Q_{t|t-1} - \mathbf{K}_t \mathbf{x}_t^T Q_{t|t-1}$$
(2.7.76)

$$K_{t} = Q_{t|t-1}x_{t} \left(x_{t}^{T} Q_{t|t-1}x_{t} + D_{t}\right)^{-1}$$
(2.7.77)

以下では表記を簡単にするため、一時的に  $w_t \to w$ ,  $w_{t|t-1} \to \hat{w}$ ,  $Q_{t|t-1} \to Q$ ,  $y_t \to y$ ,  $y_{1:t-1} \to z$ ,  $x_t \to x$ ,  $D_t \to D$  と記述する。y, z, D はスカラー、 $w, \hat{w}, x$  は K 次元ベクトル、Q は  $K \times K$  対称行列である。 与えられた条件より、

$$p(y|\mathbf{w}) = (2\pi)^{-\frac{1}{2}}|D|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(y - \mathbf{x}^T\mathbf{w})^T D^{-1}(y - \mathbf{x}^T\mathbf{w})\right]$$
(2.7.78)

$$p(\boldsymbol{w}|z) = (2\pi)^{-\frac{K}{2}}|Q|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\boldsymbol{w} - \hat{\boldsymbol{w}})^{T}Q^{-1}(\boldsymbol{w} - \hat{\boldsymbol{w}})\right]$$
(2.7.79)

と書ける。(2.7.78) 式は 1 変数の正規分布の確率密度関数であるためもっとシンプルに書けるが、ここでは (2.7.79) 式との対称性を保持するため、 $1\times 1$  行列として扱うことにする。フィルタ分布の式 (2.7.10) の分母の被積分関数と分子はこの 2 つの確率密度関数を掛けた関数である。

$$p(y|\mathbf{w})p(\mathbf{w}|z) = (2\pi)^{-\frac{K+1}{2}}|D|^{-\frac{1}{2}}|Q|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\left\{(y-\mathbf{x}^T\mathbf{w})^TD^{-1}(y-\mathbf{x}^T\mathbf{w}) + (\mathbf{w}-\hat{\mathbf{w}})^TQ^{-1}(\mathbf{w}-\hat{\mathbf{w}})\right\}\right]$$
(2.7.80)

(2.7.10) 式では w で積分することから、右辺の  $\{\ \}$  内を w の項とそれ以外の項に分けることを考える。 $\{\ \}$  内を展開して整理すると、

$$\{ \} = \boldsymbol{w}^{T} Q^{-1} \boldsymbol{w} - \boldsymbol{w}^{T} Q^{-1} \hat{\boldsymbol{w}} - \hat{\boldsymbol{w}}^{T} Q^{-1} \boldsymbol{w} + \hat{\boldsymbol{w}}^{T} Q^{-1} \hat{\boldsymbol{w}} + y^{2} D^{-1} - y D^{-1} \boldsymbol{x}^{T} \boldsymbol{w} - \boldsymbol{w}^{T} \boldsymbol{x} D^{-1} \boldsymbol{y} + \boldsymbol{w}^{T} \boldsymbol{x} D^{-1} \boldsymbol{x}^{T} \boldsymbol{w}$$

$$= \boldsymbol{w}^{T} \left( \boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1} \right) \boldsymbol{w} - \boldsymbol{w}^{T} \left( Q^{-1} \hat{\boldsymbol{w}} + \boldsymbol{x} D^{-1} \boldsymbol{y} \right) - \left( \hat{\boldsymbol{w}}^{T} Q^{-1} + y D^{-1} \boldsymbol{x}^{T} \right) \boldsymbol{w} + \hat{\boldsymbol{w}}^{T} Q^{-1} \hat{\boldsymbol{w}} + y^{2} D^{-1}$$

$$= \left( \boldsymbol{w} - \boldsymbol{\xi} \right)^{T} \left( \boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1} \right) \left( \boldsymbol{w} - \boldsymbol{\xi} \right) + \hat{\boldsymbol{w}}^{T} Q^{-1} \hat{\boldsymbol{w}} + y^{2} D^{-1} - \boldsymbol{\xi}^{T} \left( \boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1} \right) \boldsymbol{\xi}$$

$$= \left( \boldsymbol{w} - \boldsymbol{\xi} \right)^{T} R^{-1} (\boldsymbol{w} - \boldsymbol{\xi}) + \Delta$$

$$(2.7.81)$$

となり、w の項とそれ以外の項に分けることができた。ここで、

$$R \equiv \left(\boldsymbol{x}D^{-1}\boldsymbol{x}^T + Q^{-1}\right)^{-1} \tag{2.7.82}$$

$$\Delta \equiv \hat{\boldsymbol{w}}^T Q^{-1} \hat{\boldsymbol{w}} + y^2 D^{-1} - \boldsymbol{\xi}^T \left( \boldsymbol{x} D^{-1} \boldsymbol{x}^T + Q^{-1} \right) \boldsymbol{\xi}$$
(2.7.83)

$$\boldsymbol{\xi} \equiv (\boldsymbol{x}D^{-1}\boldsymbol{x}^T + Q^{-1})^{-1} (Q^{-1}\hat{\boldsymbol{w}} + \boldsymbol{x}D^{-1}y)$$
(2.7.84)

であり、 $\xi$ の転置は

$$\boldsymbol{\xi}^{T} = (Q^{-1}\hat{\boldsymbol{w}} + \boldsymbol{x}D^{-1}y)^{T} \left\{ (\boldsymbol{x}D^{-1}\boldsymbol{x}^{T} + Q^{-1})^{-1} \right\}^{T} = (\hat{\boldsymbol{w}}^{T}Q^{-1} + yD^{-1}\boldsymbol{x}^{T}) (\boldsymbol{x}D^{-1}\boldsymbol{x}^{T} + Q^{-1})^{-1}$$
(2.7.85)

であることと、

$$(\boldsymbol{w} - \boldsymbol{\xi})^{T} (\boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1}) (\boldsymbol{w} - \boldsymbol{\xi}) = \boldsymbol{w}^{T} (\boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1}) \boldsymbol{w} - \boldsymbol{w}^{T} (\boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1}) \boldsymbol{\xi}$$

$$- \boldsymbol{\xi}^{T} (\boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1}) \boldsymbol{w} + \boldsymbol{\xi}^{T} (\boldsymbol{x} D^{-1} \boldsymbol{x}^{T} + Q^{-1}) \boldsymbol{\xi}$$

$$(2.7.86)$$

を用いた。この結果を(2.7.10)式に用いると、

$$p(\boldsymbol{w}|y,z) = \frac{(2\pi)^{-\frac{K+1}{2}}|D|^{-\frac{1}{2}}|Q|^{-\frac{1}{2}}e^{-\frac{\Delta}{2}}\exp\left[-\frac{1}{2}(\boldsymbol{w}-\boldsymbol{\xi})^TR^{-1}(\boldsymbol{w}-\boldsymbol{\xi})\right]}{(2\pi)^{-\frac{K+1}{2}}|D|^{-\frac{1}{2}}|Q|^{-\frac{1}{2}}e^{-\frac{\Delta}{2}}\int\exp\left[-\frac{1}{2}(\boldsymbol{w}-\boldsymbol{\xi})^TR^{-1}(\boldsymbol{w}-\boldsymbol{\xi})\right]d\boldsymbol{w}}$$

$$= (2\pi)^{-\frac{K}{2}} |R|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\boldsymbol{w} - \boldsymbol{\xi})^T R^{-1} (\boldsymbol{w} - \boldsymbol{\xi})\right]$$
 (2.7.87)

となることから、

$$\boldsymbol{w}|y,z \sim N(\boldsymbol{\xi},R) \tag{2.7.88}$$

であることがいえる。

続いて、 $\xi$ と R を変形することで (2.7.30) 式  $\sim$ (2.7.33) 式を導く。まず  $\xi$  については、

$$\xi = (xD^{-1}x^{T} + Q^{-1})^{-1} (Q^{-1}\hat{w} + xD^{-1}y) 
= (xD^{-1}x^{T} + Q^{-1})^{-1} [(xD^{-1}x^{T} + Q^{-1})\hat{w} + xD^{-1}(y - x^{T}\hat{w})] 
= \hat{w} + (xD^{-1}x^{T} + Q^{-1})^{-1} xD^{-1}(y - x^{T}\hat{w}) 
= \hat{w} + Qx (x^{T}Qx + D)^{-1}(y - x^{T}\hat{w}) 
= \hat{w} + K (y - x^{T}\hat{w})$$
(2.7.89)

となる。ここでKは

$$K \equiv Qx \left(x^{T}Qx + D\right)^{-1} \tag{2.7.90}$$

である。上記の式変形では、恒等式  $xD^{-1}\left(x^TQx+D\right)=\left(xD^{-1}x^T+Q^{-1}\right)Qx$  に左から  $\left(xD^{-1}x^T+Q^{-1}\right)^{-1}$ を、右から  $\left(x^TQx+D\right)^{-1}$  を掛けると、

$$(xD^{-1}x^T + Q^{-1})^{-1}xD^{-1} = Qx(x^TQx + D)^{-1}$$
(2.7.91)

が得られることを用いた。最後にRについては、

$$R = (xD^{-1}x^{T} + Q^{-1})^{-1}$$

$$= (xD^{-1}x^{T} + Q^{-1})^{-1} [(xD^{-1}x^{T} + Q^{-1}) Q - xD^{-1}x^{T}Q]$$

$$= Q - (xD^{-1}x^{T} + Q^{-1})^{-1} xD^{-1}x^{T}Q$$

$$= Q - Qx (x^{T}Qx + D)^{-1} x^{T}Q$$

$$= Q - Kx^{T}Q$$
(2.7.92)

ここで再び(2.7.91) 式を用いた。(2.7.88) 式を元の表記で書くと

$$\mathbf{w}_t | \mathbf{y}_{1:t} \sim N(\boldsymbol{\xi}, R) \tag{2.7.93}$$

となる。これは  $w_t$  の時刻 t におけるフィルタ分布の期待値が  $\xi$ 、分散共分散行列が R であることを意味するので、 $w_{t|t}\equiv \xi$  および  $Q_{t|t}\equiv R$  とおく。以上の結果をまとめて元の表記で書くと、

$$\mathbf{w}_t | y_{1:t} \sim N(\mathbf{w}_{t|t}, Q_{t|t})$$
 (2.7.94)

$$\mathbf{w}_{t|t} = \mathbf{w}_{t|t-1} + \mathbf{K}_t \left( y_t - \mathbf{x}_t^T \mathbf{w}_{t|t-1} \right) \tag{2.7.95}$$

$$Q_{t|t} = Q_{t|t-1} - K_t x_t^T Q_{t|t-1}$$
(2.7.96)

$$K_{t} = Q_{t|t-1}x_{t} \left(x_{t}^{T}Q_{t|t-1}x_{t} + D_{t}\right)^{-1}$$
(2.7.97)

が得られる。ここで $K_t \equiv K$ とした。

### 2.8 診断手法 1

### 2.8.1 概要

診断手法は、前節までに述べた統計手法と異なり、過去の研究や目的変数の定義から予測式を決定し、ガイダンスの予測値を算出する手法である。診断手法が利用されるようになった背景には、数値予報モデルの精緻化や予測精度の向上がある。近年の数値予報は、ガイダンスの運用が開始された頃に比べ、統計的な補正を行わない簡易的な手法でも、実用的な精度を持つガイダンスを作成することが可能になってきている。

統計手法と比べて、診断手法は次のような特長を持つ。

観測や数値予報モデルの長期間のデータが不要。 観測密度に起因する予測精度の不均一性がない。 メリハリの効いた予測が可能。

統計手法では係数を学習するために長期間の観測と数 値予報モデルのデータが必要となるが、診断手法では 係数を学習する必要がないため、長期間のデータの蓄 積は不要 $^2$ である()。また、統計手法では、係数を 学習するために予測対象とする地点や領域で十分な数 の観測データが必要となるが、診断手法では観測デー タが少ない海上や上空でも他の地点と同程度の精度を 持つ予測値が得られる()。さらに、統計手法では過 去データ全体に対して平均的な誤差が小さくなるよう に予測式が求まる。その結果、稀な現象の予測頻度は 低くなり、また、MOS の場合には予報時間が長くなる と予測が気候値に近づいていく (第2.9節)。これに対 し、診断手法では数値予報モデルの誤差を考慮しない ため、メリハリの付いた予測が可能となる()。これ は現象の発生頻度や予報時間に応じた予測の不確実性 によらず同じ予測式を適用することで得られる特徴で ある。

診断手法には統一的な作成手法はなく、予測要素に 応じて手法を開発する必要がある。本節では、診断手 法を利用したガイダンスを例示し、診断手法を利用す る上での留意事項を述べる。

### 2.8.2 診断手法を利用したガイダンスの運用

前項で診断手法は過去の研究や目的変数の定義から 予測式を決定することを述べたが、気象庁で診断手法 を利用するガイダンスの多くは、実際にはモデルの予 測特性や観測値に合わせて予測式を調整している。こ れは診断手法の利点を生かしつつ、数値予報モデルの 系統誤差などを出来るだけ軽減し、ガイダンスの予測 精度の向上を目指すためである。利用するモデルに合 わせた調整の実際を、MSM 視程分布予想を例として 紹介する。

MSM 視程分布予想は MSM で予測された降水量、降雪量、相対湿度、風速から、光の消散率 (消散係数  $\sigma$ ) を求め、これを視程  ${
m VIS}(=3/\sigma)$  に変換する。詳細は第 4.9 節で述べるが、この消散係数の一つである雲粒による消散係数  $\sigma_c$  の予測式は、Gultepe et al. (2006) を参考に METAR<sup>3</sup> 及び SPECI<sup>4</sup> の視程と MSM の雲水量の関係を元に調整した (井藤 2011)。 Gultepe et al. (2006) は測器観測による雲水量 LWC [g m $^{-3}$ ] から、雲粒の消散係数  $\sigma_c$  による視程 VIS を以下のように見積もった。

VIS = 
$$3/\sigma_c = 0.0219 \times LWC^{-0.9603}$$
 (2.8.1)

m MSM 視程分布予想(運用開始時)は、この式を参考に、m MSM の雲水量  $m QC~[g~kg^{-1}]$  と観測値の関係から、以下を予測式とした。

VIS = 
$$3/\sigma_c = 0.333 \times QC^{-0.9}$$
 (2.8.2)

この例のように、診断手法を用いたガイダンスの多くは、利用するモデルや観測値に合わせて予測式の調整を行っている。ただし、診断手法で得られる予測式を用いて予測値の分布や特性を確認するといった簡易的な調整がほとんどであるため、調整に必要なデータは統計手法に比べて少なくて済む。

### 2.8.3 診断手法の適用例

ここでは診断手法を利用しているガイダンスを紹介 する。なお、各ガイダンスの詳細は第4章で解説する ため、ここでは概略を述べるに留める。

#### (1) 視程分布予想

詳細は第4.9節を参照。視程分布予想は地上及び海 上の水平視程を面的に予測するガイダンスで、濃霧等 の視程障害の予報に利用されている。視程の観測点は、 気温や雨の観測点に比べて数が少ないため、視程分布 予想の運用を開始するまでは視程障害に関する面的な 予測資料がなかった。しかし、数値予報モデルの改良に よって地表付近の雲水量が精度よく予測できるように なり、診断手法による予測が視程障害を予測する際の 参考資料として実用的な精度を持つことが確かめられ たため、2011 年に MSM 視程分布予想の運用を開始し た。視程分布予想は GSM, MSM,  $LFM^5$  で運用されて いる。各数値予報モデルの特性を考慮するために、光 消散係数の各要素 ( 浮遊塵  $\sigma_p$ 、 雲粒  $\sigma_c$ 、 雨粒  $\sigma_r$ 、 雪  $\sigma_s$ ) の予測式をモデルに合わせて調整している (井藤 2011, 2013; 金井ほか 2015; 後藤 2017)。ただし、領域 によって予測式を変えていない 6 ため、例えば、北日

<sup>1</sup> 後藤 尚親

<sup>&</sup>lt;sup>2</sup> 第 2.8.2 項で示すように、予測式の調整に数値予報モデルのデータを利用するガイダンスの場合は、データの蓄積とモデル更新時に作業が必要となる。ただしその場合でも統計手法に比べると、必要なデータの量や作業は少なくて済む。

<sup>3</sup> 航空気象定時観測気象報

<sup>4</sup> 航空気象特別観測気象報

 $<sup>^5</sup>$  LFM 航空悪天 GPV の一要素として作成している。

<sup>&</sup>lt;sup>6</sup> 例外として、GSM 視程分布予想ではオホーツク海とそれ 以外の領域で予測式を変えている。

本・東日本太平洋側では霧が多く予測されるが、南シ ナ海では霧の予測が少ないといった問題がある。

#### (2) 降水種別ガイダンス

詳細は第 4.3 節を参照。降水種別ガイダンスは雨、雨か雪、雪か雨、雪の 4 つの降水の種別を予測するガイダンスである。降水の種別を観測する地点は少ないが、面的な天気ガイダンスや最大降雪量ガイダンスには面的な降水種別の予測が必要であることから診断手法が用いられている。降水種別の判別には柳野 (1995) の地上気温と相対湿度で降水種別を判定する雨雪境界線を基に、2004 年から 2008 年の 5 年間の冬季の 51 の気象官署の観測データを対象に調整を行った結果を利用している。

降水種別ガイダンスは、実況の気温と相対湿度から 作成した雨雪境界線に数値予報モデルの系統誤差を補 正した格子形式気温ガイダンスと数値予報モデルの相 対湿度を入力し、上空の気温等で補正することで降水 種別を予測している。相対湿度については数値予報モ デルの予測値をそのまま利用しているため、系統誤差 が補正されていない課題がある。

### (3) 着氷指数

詳細は第 4.12 節を参照。着氷指数は航空悪天 GPV で算出している要素の一つで、航空機への着氷を予測する指数である。飛行中に発生する着氷の予測には過冷却水滴の有無やその量が重要になるが、数値予報モデルで過冷却水滴を正確に予測することは現状ではまだ難しい。また、航空機による着氷の事例数が極端に少ないことから、一般的な統計手法を利用することは難しい。そこで過去の気温と着氷頻度及び湿数と着氷頻度の調査に基づき着氷指数を開発し運用している(工藤 2008)。この指数は数値予報モデルの気温・湿数と観測された着氷頻度の関係を関数で当てはめ、2 つの関数を掛けあわせて予測式を作成している。予測式の作成にモデルの気温と湿数を用いることで、利用するモデルに適した予測式となるように調整している。

### (4) 積乱雲量・積乱雲頂高度

詳細は第 4.13 節を参照。積乱雲量・積乱雲頂高度は 航空悪天 GPV で算出している要素の一つで、パーセ ル法に基づいて、数値予報モデルの気温や相対湿度な どから積乱雲の雲量と雲頂高度を算出している。空域 予報においては陸上だけでなく海上も含めた領域での 予測が必要であることや、積乱雲の雲量や雲頂高度の 実況を面的に入手することが出来ないことから、診断 手法を用いて積乱雲の予測を行っている。

### (5) 圏界面

圏界面は航空悪天 GPV で算出している要素の一つで、高層気象観測指針にある気圧と気温を用いた第一圏界面の定義を元に、モデルの気圧と気温から算出している。圏界面の実況は高層観測地点でしか得られな

いが、航空機の運航を支援するためには面的な予測が必要となるため、診断手法により圏界面の気圧などを 面的に算出している。

### 2.8.4 利用上の留意点

診断手法は過去の研究や目的変数の定義に基づく予測手法であり、統計手法を利用する場合と比べて、長期間のデータ蓄積が不要、観測データが少ない地点や領域での予測が可能、メリハリのある予測が可能、という利点があることを述べた。一方で、診断手法には以下のような問題点もある。

数値予報モデルの系統誤差を補正できない。 予測対象地点や時刻ごとの誤差特性を反映した予 測式を作成することが難しい。

数値予報モデルの特性が大きく変わった場合に予 測精度の低下や予測特性の変化が起きる可能性が ある。

については、本節で述べたように、モデルごとに予測式を調整することである程度の系統誤差を取り除くことができる。ただし、予測式を調整しすぎるとの問題も生じる。降水種別ガイダンスで行っているように、数値予報モデルを入力とする代わりに、系統誤差が補正されたガイダンスを利用することで解決できる場合もあるが、視程分布予想における雲水量や積乱雲予測における上空の気温や相対湿度など、多くの場合はガイダンス値が得られない。このため診断手法では、同じモデル予測値を利用するとしても、可能な限り系統誤差の少ないデータを入力とすることが望ましい。

は診断手法にとってはデメリットであり、これを解決するには、多くの観測データを入手して統計手法を利用する必要がある。 について、同様のことは統計手法を用いたガイダンスでも起こりうるが、例えば第3.2節で述べるように、統計手法の場合にはモデル更新時に長期間の再予報を実施して係数を再学習することで、更新前と同程度の予測精度を得られる場合が多い。診断手法でも、係数を調整することで予測精度の低下を回避・軽減することは可能であるが、係数の調整のみでは十分な予測精度が得られない場合がある。このような場合には予測手法自体の大幅な見直しも必要になる。診断手法によるガイダンスを開発・運用するに当たっては、これらの点に留意が必要である。

### 参考文献

後藤尚親, 2017: MSM ガイダンスの特性の変化. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 56-60.

Gultepe, I, M.D.Müller, and Z.Boybeyi, 2006: A New Visibility Parameterization for Warm-Fog Applications in Numerical Weather Prediction Models. J. Appl. Meteor. Climat., 45, 1469–1480.

井藤智史, 2011: 視程分布予想の解説. 平成 23 年度数値予報研修テキスト, 気象庁予報部, 25-29.

- 井藤智史, 2013: GSM 視程分布予想の開発. 平成 25 年 度数値予報研修テキスト, 気象庁予報部, 58-62.
- 金井義文, 満満男, 工藤淳, 2015: 下層悪天予想図及び新しい狭域悪天予想図. 航空気象ノート第77号, 気象庁航空気象管理官.
- 工藤淳, 2008: 国内航空悪天 GPV. 平成 20 年度数値予報研修テキスト, 気象庁予報部, 92-97.
- 柳野健, 1995: ニューラルネットによるガイダンス. 平成7年度量的予報研修テキスト, 気象庁予報部, 54-69.

### 2.9 頻度バイアス補正<sup>1</sup>

### 2.9.1 はじめに

頻度バイアス補正は、予測頻度を実況頻度に合わせるように予測値を補正する CDF (Cumulative Distribution Function; 累積分布関数) マッチング<sup>2</sup>の一手法である。気象庁では予測頻度の実況頻度に対するバイアスを補正する手法という意味で「頻度バイアス補正」という名称で呼んでいる。海外の気象機関でもバイアス補正または quantile mapping (Maraun 2013) と呼ばれている。

気象庁のガイダンスには、1996 年 3 月、降水量ガイダンスにカルマンフィルタ (KF) を導入した際に合わせて導入された (國次・藤田 1996; 藤田 1996)。KF の導入と同時に導入されたが、KF の手法に問題があって導入されたわけではなく、線形重回帰 (MLR) を手法として使っていた際にも課題であった大雨の予測頻度が少ない点を修正するために導入されている。その後、頻度バイアス補正は降水量ガイダンス以外にも適用され、2018 年現在、風ガイダンス、視程ガイダンス、降雪量地点ガイダンス、雲ガイダンスにも使われている。これらのガイダンスではいずれも KF またはニューラルネットワーク (NN) の後処理として頻度バイアス補正が行われている。

ここでは、頻度バイアス補正がガイダンスの後処理 として何故必要かを解説し、その後、気象庁での頻度 バイアス補正の具体的な手法及びその逐次更新の手法 を解説する。最後にまとめとして気象庁以外での利用 について簡単に述べる。

### 2.9.2 頻度バイアス補正の必要性

例えば降水量の実況頻度は弱い降水に偏っており、強い降水の頻度は極端に少ない。このような場合に、予測の誤差を減らそうとすると、頻度の多い弱雨のデータを重視した予測となる。図 2.9.1 に予測式と、各説明変数における誤差の頻度分布のイメージを示した。弱雨の頻度が多いことにより、 弱雨を予測する所で誤差の頻度分布のピークが大きくなること、 誤差分布は正規分布ではなく、弱雨側にピークがあり、この誤差分布に引きずられて予測式は弱雨側に設定される 3。また、気象予測の場合は低気圧等の擾乱の位置ずれなどで誤差が生じ、同じ説明変数でも実況は大雨と弱雨となる場合があるため、予測は中間的な値となって、大雨が予測されにくくなることも原因の一つと考えられ

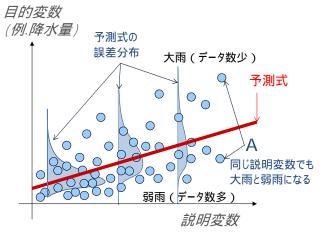


図 2.9.1 頻度が偏った目的変数を予測した場合の予測式と、その誤差の頻度分布のイメージ。説明変数を 1 個と仮定した場合の例。青の丸は個々の説明変数と目的変数のデータ、青のハッチの分布は予測式の各説明変数値での誤差の頻度分布(横軸側に頻度分布)イメージを示す。

### る(図2.9.1のA)。

図 2.9.2 に、風ガイダンスの場合における、頻度バイ アス補正前と後の予測頻度と実況頻度を示す。補正前 は、頻度の多い弱風の予測頻度が実況頻度より多く、頻 度の少ない強風の予測頻度は実況頻度より少なくなっ ていることがわかる。強風の予測が実況より頻度が少 ないということは、見逃しが多いことを意味するため、 防災気象情報の作成を支援するガイダンスとしては問 題となる。これに対し頻度バイアス補正後は、実況と 予測の頻度が概ね一致し、強風の予測事例が増えてい る。図 2.9.3 には同じ風ガイダンスの予測と実況の散 布図を示した。頻度バイアス補正前には弱めに予測さ れていること、頻度バイアス補正後は風速を強くする ように補正が行われていることがわかる。風速を強く することによって、強風の捕捉率が上昇しており(赤 のハッチ部分)防災気象情報にとってはより有効な予 測となっている。一方、頻度バイアス補正は実況の頻 度に合わせるように風を強めに補正するため、空振り も増えている(青のハッチ部分)。

# 2.9.3 頻度バイアス補正の手法

頻度バイアス補正の手法は単純で、予測の頻度が実 況の頻度と合うように予測値を補正するだけである。 図 2.9.4 は、風ガイダンスを例とした、CDF マッチン グとしての頻度バイアス補正のイメージ図である。補 正前のガイダンスは弱風の頻度が実況より多く、強風 の頻度が少ないため、累積した頻度分布である CDF は実況の CDF より左側にある。その予測の CDF を 実況の CDF に補正する手法が頻度バイアス補正であ る。つまり、図中に示した太矢印のようにガイダンス (補正前)の出力値を上方補正し、実況の CDF に合わ せる。図 2.9.5 に風ガイダンス(定時風)等で実際に 使われている頻度バイアス補正の方法を示す。頻度バ イアス補正では、観測値と予測値に数個の閾値を設定

<sup>1</sup> 高田 伸一

<sup>&</sup>lt;sup>2</sup> 2 つの母集団があり、片方の頻度分布をもう片方の頻度分布に合わせるように値を補正する手法。以下で説明する手法 の他に、変換係数で合わせる手法 (幾田 2017) がある。

<sup>&</sup>lt;sup>3</sup> KF は観測ノイズが正規分布に従うことを仮定しており、この仮定が成り立っていない影響が出ているとも言える。今後は目的変数を変換したり、正規分布を仮定しない他の手法の導入を検討する必要がある。

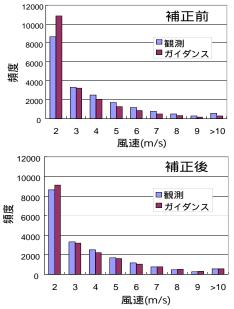


図 2.9.2 風ガイダンス (定時風)の予測頻度と実況頻度。上が頻度バイアス補正前、下が頻度バイアス補正後。1997年 12 月から 1998 年 2 月の全アメダス地点の合計。

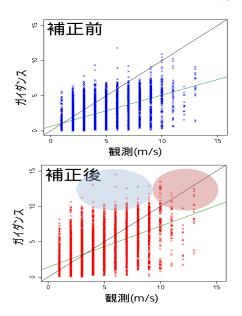


図 2.9.3 風ガイダンス (定時風)の実況と予測の散布図。上が頻度バイアス補正前、下が頻度バイアス補正後。図 2.9.2 と同じ 1997 年 12 月から 1998 年 2 月の全アメダス地点のデータを元に作成。下の赤のハッチ部分は、頻度バイアス補正によって 10 m/s 以上の風の捕捉率が増えたことを示す。青のハッチ部分は逆に 10 m/s 以上の予測の空振りが増えたことを示す。緑線は観測とガイダンスの回帰直線を示す。

し、各カテゴリでの頻度が等しくなるように補正を行う。風ガイダンスの場合は、観測側に4つの閾値(2.5, 5.5, 9.5, 13.0 m/s)を設定し、カテゴリを4つに分ける。その4つのカテゴリ毎に観測の頻度を数え、それと同じ頻度となる予測側の閾値を計算する。この例の場合は4つの観測側の閾値に対応し、4つの予測側の閾値(1.9, 3.8, 7.1, 9.8 m/s)を計算で求めている。こ

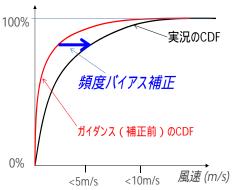


図 2.9.4 風ガイダンスを例とした、実況の CDF (累積分布 関数)、補正前のガイダンスの CDF 及び頻度バイアス補 正のイメージ。

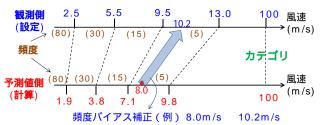


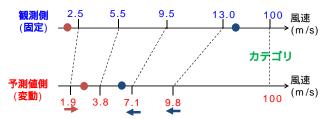
図 2.9.5 風ガイダンス等で実際に使われている頻度バイアス補正の手法。観測側の閾値(青色)を設定して各カテゴリ(緑色)に分け、その各カテゴリの頻度と同じになるように予測値側の閾値(赤色)を計算する。頻度バイアス補正は太矢印(薄青色)のように行う。

の両方の閾値に沿って例えば予測値が  $7.1~\mathrm{m/s}$  であれば  $9.5~\mathrm{m/s}$  に補正する。途中の  $8.0~\mathrm{m/s}$  であれば線形補間で  $10.2~\mathrm{m/s}$  と補正値を決定する。なお、この例では一番右側に補正を行う上限値として  $100~\mathrm{m/s}$  が設定してある。この上限値以上では補正を行わず、ガイダンス(補正後) = ガイダンス(補正前)とする。

頻度バイアス補正を行う場合に留意すべき点として、各カテゴリに含まれるサンプル数が少なくなりすぎないように閾値を設定することが挙げられる。サンプル数が少ない状態で CDF マッチングすると、過大な予測がされる可能性がある。図 2.9.4 の 100%近いところ(図の上部)では補正量が大きく、またわずかな頻度の違いで CDF の線が変わる所でもある。強風、大雨等の予測に大きな影響を与えることから、無理な補正とならないよう、サンプル数に十分気をつけて観測側の閾値を設定する必要がある。

### 2.9.4 頻度バイアス補正の逐次更新

逐次学習型ガイダンスの場合には、前項で説明した 予測側の閾値を逐次更新する。KF や NN で予測式の 係数が更新する際に、頻度バイアス補正の予測側の閾値も同時に更新する。閾値更新のイメージ図を図 2.9.6 に示す。頻度バイアス補正では、予測値と観測値のカ テゴリが異なった場合に予測側の閾値を更新する。観 測値の方が低いカテゴリに入った場合は予測側の閾値 を上げるように、観測値の方が高いカテゴリに入った



例1)観測が、予測がのカテゴリーの場合予測閾値1.9m/sを上げる 例2)観測が、予測がのカテゴリーの場合予測閾値7.1,9.8m/sを下げる

図 2.9.6 風ガイダンスを例とした、頻度バイアス補正の逐次的な更新方法。丸(赤、紺)は新しい予測と観測の事例を示し、矢印(赤、紺)はその時の予測値側閾値の修正の方向を示している。その他は図 2.9.5 に同じ。

場合は予測側の閾値を下げる。この閾値をどの程度変化させるかは、ガイダンスの予測精度を調べながら調整して決めている。図の紺色の例のように、カテゴリが2つ以上ずれた場合には、その途中の閾値を全て更新する。また、あるカテゴリの予測値側の閾値を全で化させる場合に、その上下の閾値を越えないように調整を行う。閾値を下げる場合と上げる場合で、閾値を変化させる割合は同じ値が用いられることが多いが、降水量ガイダンスでは同じ割合とすると大雨の予測頻度が過剰になるため、予測値の閾値を上げる割合の方が多い設定としている(白山 2017)。

# 2.9.5 アンサンブル平均での頻度バイアス補正の 利用

上記では、実況の頻度に偏りがある場合の対策とし て、頻度バイアス補正を利用することを説明した。一 方、アンサンブル予報のメンバーの予測を平均すると、 同じように強雨や強風の頻度が減る。なぜなら、例え ば頻度の少ない強雨域は通常狭く予測されるため、強 雨域が異なった予測を平均すると、強雨と弱雨の平均 値となり強雨のピークが減るからである。また、メン バー間の強雨の時間的なずれによっても強雨のピーク が減る。この平均によって降水量の誤差は減るかもし れないが、全体的に鈍った予測となり、強雨の予測が 出にくくなる。このようなアンサンブル平均の問題点 を補正するためにも、頻度バイアス補正が使われてい る。例えば、第 1.4 節で説明した米国気象局の NBM (National Blend of Models) の降水量予測においては、 多くの予測が結合された後に、頻度バイアス補正(論 文中では quantile mapping ) によって大雨が予測され るように補正されている (Hamill et al. 2017)。また、 第5.1 節で説明する、開発中の LFM 降水量ガイダンス でも、初期値アンサンブルを行った後に頻度バイアス 補正を行う予定である。

### 2.9.6 海外での状況

米国気象局では、風ガイダンスにおいて強風の予測が 出にくいことから、風速のバイアス補正として inflation という方法を 1975 年から利用していた (Schwartz and Carter 1982)。ただし、この手法は頻度バイアス補正 ではなく、 ${
m MLR}$  で予測式を作成した際の重相関係数 R (第 2.4 節 ) を使って以下のように風速を強めに補正していた。

$$S_{\text{after}} = S_{\text{mean}} + \frac{S_{\text{before}} - S_{\text{mean}}}{R}$$
 (2.9.1)

ここで  $S_{\rm after}$  は補正後の風速、 $S_{\rm before}$  は補正前の風速、 $S_{\rm mean}$  は重回帰分析を行った期間の平均風速であり、 $S_{\rm before}$  が  $S_{\rm mean}$  より小さい場合は補正を行わない。この手法は重相関係数 R が大きい(予測精度が高い)場合は補正を小さく、R が小さい(予測精度が低い)場合は補正を大きくする手法である。図 2.9.1 で予測式の誤差分布のイメージを示したが、予測精度が低いと(誤差が大きいと)、中間的な予測となりやすく強風がより予測されにくくなるため、このような予測精度を考慮した補正方法が有効だったと考えられる。なお、2018年現在はこの手法でなく、予測と実況を比較し、その差を逐次的に計算した量を inflation として風速を増大する手法が使われている (Glahn et al. 2014)。

フランス気象局では、WMO の Technical Progress Report によると風速、降水量ガイダンスにおいてモデル予測値を校正 (Calibration) していると記述されているなど頻度バイアス補正を利用しているようであるが、具体的な手法は不明である。

# 参考文献

藤田司, 1996: ガイダンスの検証. 平成 8 年度数値予報 研修テキスト, 気象庁予報部, 34-40.

Glahn, B., D. Rudack, and B. Veenhuis, 2014: On bias correcting MOS wind speed forecasts. *MDL Office Note*, **14-1**, 1–19.

Hamill, T. M., E. Engle, D. Myrick, M. Peroutka, C. Finan, and M. Scheuerer, 2017: The U.S. National Blend of Models for Statistical Postprocessing of Probability of Precipitation and Deterministic Precipitation Amount. *Mon. Wea. Rev.*, 145, 3441–3463.

幾田泰酵, 2017: 局地数値予報システムにおける新規観 測データの利用開始及び同化手法の高度化. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 82-85.

國次雅司, 藤田司, 1996: カルマンフィルターによるガイダンス. 平成7年度量的予報研修テキスト, 予報部 予報課, 44-53.

Maraun, D., 2013: Bias correction, quantile mapping, and downscaling:Revisiting the inflation issue. *J. Climate*, **26**, 2137–2143.

Schwartz, B. E. and G. M. Carter, 1982: An evaluation of a modified speed enhancement technique for objective surface wind forecasting. *TDL Office Note*, **82-1**, 10pp.

白山洋平, 2017: 降水ガイダンスの改良. 平成 29 年度 数値予報研修テキスト, 気象庁予報部, 86-93.

#### 2.10 その他の統計手法 1

本節では第 2.4 節  $\sim$  第 2.9 節で述べた手法以外の統計手法について、その一部を簡単に紹介する。本節で述べる統計手法はいずれも 2018 年現在の気象庁のガイダンスには利用されていないが、海外気象機関のガイダンスでは利用されている手法もあり、今後新たな統計手法を導入するに当たって候補となりうる手法である。

### 2.10.1 決定木

決定木はフローチャートのように設定した説明変数の閾値によって、事象(晴、曇、雨など)をいくつかのクラスに分類する手法である。例えば図 2.10.1 の左図のように、2 つの説明変数  $x_1$ と  $x_2$  を用いて現象あり( )となし( x )を分けることを考える。このとき閾値  $\theta_1 \sim \theta_3$  を用いると、右図のような判別により現象のあり・なしを分類できる。図からわかるように、決定木では分類するクラスが線形分離不可能な場合でも判別可能である。予報作業で用いられるフローチャートは過去の事例を元に開発者が分類方法や閾値を決定する場合が多い $^2$  が、決定木ではそれを統計手法によって行う。決定木は主に分類問題に利用されるが、連続値に対する回帰を行うこともできる。回帰の場合には、分類された結果に属する学習データの平均値を出力値とする。このため出力される値は離散的な値となる。

決定木で分類方法を決定する代表的なアルゴリズム に CART(Breiman et al. 1984) や C4.5(Quinlan 1996) がある。これらは以下の手順で分類を行う。

- 1. 全ての学習データに対して設定した誤差関数が小さくなるようにデータを2分割する
- 2. 分割されたデータに対して誤差関数が小さくなる ようにデータをさらに 2 分割する
- 3. 設定した停止条件(分岐の数や階層の数など)が 満たされるまで手順2を繰り返す
- 4. 過学習を防ぐため、設定したパラメータに基づい て分岐を剪定する

決定木では最上位の分岐をルート、各分岐をノード、最下位の判別結果をターミナルノードと呼び、各ノードに含まれる学習データの数を例題と呼ぶ。誤差関数としては負のエントロピーやジニ係数 3 などが用いられ、2 つに分けたデータの不純度が低くなるように(片方のクラスのデータが多く含まれるように)データが分割される。手順4の分岐の剪定では、ターミナルノードの数や階層の数などに比例したペナルティー項を誤

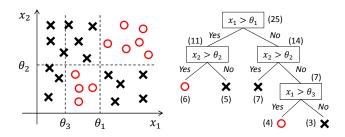


図 2.10.1 決定木での分類の例。 は現象あり、×は現象なしを表す。括弧内の数字は各ノードに含まれる例題(学習データ)の数。

差関数に加えることで、分類精度への影響の小さい分 岐を剪定(削除)する。

決定木のメリットとして、判別の過程が理解しやすいことと、線形分離不可能な場合でも判別可能であることが挙げられる。デメリットとしては、判別結果の分散が大きいことが挙げられる。このことは、学習データが少し変わっただけで分割方法が変わってしまうことを意味する。特に上位の階層で分割が変わるとそれ以下の階層にも影響が伝播するため、決定木の構造が大きく変化してしまう。これは予測精度の低下を招く要因となる。

### 2.10.2 集団学習

線形重回帰やロジスティック回帰、ニューラルネットワークなどの統計手法は、目的変数と説明変数の関係を学習する機械(学習器)といえる。これらの学習器に説明変数を与えると予測値が得られる。ガイダンスの予測精度を向上させようとした場合、表現能力の高い学習器を用いることもできるが、そのような学習器は一般に計算コストが高く、またチューニングのためのパラメータも多くなる。そこで、表現能力は低いが計算コストやパラメータの少ない学習器を多く作成し、それらを組み合わせることで予測精度を向上させようという手法が集団学習である。集団学習はそれ自体は学習器ではないが、回帰手法や分類手法の精度を大きく向上させる可能性のある手法である。代表的な集団学習にバギング、ブースティング、ランダムフォレストがある。

## (1) バギング

バギング  $^4$ (Breiman 1994) では以下の手順で予測を行う(図 2.10.2)。

- 1. N 個の学習データからブートストラップ法 (第 2.3.12 項) でサンプル数 N の学習データを R 組 生成する
- 2. R 組の学習データに元の学習データを加えた R+1 組の学習データに対して、ある一種類の統計手法 (線形重回帰や決定木など)を用いて R+1 個の

<sup>1</sup> 工藤 淳

<sup>&</sup>lt;sup>2</sup> このためフローチャートは "subjective decision tree" (主観的決定木)とも呼ばれる。

 $<sup>^3</sup>$  あるノードに含まれる例題を 2 つに分割する割合を p としたとき、負のエントロピーは  $p\ln p + (1-p)\ln(1-p)$ 、ジニ係数は 2p(1-p) となる。

<sup>&</sup>lt;sup>4</sup> Bagging, Bootstrap AGGregatING の略。

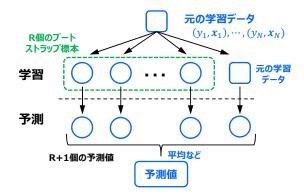


図 2.10.2 バギングによる予測の模式図

予測式を作成する

3. R+1個の予測式を用いて予測を行い、予測対象が連続値(回帰)の場合には各予測の平均を、クラス分類の場合には最も多く予測されたクラスを予測値とする

バギングは手法がシンプルであるため導入が容易であり、また各学習データに対する予測式作成を並列に実行できるため比較的短時間で学習することができる。 学習器として回帰を利用した場合には、複数の予測式で予測する代わりに係数の平均を用いて予測を行えばよいため運用も容易である。

バギングでは、学習に使用するデータはランダムに選ぶことになるため、各プートストラップ標本ごとに約 1/3 のデータは学習に使用されないことになる 5 。このことを利用すれば、交差検証と同様の効果が得られる。具体的には、学習に使用されなかったデータを検証データとして各プートストラップ標本について予測精度を求め、それを全標本で平均する。これを Out-Of-Bag (OOB) 推定といい、OOB 推定に用いられる標本を OOB 標本と呼ぶ。

バギングは学習データが十分には得られない場合ほど予測精度改善の効果が大きいと考えられる。第 2.4.5 項で述べたように、学習データが少ない場合には係数の分散が大きく(推定の精度が低く)なり、そのことが予測精度を低下させる要因となっている。バギングでは精度の低い予測を平均することで係数の推定精度を高めることができる。

### (2) ブースティング

ブースティング (Kearns and Valiant 1994) は精度の低い学習器 (弱学習器)を逐次的に組み合わせることで、精度の高い予測を行う手法である。ここでは代表的な 2 クラス分類に対するブースティングの手法である AdaBoost (Freund and Schapire 1997) について述べるが、ほかにも多クラス分類に対する AdaBoost の

拡張である AdaBoost.M1, AdaBoost.M2 (Freund and Schapire 1997), AdaBoost.MH (Schapire and Singer 1999) や、ロジスティック回帰に適用した Logit Boost (Friedman et al. 2000) など様々な拡張手法が提案されている。

バギングでは学習データから均一な割合でサンプリングすることで新たな学習データを生成したが、AdaBoostでは判別できなかったデータほど抽出されやすいように重みを付けることで、誤差が大きいデータを重点的に学習するようにする。ただし実際には重みに応じてサンプリングするのではなく、弱学習器を重みに応じて選択することでサンプリングに代えている。

今、N 個の学習データ  $(y_1,x_1)$ ,  $\cdots$ ,  $(y_N,x_N)$  と M 個の弱学習器  $f_m$   $(m=1,\cdots,M)$  があるとする。各 x は K 次元の説明変数ベクトル、 $y_n$  は目的変数で-1 または +1 である。 $f_m$  としては、例えば k 番目の説明変数を 1 つだけ用いた、1 階層しかない決定木(閾値  $\theta_m$ )などが用いられる。

$$f_m(x_{nk}) = \operatorname{sgn}(x_{nk} - \theta_m) \tag{2.10.1}$$

このような弱学習器の集合  $\{f_1,\cdots,f_M\}$  を  $\mathcal F$  とする。 最終的に求めたいものは、弱学習器を重み  $\mathcal B$  で線形結 合した  $F(x)=\sum_{m=1}^M\beta_mf_m(x)$  である。ここでは Fを判別関数と呼ぶことにする。AdaBoost による繰り 返し学習のステップ数を  $r=1,\ \cdots,\ R$  とし、r 回目の 学習ステップでの判別関数を  $F_r$ 、重みを  $w_{n,r}$  と書く。 AdaBoost では以下の手順で判別関数を求める。

- 1. 重みの初期値  $w_{n,0}$  は全て 1/N、判別関数の初期値は  $F_0(x)=0$  とする。
- 2. 学習ステップ  $r=1, \cdots, R$  について以下の手順  $3\sim 4$  を繰り返す。
- $3.~\mathcal{F}$  の中から以下の誤り率  $\varepsilon_r$  が最も小さい弱学習器を選択し、それを  $g_r$  とする。

$$\varepsilon_r(f) = \sum_{n=1}^{N} w_{n,r} I[f(\boldsymbol{x}_n) \neq y_n] \qquad (2.10.2)$$

$$g_r(\boldsymbol{x}_n) = \operatorname{argmin}_{f \in \mathcal{F}} \varepsilon_r(f)$$
 (2.10.3)

ここで関数 I[] は [] 内の条件が真ならば 1、偽ならば 0 を返す関数である。

4. 次式により重みと判別関数を更新する。

$$w_{n,r+1} = \frac{\exp\left[-F_r(\mathbf{x}_n)y_n\right]}{Z_{r+1}}$$
 (2.10.4)

$$F_r(\boldsymbol{x}_n) = F_{r-1}(\boldsymbol{x}_n) + \alpha_r g_r(\boldsymbol{x}_n) \qquad (2.10.5)$$

$$Z_{r+1} = \sum_{n=1}^{N} \exp\left[-F_r(\boldsymbol{x}_n)y_n\right]$$
 (2.10.6)

$$\alpha_r = \frac{1}{2} \ln \frac{1 - \varepsilon_r}{\varepsilon_r} \tag{2.10.7}$$

5. 判別関数を  $F = F_R$  とし、F(x) または  $\operatorname{sgn}(F(x))$  を出力する。

 $<sup>^{\</sup>overline{5}}$  N 回のサンプリングで、あるデータが一度も選ばれない確率は  $(1-1/N)^N$  であり、N が大きい場合は  $\lim_{N o\infty}(1-1/N)^N=e^{-1}\simeq 0.368$ 

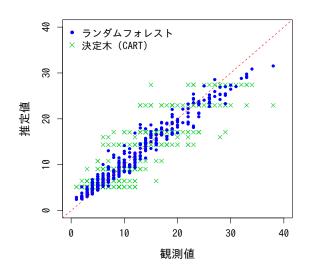


図 2.10.3 ランダムフォレストと CART を用いた決定木によるオゾン濃度の推定。横軸は統計ツール R のパッケージ "earth"に含まれている、ロサンゼルスで観測されたオゾン濃度の日最大値、縦軸は同データセットに含まれているロサンゼルスの地上の風や湿度などの観測値を説明変数として回帰を行った場合の推定値。

ブースティングは重みを逐次的に学習するため、並列 実行可能なバギングと比べて計算に時間が掛かるとい うデメリットはあるが、バギングよりも予測精度が高 くなる傾向がある。

#### (3) ランダムフォレスト

ランダムフォレスト (Breiman 2001) は、決定木の集団学習法の一種で、集団学習の方法としてはバギングを採用している。ランダムフォレストではブートストラップ法でサンプリングした複数個の学習データに対して決定木で学習を行うが、決定木の学習とは以下の点が異なる。

- 学習に利用する説明変数は、全ての説明変数(K個)を使うのではなく、ランダムに抽出された $\widetilde{K}$ 個だけ利用する。分類の場合には $\widetilde{K}=K/3$ 、回帰の場合には $\widetilde{K}=\sqrt{K}$ が推奨されている。
- ターミナルノードの例題の数は、分類の場合は 1 個、回帰の場合は 5 個が推奨されている。
- ◆ 決定木では過学習を防ぐために剪定を行うが、ランダムフォレストでは行わない。

ランダムフォレストの予測はバギングと同様に、分類 の場合には各クラスに分類された例題の割合を、回帰 の場合には平均値を用いる。

図 2.10.3 は、統計ツール R のパッケージ "earth" に含まれている、ロサンゼルスで観測されたオゾン濃度の日最大値を、地上の風や湿度などの観測値を説明変数としてランダムフォレストと CART による決定木で回帰を行った結果である。回帰の場合、ランダムフォレストと決定木はいずれも離散的な値をとるが、ラン

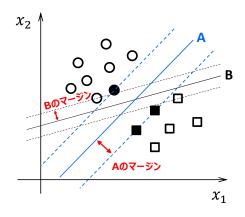


図 2.10.4 サポートベクターマシンでのマージン最大化の模式図。 $x_1, x_2$  は説明変数、 と は分類したい 2 つのクラス、 と はサポートベクターを表す。

ダムフォレストでは複数の決定木を用いて回帰を行う ため、連続値に近い推定値が得られている。またラン ダムフォレストの方がばらつきが小さく推定の精度が 高い。

ランダムフォレストは決定木のシンプルさを保持しつつ、予測精度が低い問題を解決している。また OOB 推定を用いることで、学習データだけを用いて予測データに対する精度を推定できるほか、説明変数の重要度 (Gregorutti et al. 2017) を評価することもできる。その方法は以下のとおりである。まず OOB 標本を用いて予測誤差を評価する。各 OOB 標本の誤差の平均値を  $E_{OOB}$  とする。続いて、OOB 標本の中の k 番目の説明変数をランダムに入れ替えて同様に予測誤差を評価する。このときの各 OOB 標本での予測誤差の平均値を  $E_{OOB}^{(k)}$  とする。k 番目の説明変数の重要度  $J_k$  は、

$$J_k = E_{OOB}^{(k)} - E_{OOB} (2.10.8)$$

と定義される。入れ替えた説明変数の重要度が高いほど説明変数を入れ替えたことによる誤差が大きくなるため  $J_k$  が大きくなる。

# 2.10.3 サポートベクターマシン

サポートベクターマシン (Vapnik and Lerner 1963) は、クラス分類や回帰、パターン認識に用いられる統計手法の一つであり、優れた判別能力を持つことが知られている。サポートベクターマシンの特徴として、マージンを最大化するようにクラス分類を行うことと、カーネル法 (Boser et al. 1992) を用いることで線形分離不可能な場合でも適用できることが挙げられる。

サポートベクターマシンでクラス分類を行う例を図2.10.4 に示す。図の と が分類したい2つのクラスで、 $x_1, x_2$  は説明変数である。この場合2つのクラスを直線で分離する方法は、例えば図中にAとBで示した直線など無数に存在するのだが、サポートベクターマシンでは直線に最も近い各クラスの点との距離(これをマージンと呼ぶ)が最大になる直線を選択する。

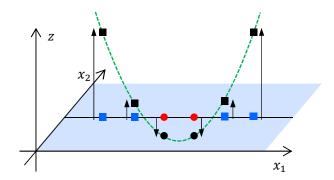


図 2.10.5 線形分離不可能なデータを高次元に射影すること で線形分離可能になる例

図の例では青線で示した A の直線が選ばれることにな る。このようにすることで2つのクラスが最も明確に 分離されるとともに、直線 A を決定しているのは図中 にと で示した 3 点のみであり、元のデータと比 べて非常に少ない数のデータでパラメータが決定され ることになる。 と のように直線を決定するデー タをサポートベクターと呼ぶ。

この例では一本の直線で誤りなく分離できているが、 実際の問題では線形分離できるとみなせるような場合 の一部は 側に、 であったとしても、 側に含まれる場合が多いだろう。このような場合には、 多少の誤りを許してマージンを決定することになる。 これをソフトマージンという。

ソフトマージンを用いたとしても、そもそも線形分 離不可能な場合にはクラス分けすることはできない。 そこでサポートベクターマシンではカーネル法が利用 される。カーネル法では元のデータを高次元空間に射 影することで線形分離不可能なデータでも線形分離可 能にすることができる。図 2.10.5 に簡単な例を示す。 この図で $x_1-x_2$ 平面上にある青四角と赤丸は一本の直 線で分けることができない。すなわち線形分離不可能 である。カーネル法ではこれを z 方向を含むような元 の空間よりも高次元の空間に射影する。射影した空間 上では元の青四角と赤丸は黒四角と黒丸のようになり、 線形分離可能になる。ただし高次元に射影しただけで は計算量が増えてしまうため、射影したデータをその まま利用するのではなく、射影したデータの内積をカー ネル関数に置き換えるという処理を行う。

### 2.10.4 拡張カルマンフィルタ

第2.7節で述べたように、カルマンフィルタは線形・ ガウス状態空間モデルに基づく手法であり、システム モデルや観測モデルが非線形である場合には適用する ことができない。そこで、非線形項を1次の項までで テーラー展開して線形化することで非線形関係も扱う ことができるようにする。ここではガイダンスのカル マンフィルタにおいて、観測モデルが非線形である以 下の状態空間モデルを考える。

$$\boldsymbol{w}_t = \boldsymbol{w}_{t-1} + \boldsymbol{u}_t \tag{2.10.9}$$

$$y_t = f_t(\boldsymbol{w}_t) + v_t \tag{2.10.10}$$

変数や添字の定義は第2.7節を参照していただきたい。 ガイダンスで予測を行う時点での $w_t$ の最適な推定量は 一期先予測の係数  $oldsymbol{w}_{t|t-1}$  であるから、 $f_t(oldsymbol{w}_t)$  を  $oldsymbol{w}_{t|t-1}$ のまわりでテーラー展開する。

$$f_{t}(\boldsymbol{w}_{t}) \approx c_{t|t-1} + \boldsymbol{f}_{t}^{\prime T}(\boldsymbol{w}_{t|t-1}) \ \boldsymbol{w}_{t}$$

$$c_{t|t-1} \equiv f_{t}(\boldsymbol{w}_{t|t-1}) - \boldsymbol{f}_{t}^{\prime T}(\boldsymbol{w}_{t|t-1}) \ \boldsymbol{w}_{t|t-1}$$
(2.10.11)

(2.10.12)

ここで  $f'_t$  は  $f_t$  の 1 階微分で  $w_t$  と同じ K 次元ベクトル、  $c_{t|t-1}$  は非確率変数である。上式を用いると、(2.10.10)式は、

$$y_t - c_{t|t-1} = \mathbf{f}_t^{T}(\mathbf{w}_{t|t-1}) \ \mathbf{w}_t + v_t$$
 (2.10.13)

となり、ガイダンスのカルマンフィルタの観測方程式 (2.7.4) において、 $y_t o y_t - c_{t|t-1}$ 、 $oldsymbol{x}_t o oldsymbol{f}_t'(oldsymbol{w}_{t|t-1})$ とした線形モデルとみなすことができる。この場合、 ガイダンスのカルマンフィルタとの違いはフィルタの 式のみである。

$$\nu_t = y_t - f_t(\mathbf{w}_{t|t-1}) \tag{2.10.14}$$

$$Q_{t|t} = Q_{t|t-1} - \mathbf{K}_t \mathbf{f}_t^{\prime T} (\mathbf{w}_{t|t-1}) Q_{t|t-1}$$
 (2.10.15)

$$Q_{t|t} = Q_{t|t-1} - \mathbf{K}_t \mathbf{f}_t^{\prime T}(\mathbf{w}_{t|t-1}) Q_{t|t-1}$$
 (2.10.15)  
$$\mathbf{K}_t = \frac{Q_{t|t-1} \mathbf{f}_t^{\prime}(\mathbf{w}_{t|t-1})}{\mathbf{f}_t^{\prime T}(\mathbf{w}_{t|t-1}) Q_{t|t-1} \mathbf{f}_t^{\prime}(\mathbf{w}_{t|t-1}) + D_t}$$
 (2.10.16)

これらの式を用いることで、観測モデルが非線形関数 で表される場合でも、カルマンフィルタと同様に係数 を更新することができる。カルマンフィルタの場合に は係数の初期値は乱数や0などの適当な値で問題なかっ たが、ここでは $w_t$ が $w_{t|t-1}$ に近いと仮定して近似し ているため、係数の初期値は何らかの方法 (y = f(w)を用いた非線形回帰など)で見積もった真値に近い値 に設定する必要がある。

#### 参考文献

Boser, B. E., I. M. Guyon, and V. N. Vapnik, 1992: A training algorithm for optimal margin classifiers. Proceedings of the 5th Annual Workshop on Computational Learning Theory, 144–152.

Breiman, L., 1994: Bagging predictors. Technical Report No.421, Department of Statistics, University of California, 1–19.

Breiman, L., 2001: Random forests. Machine Learning, **45(1)**, 5–32.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1984: Classification and Regression Trees. CRC Press, 368 pp.

- Freund, Y. and R. E. Schapire, 1997: A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, **55(1)**, 119–139.
- Friedman, J., T. Hastie, and R. Tibshirani, 2000: Additive logistic regression: A statistical view of boosting. *The Annals of Statistics*, **28(2)**, 337–407.
- Gregorutti, B., B. Michel, and P. Saint-Pierre, 2017: Correlation and variable importance in random forests. *Statistics and Computing*, **27(3)**, 659–678.
- Kearns, M. and L. G. Valiant, 1994: Cryptographic limitations on learning Boolean formulae and finite automata. *Journal of the Association for Computing Machinery*, **41(1)**, 67–95.
- Quinlan, J. R., 1996: Bagging, boosting, and C4.5. Proceedings of the 13th National Conference on Artificial Intelligence, 725–730.
- Schapire, R. E. and Y. Singer, 1999: Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, **37(3)**, 297–336.
- Vapnik, V. and A. Lerner, 1963: Pattern recognition using generalized portrait method. Automation and Remote Control, 24(6), 774–780.

# 第3章 ガイダンスの開発と運用

# 3.1 はじめに<sup>1</sup>

ガイダンスの予測精度や特性は入力となる数値予報 モデルや観測データの特性に強く依存する。ガイダン スの予測手法が入力データの特性に合っていない場合 には、十分な予測精度が得られなかったり、異常な値を 予測するなど予測が不安定になったりする。また、数値 予報モデルの更新や観測地点の移設などで入力データ の特性が変化した場合には、ガイダンスの予測精度が 低下することがある。このため、ガイダンスの開発や 運用に当たっては、入力データの特性やその変化に留 意し、適切な手法や説明変数を用いるとともに、デー タの特性が変化した場合には、係数を再作成するなど の処置を検討する必要がある。

本章ではガイダンスの開発と運用に当たって検討すべき事項を述べる。まず第 3.2 節ではガイダンスの設計時に留意すべき事項を述べ、続いて第 3.3 節と第 3.4 節では、ガイダンスの運用において重要な数値予報モデルの更新と観測所の移設時の対応について述べる。

### 3.2 ガイダンス設計時の検討事項

本節では、ガイダンスを新規に開発する場合や手法を大きく変更する場合など、ガイダンスの設計段階で 留意すべき事項を列挙する。

### (1) 手法の選択

ガイダンスに用いる統計手法や学習方法、説明変数、層別化などは、入力データの特性に合ったものを選択しなければならない。そのために、まずは目的変数と説明変数の特徴を把握することが大切である。観測データが連続値であるか、2 値データであるか、3 クラスデータであるかなどによって用いるべき手法は異なる。また、目的変数と説明変数の関係が線形であるか、非線形であるか、特定の分布に従うかや、予測誤差が正規分布に従うか、他の分布に従うかなどによっても用いるべき手法は異なる。第2 章で述べた各種手法の理論や特徴、利用上の注意点を参考に、適切な手法を選択する必要がある。

#### (2) 保守性

ガイダンスでは予測精度の観点から MOS 方式が採用されることが多いが、第 2.2.2 項で述べたように PPM方式にも様々なメリットがある。また、数値予報モデルの更新による特性変化の影響を軽減するために逐次学習方式は有効であるが、第 2.2.3 項で述べたように、一括学習方式は逐次学習方式と比べて予測結果がわかりやすいことや、開発および維持・管理するためのコストが少ないこと、頻度の少ない現象に対して予測精

度が高いなどといったメリットがある。ガイダンスの 設計段階では様々な手法や方式を試し、予測精度だけ ではなく、保守性も含めて検討するとよい。開発や維 持管理コストの少ないガイダンスは改良や拡張が容易 であり、長期的に大きなメリットとなる。

### (3) 観測データの品質管理

目的変数として使用する観測データには、異常値が含まれていたり、地域によって品質が不均一であったり、欠測値が含まれていたりする場合がある。このような品質の低い観測データを利用すると予測精度が低下してしまうため、前処理として観測データの品質管理や補正等が必要になる。観測データの品質管理は一括学習型、逐次学習型のいずれのガイダンスにおいても重要であるが、運用中に異常な学習をする可能性があるため、逐次学習型のガイダンスにおいて特に重要である。

### (4) 説明変数の作成

説明変数は統計手法と矛盾がないように設計しなければならない。線形重回帰とカルマンフィルタを利用する場合には目的変数との線形性を確認する。ロジスティック回帰を利用する場合には実況のロジットとの線形性を確認する。ニューラルネットワークを利用する場合にはこのような制限はないが、説明変数を標準化して入力値のオーダーを揃えておく必要がある。説明変数間に強い多重共線性がある場合には、一つだけを説明変数として採用してほかは採用しないか、正則化や説明変数の選択、主成分分析などを用いることで予測への悪影響を減らすことができる。

### (5) 層別化

第2.2.5 項でも述べたように、適切な層別化は予測 精度の向上をもたらすが、細かく層別化しすぎると学 習データの数が少なくなるため、十分な学習ができず、 予測が不安定になることに注意する。層別化を検討す る場合は、数値予報モデルの特性(日中と夜間でバイ アス傾向が異なるなど)や気象学的な特性(地点によっ て風が吹きやすい方向が異なる、天気によって視程の 特性が異なるなど)を考慮しつつ、実際に学習期間の データをプロットしたり検証したりして層別化する方 法を決めることになる。時間方向の層別化には、初期 時刻層別、予報時間層別、対象時刻層別の3種類の方 法があり、データの特性に応じて1種類か2種類の層 別化が用いられる。予報時間方向に層別化すれば、予 報時間が進むにつれて数値予報モデルのランダム誤差 が大きくなることや、系統誤差が予報時間によって変 化することが考慮されることになる。また対象時刻で 層別化すれば、時間帯による現象の発生のしやすさや バイアス特性の違いなどを考慮することができる。初

<sup>1</sup> 第 3.1 節、第 3.2 節 工藤 淳

期時刻と予報時間、または、予報時間と対象時刻で層別化すれば、予報時間と対象時刻方向の層別化がともに考慮されることになるが、運用上・開発上の利便性からガイダンスでは初期時刻と予報時間の組み合わせが用いられることが多い。

層別化によって十分な学習データが得られなくなる場合は、類似した特性を持つ地点や時刻をまとめることでデータ数を増やすことができる。単純には指定した範囲の緯度経度や日付、予報時間で同じ予測式を用いるというような方法がある。より高度な方法としては、クラスター分析を用いて複数の地点や格子をまとめて扱ったり、季節変化を日付で表す代わりに-10  $^{\circ}$ C 高度で層別化するという手法も用いられている。

### (6) 予測の限界

統計手法は過去の予測と実況の関係に基づいて未来を予測する手法であるため、学習期間中にほとんど発生しなかった現象(層別化された単位でみて頻度の少ない大雨や強風など)を統計的な手法で予測することは適切ではない。学習期間を延ばすことができれば良いのだが、MOS方式を採用した場合、数値予報モデルの更新によって説明変数の特性が変化することや、数値予報データを保存するためのストレージに制約があることから、学習期間を延ばすことには限界がある。ガイダンスの開発に当たっては、統計的な限界や開発資源の制約を踏まえ、予測の限界を把握し利用者に提示することも必要である。

### (7) 予測精度と利便性

ガイダンスにとって予測精度が高いことは重要な要素ではあるが、利用のしやすさ、解釈のしやすさといった利便性も同様に重要である。複雑な構造をもつ手法を用いたり、目的変数との相関が弱い変数も説明変数として利用すれば、シンプルなガイダンスよりも予測精度は若干向上するかもしれない。しかしそれでは解釈が難しくなり、予測のシナリオや実況の推移に合わせてガイダンスを修正することは困難になってしまう。ガイダンスの目的は防災気象情報や天気予報の作成を支援することであるから、予測精度と利便性のバランスが取れたガイダンスを作成することが理想的である。開発においては予測精度の向上に重点を置きがちだが、予報現場での作業形態や利用方法の把握にも努める必要がある。

### 3.3 モデル更新への対応 1

ガイダンスは MOS (Model Output Statistics) また は statistical post-processing と呼ばれるように、主に 数値予報を使って統計処理を行っているため、数値予 報モデルの更新の影響は大きい。特にモデルの水平格 子間隔の変更はモデル地形、海陸分布の変化を伴うた め、ガイダンスが利用する数値予報格子点値 (GPV) の 特性が大きく変化する。また、モデルの物理過程の変 更も GPV の大きな特性変化を生み、時にはガイダンス の精度劣化をもたらす。ガイダンスの作成には、通常 2年間以上の数値予報データが必要となるため (Carter 1986)、モデルが更新されて特性変化が起こった場合、 新たなモデルに対応したガイダンスの予測式の作成ま で2年もの間ガイダンスの精度劣化が続く可能性があ る。これは日本に限らず、ガイダンスを維持・管理す る上で大きな問題であり、特に1980年代以降は頻繁な モデルの更新が行われるようになったため、この問題 が大きくなった (Wilson and Vallée 2002)。第1.3 節で 述べたように約40年の気象庁のガイダンスの経過もモ デル更新への対応の連続でもあり、海外の気象機関も 含めて現在もまだこの問題に取り組んでいる。

以下では、気象庁のモデル更新への対応の経過について海外の例と比較しながらまとめ、その後、ここ数年でのモデル更新での実際の対応例、海外のモデル更新への特殊な対応方法について解説する。

### 3.3.1 モデル更新への対応の歴史

第 1.3 節では気象庁のガイダンスの経過として、モデル更新の対応を既に述べた。ここでは気象庁のそれぞれの対応の詳細について、海外の事例と比較しながら解説する。

### (1) 旧モデルで作成した予測式を継続利用

1982 年 3 月の FLM-6L (格子間隔 150 km、鉛直 6層)から FLM-10L (格子間隔 127 km、鉛直 10 層)への更新の際には、旧モデルで作成した予測式を継続して利用した。この際に格子間隔の違いを緩和するために、新格子間隔の予測を旧格子間隔に内挿して影響を緩和した。

1996 年 3 月の ASM(格子間隔 75 km、鉛直 16 層)から RSM(格子間隔 20 km、鉛直 36 層)への更新の際、発雷確率、天気(雨雪判別)、視程・雲ガイダンスはそれまで運用されていた JSM(格子間隔 30 km、鉛直 19 層)でいったん予測式を作成し、それを RSMに適用した。その後も、発雷確率、雪水比ガイダンスなどは一括学習型であったため、モデル更新の際には旧モデルで作成した予測式を使用していた。

米国気象局 (NWS) では現在でも全て一括学習型の 線形重回帰 (MLR) で予測式を作成しており、NWS の ホームページには、最近のモデル更新時において予測式を継続した場合の精度評価結果が掲載されている。 精度評価結果で大きく精度が劣化した際には、その影響を緩和するために、気温や風ガイダンスの予測値に一週間から一ヶ月で計算したバイアスを加える後処理を実施する等、モデル更新への対応に苦慮している報告がある(Antolik 2012)。

### (2) 新旧モデルの並行運用

1988年3月の FLM-12L (格子間隔 127 km、鉛直 12 層) から ASM (格子間隔 75 km、鉛直 16 層) への更新の際に、ASM のデータが 2 年分蓄積されるまで新旧モデルの並行運用を行い、ガイダンスは旧モデルを利用して作成されていた。ただし、新旧モデルの並行運用のコストが大きいこと、数値予報 FAX 図が利用しているモデルとガイダンスが利用しているモデルが異なることなどの問題も生じ、これ以降はこのような対応は取られていない。

また、NWSでも2006年に領域モデルを非静力学モデルに更新する際に同じように新旧モデルの並行運用を行った例があるが、予報官が参照する数値予報は新モデルのみであり、ガイダンスに入力する旧モデルを参照できない問題があったとの報告がある(Antolik and Baker 2009)。

### (3) 逐次学習型ガイダンスの導入

1996年3月のASMからRSM(格子間隔20km、鉛直36層)への更新の際に、モデルの更新に対応可能な手法として、カルマンフィルタ(KF)及びニューラルネットワーク(NN)を利用した逐次学習型の手法が導入された。このとき、一括学習型ガイダンスも一部残ったものの、大部分のガイダンスが一括学習型から逐次学習型に移行し、現在も継続している。ただし、逐次学習型のガイダンスでもすぐにモデル更新に対応できるわけでなく、事例数が少ない現象(強風、大雨等)においてはモデル更新による影響が数か月以上残る。また、第1.2節で述べたように、逐次学習型ガイダンスは係数が不適切に変化する場合があること、予測根拠がわかりにくいこと、運用コストが高いことなどの欠点もある。

海外においても、英国、フランス、韓国等で KF が 導入されている。ただし、第 1.4 節で述べたように気 象庁ほど KF を多くの要素に利用している国はない。

# (4) 新モデルの再予報の利用

季節予報用ガイダンスでは、更新予定のモデルを過去に遡って計算する「再予報」が行われ、再予報を用いて予測式を再作成するのが一般的である。短期予報の場合は、全球モデルや領域モデルは計算機の資源を多く使うことから、再予報は現実的でなく、季節予報用ガイダンスのような処置は取られていなかった。しかし、2018年現在はある程度の期間なら再予報が可能

<sup>1</sup> 高田 伸一

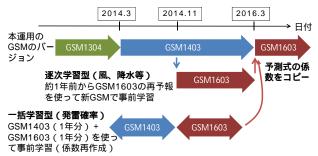


図 3.3.1 2016 年 3 月の GSM 更新におけるガイダンスの 事前学習の説明図。GSM1304, GSM1403, GSM1603 は GSM のバージョンで、ぞれぞれ 2013 年 4 月、2014 年 3 月、2016 年 3 月に導入されたことを示す。

な計算機資源が確保されているので、モデルの業務化試験<sup>2</sup>以外にガイダンスのための再予報が実施されるようになった。最近ではこれらの再予報のデータと旧モデルのデータを混ぜることによって予測式の再作成を行い、ガイダンスへの影響を小さくし、かつモデル改良がガイダンスの精度向上につながるような対処が行われている。この実例については次節で説明する。NWS も、領域モデルを非静力学化した際に、新旧モデルを混ぜてガイダンスの予測式を再作成したとの報告がある(Antolik and Baker 2009)。

### 3.3.2 現在のモデル更新への対応例

ここでは、2016 年 3 月の GSM の物理過程の変更における、再予報の実施とガイダンスの対応例について述べる。これと同様な処理が、2017 年 5 月の GSM 物理過程の変更、2017 年 2 月の MSM の力学及び物理過程の変更においても行われている。

2016年3月のGSMの更新では、物理過程の変更に よって、地上風や地上気温などの GSM の予測特性が大 きく変化し、ガイダンスへの影響調査の結果、風や発 雷確率などのガイダンスにおいて精度が劣化すること がわかった。このため、1年程度の再予報が実施され、 これを用いて事前学習(予測式の最適化)を行った結 果、概ね精度の劣化を防くことができた(後藤 2016)。 この時に行われた事前学習の処置を図3.3.1に示した。 更新予定の GSM (GSM1603) の業務化試験と再予報等 の数値予報データが、2014年11月から2016年3月 まで継続して作成され、風や降水等の逐次学習型ガイ ダンスはこの間の数値予報データを使って事前学習を 行った。また、一括学習型ガイダンスである発雷確率 は、更新前の GSM (GSM1403) と GSM1603 の数値予 報データを混ぜて予測式の再作成を行った。そして、 GSM1603 が運用開始になる当日に、予測式の係数を 差し替えた。

この事前学習の効果をみるために、逐次学習型である風ガイダンスの精度の変化を図3.3.2 に示した。図

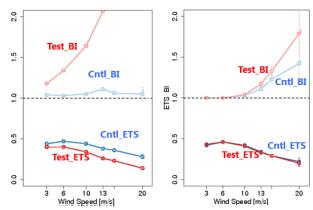


図 3.3.2 GSM1603 の業務化試験(冬)における風ガイダンスの統計検証結果と事前学習を行った後の統計検証結果。左が業務化試験結果を使った 2014 年 12 月 ~ 2015 年 2 月での検証結果、右が事前学習後の 2015 年 12 月 ~ 2016 年 2 月での検証結果。風速の各閾値でのバイアススコア (BI)とエクイタブルスレットスコア (ETS)で、BI と ETS の前についている Cntl と Test は、Cntl が当時の運用モデルである GSM1403 を入力とした風ガイダンス、Test がGSM1603 を入力とした風ガイダンスのスコアを示している。エラーバーは 95%信頼区間を示す。

の左は業務化試験(ガイダンスの事前学習なし)の冬における風ガイダンスの統計検証の結果、右は事前学習を行った場合の風ガイダンスの統計検証結果を示す。ここでは、事前学習は 2014 年 11 月から統計検証期間前の 2015 年 11 月まで 1 年間行っている。事前学習を行わない場合は風ガイダンスのバイアススコア (BI) が 1 を大きく超えて予測過剰となり、エクイタブルスレットスコア (ETS) も低くなって、精度が劣化していることがわかる。事前学習を行った結果、BI が 1 に近くなり、かつ ETS もコントロールと同程度となって、精度の劣化を抑えることができている。

また、一括学習型ガイダンスの発雷確率ガイダンスでは、図 3.3.1 に示したように GSM1403 と GSM1603 のデータを混ぜて係数を作成している。業務化試験の統計検証結果では、プライアスキルスコア (BSS) が冬に劣化、夏に中立となった(図略)。再予報を含めて係数を再作成した結果、1 か月抜き交差検証(第 2.3 節参照)からは、図 3.3.3 のように精度が向上していることがわかる。

### 3.3.3 その他のモデル更新への対応

### (1) UMOS

1980 年代から、数値予報モデルは頻繁に改良が行われ、新しいモデルに自動的に対処できる手法が考えられた。その1つは前述した KF であり、も51つは更新可能な MOS である。カナダ気象局では、KF の開発・運用コストが大きすぎることを懸念し、後者の手法である UMOS (Updateable Model Output Statistics) を選択し、2000 年から運用している (Wilson and Vallée 2002)。図 3.3.4 に UMOS システムの概念図を示す。

<sup>&</sup>lt;sup>2</sup> モデルの更新や新規モデルの運用開始において、業務化を 行うかどうかを判断するために行う過去期間の試験。

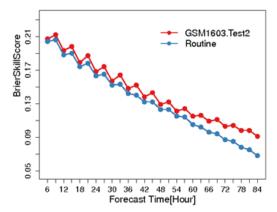


図 3.3.3 GSM1603への更新時における発雷確率ガイダンスの係数再作成を行った後の統計検証結果。2015年の1年間における予報時間ごとのプライアスキルスコア (BSS)で、青線が係数再作成を行わずに GSM1603を入力としたスコアで、赤線が係数を再作成し GSM1603を入力としたスコア。後者の検証は1か月抜き交差検証で行っている。

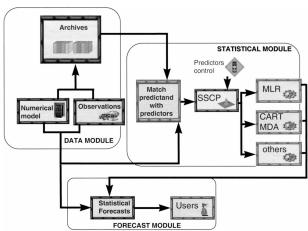


図 3.3.4 CMC (カナダ気象局)が運用する UMOS システムの概念図。predictand は目的変数、predictor は説明変数、SSCP は積和行列、MDA は判別分析。Wilson and Vallée (2002) より。

UMOS では、MLR 等に必要な目的変数と説明変数を蓄積しておき、説明変数の積和行列 (SSCP) において、旧モデルの積和と新モデルの積和の重みを徐々に変え、最終的には新モデルのサンプル数がある数 (要素によって異なるが約300個)以上になると新モデルのみを使って MLR 等を行うシステムである (図中の Predictors control が対応)。これにより、モデル更新に対応でき、かつ安定で精度の高い予測式が作成できる。しかし、この方法もコストが小さくなく、かつ新旧モデルの特性が大きく違う場合には、モデルの更新からしばらくは精度の劣化が生じると思われる。

### (2) PPM 方式の採用

第 2.2 節で述べたように PPM 方式は、予測式の作成時に説明変数として数値予報ではなく解析や目的変数以外の観測を使うため、モデル更新による予測式の

再作成の必要性はない。MOS 方式では利用するモデルのバイアスが補正されて PPM 方式より精度が高くなるが、モデル更新での数値予報のバイアスが大きく変わった場合には、PPM 方式より精度が悪くなる可能性がある。実際、NWS ではモデル更新により、PPM方式の方が高い精度となったため、MOS 方式から一時 PPM 方式のガイダンスが運用されていた例がある(Carter 1986)。

PPM 方式は、通常 MOS 方式に比べて精度が低くなるが、その精度劣化を軽減する方法として PPM 方式で作成した予測式にガイダンスを入力する方法がある。気象庁では雪水比、降水種別及び最大降水量ガイダンスに PPM 方式を利用しているが 3、予測式に入力するデータは、数値予報ではなく数値予報の系統誤差を除去したガイダンスとしている。これにより、PPM 方式で問題になる数値予報のバイアスによる精度の劣化を軽減でき、かつモデル更新にも対応できる。ただし、作成されているガイダンスの要素は限られており、かつ予測式に入力するガイダンスがモデル更新に対応する必要があるため、このような対処ができる場合は限られている。

### 参考文献

Antolik, M. S., 2012: Model Output Statistics (MOS) - Objective Interpretation of NWP Model Output, https://www.weather.gov/media/mdl/Antolik2012.pdf.

Antolik, M. S. and M. N. Baker, 2009: On the ability to develop MOS guidance with short development samples from an evolving numerical model, http://www.nws.noaa.gov/mdl/synop/amspapers/Antolik\_Baker\_2009.pdf.

Carter, G. M., 1986: Moving towards a more responsive statistical guidance system. Preprints, 11th Conf. on Weather Forecasting and Analysis, Kansas City, Amer. Meteor. Soc., 39–45.

後藤尚親, 2016: ガイダンスの特性の変化. 平成 28 年度数値予報研修テキスト, 気象庁予報部, 36-41.

Wilson, L. J. and M. Vallée, 2002: The Canadian Updateable Model Output Statistics (UMOS) System: Design and Development Tests. Wea. Forecasting, 17, 206–222.

<sup>3</sup> 最大降水量ガイダンスは、主要な説明変数である降水量にのみ PPM 方式を利用している。

#### 3.4 観測所の移設等の影響と対応1

地域気象観測システム(アメダス)の移設及び新設は、数か月に一度の頻度で行われる。また、頻度は低いが、空港の移設及び開港も数年に一度程度行われる。現在運用されているガイダンスでは、どちらの場合も観測所情報を更新することによって、システム的には観測所の移設と新設に自動的に対応できる仕組みとなっている。この節では、まずアメダスの移設時でのガイダンスの対応について述べ、その後アメダスの観測所移設等の影響とその緩和策について実例を交えて解説する。最後に空港の移設等の対応について述べる。

### 3.4.1 アメダス移設・新設時の対応

アメダス観測を目的変数とした地点形式ガイダンス である、気温(最高・最低・時系列)、風(定時・最 大・最大瞬間風速 〉降雪量(地点)湿度(最小・時 系列)<sup>2</sup> ガイダンスは、アメダスの移設・新設に対応す る必要がある。これらのガイダンスでは、アメダスの 移設・新設のどちらの場合も、新地点に距離的に最も 近いアメダス地点の予測式の係数がコピーされ3、逐 次学習によって新地点の観測特性及び数値予報の格子 点値 (GPV) の特性に次第に馴染んでゆくような仕組 みを構築している。移設・新設どちらの場合も、新地 点の観測が開始される日を設定すれば、当日に自動的 に係数等がコピーされる。なお、新設の場合には、そ の観測場所の特性に係数が馴染むまで配信を行わない 処置が取られる場合があり、精度に問題がないことを 確認した後に配信設定を変更し、配信を開始する(状 況によって異なるが、配信は観測開始から概ね数か月 後となる)。

# 3.4.2 アメダス移設時の影響と実例

アメダスの移設の際には、

高田 伸一

- 観測環境が変わり、例えば風が弱くなる等の観測 特性が変化する(目的変数の変化)
- 参照する GPV の格子点が変わり、例えばそれまで海格子を参照していた地点が陸格子を参照するといった参照 GPV の予測特性の変化(説明変数の変化)

によって、ガイダンスに影響を与える。この 2 つの影響について、それぞれ解説する。

まず、移設による観測特性の変化の影響であるが、逐次学習型ガイダンスでも新しい観測特性に馴染むまではある程度の期間が必要である。どの程度の時間がかかるかは移設前後の環境変化の程度とガイダンスの種類によって異なる。例えば、気温ガイダンスでは2週

<sup>2</sup> 湿度は気象官署(特別地域気象観測所含む)のみで、移設回数は少ない。

間程度と比較的早く新しい観測環境に馴染むが、風ガイダンスでは数か月以上かかる。ガイダンスで差が出るのは、主に予測式の層別化(風ガイダンスは風向で層別化しており、各予測式の学習回数が減る)、係数の変化を設定するカルマンフィルタ等のパラメタの違いに起因する。なお、上で示した期間の目安も平均的なものであり、特に風の場合は風向の層別化により、季節によって吹きにくい風向があるような地点では、さらに時間がかかる可能性もある。

図 3.4.1 に 2016 年 8 月 24 日の新居浜 (アメダス)の 移設における気温ガイダンスの予測を示した。この時 はアメダスが沿岸地域から内陸に移動したため、観測 特性は大きく変わったが、参照 GPV の特性は移設前 後において大きな変化はなかった(後述する GPV の 海陸に変化はなかった)。移設当日の8月24日00TC 初期値のガイダンスの予測をみると、GSM(全球モデ ル)気温ガイダンスは明らかに観測値に比べて高めに 偏っている(正バイアスがある)。これは、観測場所が 内陸に移動した結果、移設前と比べて観測値が低めに なったためと考えられる。13 日後の 9 月 6 日 00UTC 初期値のガイダンスの予測をみると、予測の外れはあ るものの、予測の正バイアスは概ね無くなっている。図 3.4.2 は同じく新居浜の移設時における風ガイダンスの 例である。移設当日の GSM ガイダンス (赤線)の風 速は GSM の風速(緑線)と同様に時間帯によって正 バイアスがある。これはアメダスが内陸に移動して風 速の観測値が弱まったためと考えられる。また、風向 の予測誤差も大きい。移設 13 日後の 9 月 6 日 00UTC 初期値においては、夜間に GSM の風速を弱めている 時間もあるが、まだ GSM の風速を十分に下方修正で きておらず、正バイアスがある。移設2か月半後の11 月9日00UTC初期値においては、ガイダンスは夜間 に GSM の風速を弱めるように下方修正しており、概 ね移設後の観測特性に合ってきている。この例からも 風ガイダンスが新しい観測特性に馴染むまでに比較的 長い期間が必要であることがわかる。

次に参照 GPV の特性変化によって生じるガイダンスへの影響について述べる。アメダスの場合、GPV の特性変化は通常はそれほど大きくないが、海岸付近に設置されていたアメダスが内陸部に移設される場合やその逆の場合は、参照する GPV が海格子から陸格子またはその逆に変化し、GPV の特性が大きく変化することがある。図 3.4.3 は秋田県男鹿半島付近での MSM地上気温の予測例である。図に示したように、海と陸では地上気温予測の差が大きく、この例では 10 °C 近くの違いが見られる。これは極端な例であるが、通常でも海と陸では数 °C 程度の違いが見られる。モデルの海岸線と実際の海岸線は一致していないため、アメダスの移設により、参照する GPV が海格子から陸格子に変わる場合、またその逆になる場合がある。これによって説明変数の特性が大きく変わり、過去の実例で

<sup>&</sup>lt;sup>3</sup> 移設の場合は、通常移設前の場所が最も近いため、移設前の地点の係数等がコピーされる。

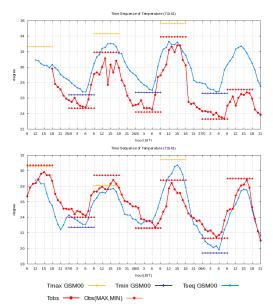


図 3.4.1 2016 年 8 月 24 日の新居浜 (アメダス)の移設時の GSM 気温ガイダンスの予測。上は移設直後の 8 月 24 日 00UTC 初期値、下は移設 13 日後の 9 月 6 日 00UTC 初期値の予測と実況。青線、黄横線、紺横線はそれぞれ GSM 気温ガイダンスの時系列、最高、最低気温ガイダンスの予測。赤線、赤横線はそれぞれ実況の時系列気温及び最高・最低気温 [°C]。最高・最低気温はどの時間帯の最高・最低気温であるかを示すために、横線としている。

も参照 GPV の海陸の違いで気温で  $3^{\circ}$ C 以上、風速でも 5 m/s 以上変化する場合があった。なお、この GPV の特性変化は影響を緩和することが可能である。前もって移設後の参照 GPV と移設前の観測で学習を行っておけば、少なくとも参照 GPV の海陸の違いは緩和できる。ただし、これらの処置は開発コストが伴うため、移設により参照 GPV の値が大きく変化する場合、かつ予報を発表している地点に限って実施している。例えば 2011 年 1 月の小田原の移設の際に、参照 GPV の海陸の違いで約 3 の変化が見られたため、MSM 気温ガイダンスにおいてこのような処置を行って影響を緩和した。

移設時に特別な対応を行った例として、2014年12月の東京のアメダスの移設がある。この場合は移設前の2013年4月から移設先の観測データが事前に並行配信された。このデータを使い、新地点での特性変化(主に観測特性の変化)を事前に学習することによって、移設直後のガイダンスの精度劣化は発生しなかった。

### 3.4.3 空港の移設、開港への対応について

空港の観測を目的変数とした航空気象ガイダンス (風、視程、雲、気温の各逐次学習型ガイダンス)においても、アメダスの移設や新設と同様に空港の移設や開港に自動的に対応する仕組みとなっている。新空港に最も距離が近い空港の係数が自動的にコピーされ、逐次学習によって徐々に新空港の観測特性や参照する

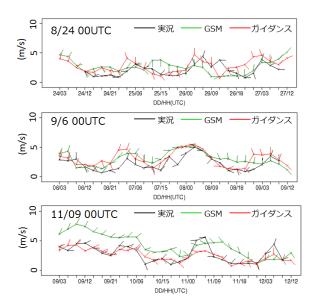


図 3.4.2 2016 年 8 月 24 日の新居浜 (アメダス) の移設時の GSM 風ガイダンスの予測。上は移設直後の 8 月 24 日 00UTC 初期値、中は移設 13 日後の 9 月 6 日 00UTC 初期値、下は移設から約 2 か月半後の 11 月 9 日 00UTC 初期値の予測と実況。緑線は GSM の予測、赤線は GSM 風ガイダンスの予測、黒線は実況の風。実線は風速、バーは風向を示す。

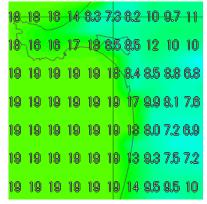


図 3.4.3 男鹿半島付近における  $\mathrm{MSM}$  の気温の予測例 ( 2016 年 10 月 18 日  $21\mathrm{UTC}$  の 24 時間後の予測 )。数字は  $\mathrm{MSM}$  の格子点値 [ $^{\circ}\mathrm{C}$ ]。細黒線は実際の海岸線。

GPV の特性に適合していく。アメダスと同様に気温ガイダンスは 2 週間程度で移設先の特性に馴染むが、その他の風、視程、雲ガイダンスは馴染むのには少なくとも数か月はかかる。航空気象予報ガイダンスのうち、ガスト発生確率及び雲底確率の 2 つは一括学習型ガイダンスであるため、予測式を更新するには 2 年以上の観測データの蓄積を待つ必要がある。観測データが蓄積された後に係数が作成されるが、それまでは移設前または近隣空港の予測式で代替することになる。なお、航空気象ガイダンスの発雷確率と降雪量については、格子形式の MSM ガイダンスから抽出しているので、空港の移設や開港の影響はない。

空港の移設例は少ないが、2006年の新北九州空港、

新種子島空港、2013年の新石垣島空港の例がある。ただし、アメダスの移設と同様に、目的変数であるMETAR等の観測が並行配信されないため、新空港への移設による観測特性の変化により、一時的にガイダンスの精度の劣化が発生する。劣化の程度は移動の状況によって変わるが、北九州空港のように海上空港に変わる場合、種子島空港のように標高が大きく変化する場合には影響が大きい4。

それに対し、空港の開港の際には、通常、開港の数か月前から観測データの配信が行われる。2005年の中部国際空港、2009年の静岡空港の際にも、開港数か月前から観測データが配信されてガイダンスの学習が行われたため、開港時には概ね精度に問題のないガイダンスが配信された。これにより、開港と同時に始まったTAFの発表作業を支援することが可能になった。

4 当時北九州空港、種子島空港ではTAF等の予報は発表されていなかったため、実際の予報作業への影響は解説への影響に限られた。

# 第4章 ガイダンスの解説

### 4.1 ガイダンス一覧<sup>1</sup>

本章では 2018 年現在運用されているガイダンスの仕様や手法、予測特性について解説する。現在運用中のガイダンスは、警報・注意報や天気予報等の発表を支援することを目的に GSM, MSM から作成される天気予報ガイダンス、飛行場予報・警報等の発表を支援する目的で MSM から作成される航空気象予報ガイダンス、空域予報の発表を支援する目的で GSM, MSM, LFM から作成される航空悪天 GPV の 3 つに大別される。以下にこれらのガイダンスの一覧を示す。MLR (線形重回帰 ) KF (カルマンフィルタ ) NN (ニューラルネットワーク ) LR (ロジスティック回帰 ) DM (診断手法 ) FBC (頻度バイアス補正 ) FC (フローチャート )  $^2$  を略語とする。

<b>=</b> 111	エケマヤギノガシュホー数
77 4.I.I	天気予報ガイダンスの一覧

で 4.1.1 人気 『報ガイタン人の一見				
	ガイダンス名	手法	対象	
	平均降水量	KF, FBC	20 km 格子 (GSM)	
第 4.2 節 降水ガイダンス	降水確率	KF	5 km 格子 (MSM)	
	最大降水量	NN, MLR	(1.12.12)	
   第 4.3 節 降雪ガイダンス	雪水比・降水種別・最大降雪量	DM	5 km 格子	
73 1.0 th P4 = 73 1 7 7 7	降雪量地点	NN, FBC	アメダス (323 地点)	
第 4.4 節 気温ガイダンス	時系列・最高・最低気温	KF	アメダス (927 地点)	
73 1.1 (1) 2(//ш/3 1 / / / /	格子形式気温	111	5 km 格子	
第 4.5 節 風ガイダンス	定時・最大・最大瞬間風速	KF, FBC	アメダス (927 地点)	
第 4.6 節 天気ガイダンス	日照率	NN	20 km 格子 (GSM)	
カモ・・・・ ハメバル・ファハ	天気	FC	5 km 格子 (MSM)	
第 4.7 節 発雷確率ガイダンス	発雷確率	LR	20 km 格子	
第 4.8 節 湿度ガイダンス	時系列湿度	KF	気象官署(153 地点、	
カモ○は一定スカーノンへ	日最小湿度	NN	特別地域気象観測所含む)	
第 4.9 節 視程ガイダンス	祖程分布予想	DM	20 km 格子 (GSM)	
ガェルは「ルゴエグコノンス	אינו קוינעדיאון אוינעדיאון	15101	5 km 格子 (MSM)	

### 表 4.1.2 航空気象予報ガイダンスの一覧

	ガイダンス名	手法	対象	
第 4.4 節 気温ガイダンス	時系列・最高・最低気温	KF		
第 4.5 節 風ガイダンス	定時・最大・ガスト風速	KF, FBC		
カモの 間 塩の 「フンハ	ガスト発生確率	LR		
第 4.9 節 視程ガイダンス	天気	FC		
カモのは「ルイモノリーノンハ	視程、視程確率	KF, FBC	国内空港(93 地点)	
   第 4.10 節 雲ガイダンス	雲量、雲底高度	NN, FBC		
おもこの 郎 去の「クンス	雲底確率	LR		
第 4.7 節 発雷確率ガイダンス 発雷確率		天気予報ガイダンス		
第 4.3 節 降雪ガイダンス	最大降雪量	から作成 <sup>3</sup>		

表 4.1.3 航空悪天 GPV の要素一覧

	手法	対象
第 4.11 節 乱気流指数	LR	   0.5 度格子 (GSM) <sup>5</sup>
第 4.12 節 着氷指数	DM	1.25 度格子 (GSM) <sup>5</sup>
第 $4.13$ 節 積乱雲量 $^4$ ・雲頂高度	DM	1.25 度倍于 (GSM)   40 km 格子 (MSM)
第 4.9 節 視程分布予想 (LFM のみ)	DM	10 km 格子 (LFM)
風、気温、湿度などの一般要素	入力とするモデルから内挿	TO KIII (EFM)

<sup>1</sup> 後藤 尚親

<sup>2</sup> 閾値を設定し事象を複数のカテゴリに分類する手法。決定木(第 2.10 節)と異なり、閾値は調査や経験則に基づく。

 $<sup>^3</sup>$  天気予報ガイダンスで作成される格子形式のガイダンスを入力とし、空港の周囲  $^4$  格子の内挿値を予測値とする。最大降雪量では周囲  $^16$  格子の最大値とその座標も提供する。

 $<sup>^4</sup>$  積乱雲量は MSM, LFM で作成している。

<sup>5</sup> 北太平洋を予測対象とする際は 0.5 度格子、全球を予測対象とする際は 1.25 度格子で作成する。

### 4.2 降水ガイダンス<sup>1</sup>

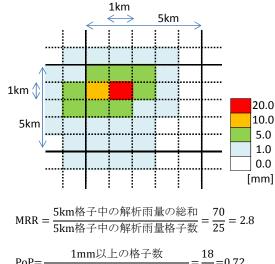
### 4.2.1 はじめに

降水ガイダンスは、日々の天気予報から警報・注意 報といった防災気象情報の発表のための予測資料とし て、また最大降雪量ガイダンス(第4.3節)や天気ガ イダンス (第 4.6 節) の入力値としても用いられるな ど、予報作業において幅広く使われている。降水に関 するガイダンスは、気象庁のガイダンスの中で最も早 くから開発が始められており、1977年に降水確率が、 1978年には平均降水量が作成され2全国の気象官署へ の配信が開始されて以降 (気象庁予報部 1991)、予測手 法の改良や予測要素の変更・追加、数値予報モデルの 更新への対応を繰り返してきた。近年も、2010年5月 に行われた予報作業支援システムの更新 (予報部予報課 2011) に合わせた変更 (小泉 2009; 小泉・蟻坂 2010) や、 24 時間最大降水量ガイダンスの予測手法の変更 (蟻坂 2013)、平均降水量、降水確率ガイダンスの改良(白山 2017) などが行われている。

本節では、降水ガイダンスを構成する、平均降水量ガイダンス (MRR)、降水確率ガイダンス (PoP)、最大降水量ガイダンス (RMAX) の 3 つのガイダンスについて解説する。まず、第 4.2.2 項で各ガイダンスの仕様と作成手法を説明し、続いて第 4.2.3 項と第 4.2.4 項で統計検証と事例検証を示す。最後に第 4.2.5 項で利用上の留意点や今後の展望について述べる。

### 4.2.2 仕様

降水ガイダンスは、GSM 及び MSM から作成する 格子形式のガイダンスである。GSM ガイダンスでは 20 km 格子(南北 12 分×東西 15 分の等緯度経度座標) MSM ガイダンスでは 5 km 格子 (南北 3 分×東西 3.75 分の等緯度経度座標)を予測対象として、MRR は一 定時間内の格子内平均の降水量を、PoP は格子内で一 定時間内の降水量が1 mm 以上となる確率を、RMAX は一定時間内の格子内最大の降水量を予測する。この 平均及び最大は、各格子に含まれる解析雨量(1 km 格 子)の平均及び最大値という意味である。図 4.2.1 は、 5 km 格子で作成する MSM ガイダンスの MRR, PoP, RMAX の目的変数の例を示しており、実線で仕切られ た格子は 5 km の MSM ガイダンスの格子、破線で仕切 られた格子は1 km の解析雨量格子を表している。図に 示すような降水量の場合には、図中央の 5 km 格子の MRR, PoP, RMAX の目的変数はそれぞれ、2.8 mm, 72.0%, 20.0 mm となる。なお、PoP の目的変数は降 水のあり・なし(つまり1または0)の格子内平均で あるため、実況降水面積率に一致する。



 $PoP = \frac{10}{5 \text{km格子中の解析雨量格子数}} = \frac{10}{25} = 0.72$ 

RMAX = (5km格子中の解析雨量の最大値) = 20

図 4.2.1 MSM ガイダンスの MRR, RMAX, PoP の目的変数の例。実線で仕切った  $5~\rm km$  の MSM ガイダンスの格子には、破線で仕切った  $1~\rm km$  の解析雨量が  $25~\rm kh$ 子含まれる。また、図中に示した計算式によって MRR, PoP, RMAXの目的変数を算出する。

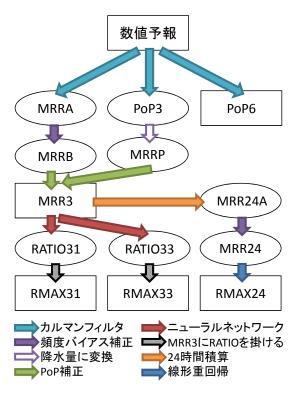


図 4.2.2 降水ガイダンスの予測要素間の関係図。矢印は凡例に示した統計手法を、数値予報以外の四角で表した要素は降水ガイダンスの主要な5つの要素を、楕円で表した要素はガイダンス計算過程の中間製品であることを表す。なお、PoP3から MRRPへの変換は、頻度バイアス補正と同様の手法を用いて行う。また、格子内最大の降水量と格子内平均の降水量の比率であるRATIO31及びRATIO33の計算には数値予報も用いるが、図では矢印を省略した。

<sup>1</sup> 白山 洋平

<sup>&</sup>lt;sup>2</sup> 一定時間の降水量が 1 mm 以上となる確率や格子内平均の 降水量という定義は現在と同じであるが、格子内に含まれる アメダス地点を対象に予測した降水確率や降水量を平均して 作成する点が現在とは異なっている。

### 表 4.2.1 平均降水量ガイダンスの仕様一覧

	V 1 3113-240 1 7 7 7 7 7 7 7 8 20
作成対象	GSM: 20 km 格子、MSM: 5 km 格子
作成方法	カルマンフィルタによる予測降水量を頻度バイアス補正後、降水確率 (PoP) で補正。
予報対象時間単位	3 時間
予報期間と間隔	GSM: FT=6 から FT=84 まで 3 時間間隔 MSM: FT=3 から FT=39 まで 3 時間間隔
逐次学習の有無	あり
説明変数	表 4.2.4 を参照
目的変数	解析雨量( $1~\mathrm{km}$ 格子)とアメダスの降水量から算出した $20~\mathrm{km}$ ( $\mathrm{MSM}$ は $5~\mathrm{km}$ )格子内の平均降水量。
層別化処理の対象	格子毎、初期時刻、予報時間(6時間区切り)
備考	頻度パイアス補正の閾値は $0.5,1,5,10,20,30,50,80,120~\mathrm{mm/3h}$ を使用 ( GSM の $11$ 月 ~ 翌 $4$ 月は $1,3,50,80~\mathrm{mm/3h}$ を使用。)。

#### 表 4.2.2 降水確率ガイダンスの仕様一覧

	代 1.2.2 片小能干ガーノンハの正体 見
作成対象	GSM: 20 km 格子、MSM: 5 km 格子
作成方法	カルマンフィルタ
予報対象時間単位	6 時間
予報期間と間隔	GSM: FT=9 から FT=81 まで 6 時間間隔 MSM: 00, 06, 12, 18UTC 初期値: FT=9 から FT=39 まで 6 時間間隔 03, 09, 15, 21UTC 初期値: FT=6 から FT=36 まで 6 時間間隔
逐次学習の有無	あり
説明変数	表 4.2.4 を参照
目的変数	解析雨量( $1~\mathrm{km}$ 格子)とアメダスの降水量から算出した $20~\mathrm{km}$ ( $\mathrm{MSM}$ は $5~\mathrm{km}$ )格子内の降水の有無の平均(実況降水面積率に同じ)。
層別化処理の対象	格子毎、初期時刻、予報時間(6 時間区切り)

### 表 4.2.3 最大降水量ガイダンスの仕様一覧

	代 1.2.0 取八件小重カーノンハの正体 見
作成対象	GSM: 20 km 格子、MSM: 5 km 格子
作成方法	1,3時間最大降水量:ニューラルネットワーク(3層:中間層はロジスティック関数、出力層は線形関数を使用)で比率(最大降水量/平均降水量)を予測し、平均降水量ガイダンス(MRR3)に比率を掛けて予測する。 24時間最大降水量:線形重回帰
予報対象時間単位	1,3 時間最大降水量: $3$ 時間、 $24$ 時間最大降水量: $24$ 時間
予報期間と間隔	GSM:1,3時間最大降水量:FT=6からFT=84まで3時間間隔 24時間最大降水量:FT=27からFT=84まで3時間間隔 MSM:1,3時間最大降水量:FT=3からFT=39まで3時間間隔 24時間最大降水量:FT=24からFT=39まで3時間間隔
逐次学習の有無	なし
説明変数	1, 3 時間最大降水量:表 4.2.4 の RMAX31, RMAX33 列を参照 24 時間最大降水量:表 4.2.4 の RMAX24 列 ( MRR24 を除く ) の主成分上位 7 位と MRR24
目的変数	解析雨量( $1~{ m km}$ 格子)から算出した $20~{ m km}$ ( $M{ m SM}$ は $5~{ m km}$ )格子内の降水量の最大値。 ただし、 $M{ m SM}$ の $1,3$ 時間最大降水量は、 $5~{ m km}$ 格子を中心とする $20~{ m km}$ 格子内の最大値。
層別化処理の対象	格子毎、平均降水量(詳細は表 4.2.5 を参照)

各ガイダンスの仕様を表 4.2.1、表 4.2.2 及び表 4.2.3 に、各ガイダンス間の関係を図 4.2.2 に示す。降水ガイダンスでは、予測要素に応じて異なる統計手法を採用しており、MRRではカルマンフィルタと頻度バイアス補正を、PoPではカルマンフィルタを用い、RMAXでは予測要素に応じてニューラルネットワークと線形重回帰を使い分けている。予測要素については、天気予報や防災気象情報などの用途に応じて対象期間を変えた5つの要素、3時間平均降水量 (MRR3)、6時間降水確率 (PoP6)、1時間最大降水量 (RMAX31)、3時間最大降水量 (RMAX31)、3時間最大降水量 (RMAX34)

を作成している $^3$ 。これら $^5$ つの他にもカルマンフィルタによる $^3$ 時間平均降水量 (MRRA) や MRRA を頻度バイアス補正した $^3$ 時間平均降水量 (MRRB)、 $^3$ 時間降水確率 (PoP3)、 $^2$ 4時間平均降水量 (MRR24) などが存在し、これらは計算過程の中間製品であったり、先の $^5$ つの予測要素の補正や説明変数として用いる。このように、降水ガイダンスは複数のガイダンスで構成

 $<sup>^3</sup>$  各要素の略称には、ガイダンス名に対象とする時間を表す数値を付したものを用いるが、1 時間最大降水量と3 時間最大降水量の場合は予報対象時刻の前3 時間内の1 時間最大降水量または3 時間最大降水量を対象としていることから、RMAX31 または RMAX33 といった表記を用いる。

された階層的な構造になっている。そのため、計算の 上流に位置するガイダンスは下流のガイダンスの予測 精度や特性に影響を与えることに留意が必要となる。

本項では、MRR3, PoP6, RMAX31, RMAX33 及びRMAX24の5つについて、順にその仕様と作成手法を説明する。なお、予測に用いる説明変数については概要を述べるのみとし、これら説明変数の作成方法は付録 4.2.A に記す。また、ガイダンスを作成する数値予報モデル名を区別する場合には、GSM-MRR3 のようにモデル名とガイダンスの要素名をつなげて表記する。

### (1) 3 時間平均降水量ガイダンス (MRR3)

MRR3は、統計手法にカルマンフィルタと頻度バイアス補正を用いて、格子内平均の3時間降水量[mm/3h]を予測するガイダンスである。格子内平均の降水量である点は数値予報モデルが予測する降水量と同じであるが、MRR3は数値予報モデルの地形と現実の地形の違いに起因する系統的な誤差を取り除いたり、モデルが予測した降水量以外の情報を基に降水量を補正したりして、その予測精度を高めている。

#### 作成方法

MRR3 は以下の手順で作成している。なお、予測の 補正に用いる PoP3 の作成方法は後述する PoP6 の作 成方法に同じである。

数値予報の GPV から説明変数を計算する。 カルマンフィルタで逐次更新した予測式で 3 時間 降水量 (MRRA) 及び PoP3 を計算する。

頻度バイアス補正で MRRA を補正する (MRRB)。 MRRB を PoP3 で補正して MRR3 とする。

まず、数値予報の GPV から表 4.2.4 の MRR3 列に示した 9 つの説明変数を計算する ( )。

で計算した説明変数を用いて3時間降水量 (MRRA) を計算し( ) この MRRA を頻度バイアス )。これは、カルマンフィルタを 補正する(MRRB、 はじめとした統計的な手法を用いて作成した予測式は 相対的に発生頻度の高い弱雨に最適化されやすく、大雨 ほど予測頻度は観測頻度に比べて低くなるため、このま までは防災気象情報を支援するのには適さないからで ある。そこで、頻度バイアス補正を用いて MRRA を補 正し、大雨の捕捉率を向上させる。頻度バイアス補正の 閾値には、GSM-MRR3、MSM-MRR3 でともに9個の 閾値「0.5, 1.0, 5.0, 10, 20, 30, 50, 80, 120 mm/3h」を 設定している。この頻度バイアス補正の閾値も MRRA の予測式と同様に逐次更新を行っており、降水頻度の 変化や数値予報モデルの改良による予測特性の変化に も対応している。

最後に、1 mm/3h 程度の弱い雨の予測精度を改善するために、MRRBをPoP3で補正して最終的な MRR3を作成する()。この補正は、1 mm 程度の弱い降水に特化した予測である PoP3 が、弱い降水についてMRRBより良い対応関係があることを利用した手法で、

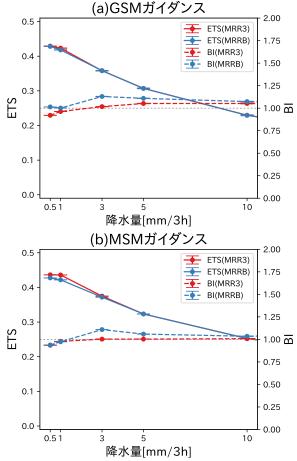


図 4.2.3 PoP 補正前後の MRRB と MRR3 の予測精度比較。(a) は GSM ガイダンス、(b) は MSM ガイダンスで、検証期間は通年(2016年1年間)。エラーバーはブートストラップ法で求めた 95%信頼区間を、赤色は MRR3、青色は MRRB、実線はエクイタブルスレットスコア、破線はバイアススコアを表す。

降水ガイダンスではこの手法を PoP 補正と呼ぶ。具体 的には、PoP3 と観測の降水量「0.5, 1.0, 3.0 mm/3h」 の対応関係を頻度バイアス補正と同じ手法を用いてあら かじめ決めておき、この対応関係を使って PoP3 を降水 量 (MRRP) に変換、PoP3 が 3.0 mm/3h に対応する閾 値以下の場合に、MRRP と MRRB を重み付き平均する ことで MRR3 を作成する。 図 4.2.3 は、PoP 補正の効果 を示す GSM 及び MSM ガイダンスの MRRB と MRR3 の2016年1年間を対象とした比較図である。図には、 10 mm/3h までのエクタイブルスレットスコア (ETS) とバイアススコア (BI) を示している。これらの図より、 PoP 補正後の MRR3 の方が GSM, MSM ガイダンスと もに BI が 1 に近づくとともに、1.0 mm/3h 前後の予 測精度が改善していることが分かる。なお、MRRB が 10 mm や 20 mm など、やや大きい場合にも PoP3 が 3.0 mm/3h に対応する閾値以下であれば PoP 補正は行 われるため、図 4.2.3 で省略している 10 mm/3h 以上 の場合にも予測値は変化するが、その数は少なく、PoP 補正前後の予測精度、頻度の差は小さくなっている。

表 4.2.4 MRR, PoP 及び RMAX の予測に用いる説明変数。なお、詳細な解説は付録 4.2.A を参照されたい。

1X 4.2.4 MIII	III, I OI 及U IIIIIAA U J AIIC用いる肌肉を致。 本の	/ IT WH '0-1	11 + DV 10 - 1 . 1	¥X 4.2.11 € ≶	~~ C 10/C V 10
	説明変数の内容	MRR3	PoP3 PoP6	RMAX31 RMAX33	RMAX24
NW85	850 hPa の北西成分の風速				
NE85	850 hPa の北東成分の風速				
NW50	500 hPa の北西成分の風速				
NE50	500 hPa の北東成分の風速				
SSI	ショワルター安定指数 (850-500hPa)				
PCWV	可降水量× 850 hPa 風速× 850 hPa 上昇流				
QWX	(上昇流×比湿×湿度×層の厚さ)				
EHQ	(基準湿度からの超過分×比湿×湿潤層の厚さ)				
OGES	地形性上昇流×比湿×湿潤層の厚さ				
DXQV	風向別降水率× 850 hPa の風速×海面と下層温位の飽和比湿差				
FRR3	数値予報モデルの3時間降水量予測				
RH85	850 hPa 相対湿度				
ESHS	(比湿×湿潤層の厚さ)/ 飽和比湿				
HOGR	地形性上昇流×相対湿度				
CFRR	数値予報モデル降水量の変換値				
D850	850 hPa <b>風向</b>				
W850	850 hPa <b>風速</b>				
OGR	地形性上昇流×比湿				
10Q4	1000 hPa 比湿と 400 hPa 飽和比湿の差				
DWL	湿潤層の厚さ				
MRR3	3 時間平均降水量ガイダンス				
Z500	500 hPa 高度				
VOR500	500 hPa 渦度				
PT500	500 hPa <b>温位</b>				
OMG850	850 hPa 鉛直 P 速度				
EPT850	850 hPa <b>相当温位</b>				
7SSI92	ショワルター安定指数 (925-700 hPa)				
QFLXZ5	500 m 高度水蒸気フラックス				
EPTZ5	500 m 高度相当温位				
7VSHZ5	500 m 高度と 700 hPa の風速鉛直シアー				
TPW	可降水量				
$IPV_{\theta}$	等温位面渦位 $(\theta = 305, 345, 355[K])$				
MRR24	24 時間平均降水量ガイダンス				

#### 説明変数

MRR3 の説明変数は、数値予報モデルの 3 時間降水量 (FRR3) をはじめ、下層の総観規模の風 (NW85, NE85) といったモデルの予測値を簡単に変換した変数、QWX や EHQ といったモデルが予測した湿りや上昇流などを掛け合わせて鉛直方向に積算した変数、DXQVのように実際に観測された格子毎の降水率をその計算に取り込んだ変数まで多岐にわたる。これらの変数は、降水ガイダンスにカルマンフィルタが導入された 1996年 3 月に、それ以前のガイダンスで用いられていた約70 種類の変数の中から多くの地点で採用されていた変数が選ばれ (瀬上ほか 1995)、さらに PCWV の定義変更や DXQV の追加が行われた (海老原 1999) ものである。

### 層別化

MRR3の予測式や頻度バイアス補正の閾値は、予測対象格子、初期時刻、予報時間で層別化している。予測対象格子による層別化は、格子毎に予測式を作成することによって数値予報と現実の地形の違いに起因する系統誤差を考慮することができる。図 4.2.4 は、20 km格子の GSM ガイダンス及び 5 km格子の MSM ガイダンスの格子定義である。MRR3と後述する PoP6では、図 4.2.4 の青、緑、黄色の 3 色で塗られた格子について予測式を作成しており、青はアメダスを含む陸格子、緑はアメダスを含まない陸格子、黄色は周辺の海格子を表している 4。黄色の格子は、解析雨量がアメ

 $<sup>^4</sup>$  図 4.2.4 の残りの赤及び灰色の格子については、MRR3 及び PoP6 では予測式を持たない。これら青、緑、黄、赤及び 灰色の格子は、後述する RMAX の予測対象格子となっており、赤及び灰色の格子については、RMAX31 及び RMAX33

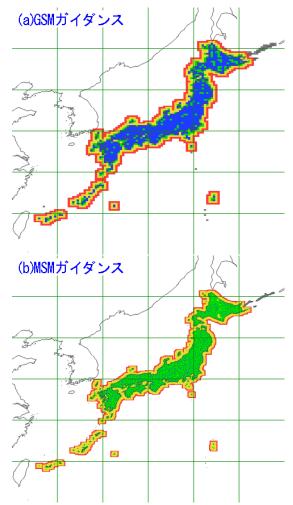


図 4.2.4 降水ガイダンスの格子定義。(a) は GSM ガイダンス、(b) は MSM ガイダンス。青はアメダスを含む陸格子、緑はアメダスを含まない陸格子、黄色は周辺の海格子を表す。 MRR 及び PoP は図に示した領域全てが予測対象だが、RMAX は上記の 3 色の格子に赤及び灰色の格子を含めた範囲を予測対象としている。

ダスで校正されない海上では精度が下がることを考慮してアメダス観測点から離れすぎていないことを条件に決定している 5。また、MRR3 と後述する PoP6 の場合は、図 4.2.4 に示した領域全体の予測値をガイダンスと称しているが、この通り全ての格子で予測式を持っているわけではない。予測式を持たない海上や日本から遠い格子については、前述の 3 色の格子の予測式を領域の外側に向かって平滑化して作成した予測式による降水量予測(補外予測)と数値予報モデルの降水量予測を平均した値をガイダンスとしている。具体的には、MRR3 の場合は数値予報モデルの 3 時間降水量と補外予測を、PoP6 の場合は説明変数の 1 つである CFRR と補外予測を、領域の外側ほど数値予報モデ

と RMAX24 でその取扱いが異なるために区別している。  $^5$  アメダス観測所は移設されたり追加されることがあるが、移設や追加の度にこの降水ガイダンスの格子定義を更新していくのは大変な手間を要するため、格子定義は固定としている。現在の格子定義は、2008 年 7 月 25 日現在のアメダス観測所情報を元に作成したものを利用している。

ルの3時間降水量や CFRR の寄与が大きくなるような 重みを付けた平均処理によってガイダンスを作成する。 なお、これら3色の格子の範囲であっても、作成した 予測式の精度が低いと考えられる場合は補外予測を利 用する。例えば、小笠原諸島は気象レーダーの観測範 囲外であるため島の格子にしか観測値が存在せず、周 囲の海上の格子は予測式を作成できないため補外予測 を利用する。南大東島周辺の海上も、レーダーの観測 範囲内であるが解析雨量がアメダスを含む島の格子で しか作成されていないため、こちらも補外予測を用い る。また、レーダーのビームカット領域に位置する格 子についても、解析雨量がその影響を受けている場合 があるため補外予測を用いてガイダンスを作成する。

初期時刻による層別化は、予報時間による層別化と組み合わせることで降水特性の日変化を予測式に反映することができる。また、予報時間による層別化は数値予報モデルの予報時間に応じた系統誤差を予測に反映することができる。MRR3と後述するPoP6では、予報時間による層別化は6時間区切りとしており、例えばGSM-MRR3では、FT=6とFT=9、FT=12とFT=15、...(以下略)のように6時間毎に予測式や頻度バイアス補正の閾値を共有する。

このほか、GSM-MRR3 の頻度バイアス補正のみ、 季節によっても層別化しており、夏から秋(5月~10 月)と冬から春(11月~翌4月)で異なる頻度バイア ス補正の閾値を用いる(白山 2017)。この季節による層 別化は、GSM-MRRA が夏季の大雨を十分に予測でき ないため、代わりに頻度バイアス補正が予測を上方修 正する特性を持ち、その特性が逐次更新では調整しき れずに冬から春においても上方修正を続けてしまい予 測頻度が観測頻度に比べて高くなることが確認された ために導入した。夏から秋には前述の9個の閾値を、冬 から春には、「1.0, 3.0, 50.0, 80.0 mm/3h」の4個の閾 値を設定している。また、切り替えは1か月かけて行 い、2 通りの閾値で求めた予測のマージ比率を徐々に変 化させることによって、層別化の切り替え時にガイダン スが急変しないようにしている。なお、MSM-MRR3 では、GSM-MRR3 のように予測頻度が過大になる傾 向が見られないため、この季節による層別化は導入し ていない。

#### 目的変数

予測式や頻度バイアス補正の閾値の逐次更新に用いる観測値には、予測対象格子に含まれる解析雨量と格子周辺のアメダスによる観測のそれぞれから計算した降水量のうち格子毎に多い方を用いる。解析雨量にはその作成過程において、すでにアメダスによる降水量観測は取り込まれているものの、気象レーダーの障害などにより解析雨量が作成されない領域が生じる場合があることやレーダーには映りにくい領域があることを考慮して、アメダスも併用した降水量を観測値に用

いている。解析雨量については、対象とする  $20~\rm km$  または  $5~\rm km$  格子に含まれる  $1~\rm km$  格子の解析雨量を平均して、 $20~\rm km$  平均または  $5~\rm km$  平均の降水量を作成する。アメダスについては、各格子点に対して半径  $20~\rm km$  以内にあるアメダス地点を対応させ、これらアメダス地点の降水量を格子点から離れるほど寄与が小さくなるような重みを付けて平均し、 $20~\rm km$  または  $5~\rm km$  の格子データへと変換する。

### (2) 6 時間降水確率ガイダンス (PoP6)

PoP6 は、統計手法にカルマンフィルタを用いて、対象とする格子内で 1 mm/6h 以上の降水がある確率 [%]を予測するガイダンスである。PoP6 は現在、気象庁が発表する天気予報の基礎資料として使われており、日本時間で 06, 12, 18, 24 JST を予報対象時刻とする 6 時間毎の予測を作成している。

#### 作成方法

PoP6 は以下の手順で作成している。なお、前述のMRR3のPoP 補正に用いるPoP3の作成方法は、予測対象時間が3時間であるだけでPoP6と同様である。

数値予報の GPV から説明変数を計算する。 カルマンフィルタで逐次更新した予測式で PoP6を計算する。

まず、数値予報モデルの GPV から表 4.2.4 の PoP 列に示す 10 個の説明変数を計算する ( )。

次に、 で計算した説明変数を用いて PoP6 を計算する( )。予測式は、MRR3 と同様にカルマンフィルタで逐次更新する。

### 説明変数

PoP6 の説明変数には、NE85 や NW85, SSI などの MRR3 と同じ説明変数を用いる一方で、MRR3 の EHQ に対応する ESHS や OGES に対応する HOGR のように、PoP6 の予測により適応しやすいように計算方法を変更した変数も用いる。これらの変数は、MRR3 の場合と同じく降水ガイダンスにカルマンフィルタが導入された 1996 年 3 月に、それ以前のガイダンスで使われていた説明変数の中から多くの地点で採用されていた変数が基本になっており (瀬上ほか 1995)、さらに海老原 (1999) において変数の定義変更や新規追加が行われたものである。

### 層別化

予測式は、MRR3と同じく予測対象格子、初期時刻、 予報時間によって層別化している。格子毎の地形特性 に起因する降水特性の違いや降水の日変化、モデルの 系統的な予報誤差に応じて層別化することによってガ イダンスの予測精度を高めている。

### 目的変数

予測式の逐次更新に用いる観測値には、予測対象格 子に含まれる解析雨量と格子周辺のアメダスによる観 測のそれぞれから計算した 1 mm/6h 以上の降水面積率のうち格子毎に高い方を用いる。解析雨量とアメダスの両方を用いる理由は、MRR3 の場合と同じである。解析雨量については、格子に含まれる解析雨量格子数に対する、1 mm/6h を超える解析雨量格子数の割合を計算する。アメダスについては、MRR3 の場合と同じく各格子点に周囲 20 km 以内のアメダスを対応させ、1 mm/6h を超えるアメダス地点数の割合を計算する。また、アメダス地点との距離に応じた重みを設定し、アメダスが格子点から離れるほど寄与を小さくする点も MRR3 の場合と同様である。

降水確率ガイダンスの逐次更新に用いる解析雨量については、MRR3 では行わない少し特別な処理を施している。解析雨量は、0.0, 0.2, 0.5, 1.0, 2.0,  $\cdots$  mm/h といった離散的な値で算出されており、また 0.2 mm/h より弱い降水を 0.2 mm/h に切り上げている。このため、解析雨量を単純に 6 時間積算すると、実際には 0.2 mm/h にも満たないようなごく弱い降水が続いた場合でも、積算値が 1 mm/6h を超える場合がある。このような観測値をそのまま降水確率ガイダンスの逐次更新に用いると、予測頻度が実際よりも高くなってしまい、特に冬季の低確率予測への影響が大きくなる。そこで、白山 (2017) は、解析雨量が 0.2 mm/h の場合には 0 mm/h として扱って積算した降水量を逐次更新に用いるようにした。

# (3) 1 時間及び 3 時間最大降水量ガイダンス (RMAX31, RMAX33)

RMAX31 及び RMAX33 は、統計手法にニューラルネットワークを用いて、格子内最大の 1 時間降水量 [mm/h] または 3 時間降水量 [mm/3h] を予測するガイダンスである。RMAX31 や RMAX33 は、現在、警報・注意報の発表や気象情報など防災気象情報の発表を検討する際の予測資料として利用されている。

はじめに、RMAX31とRMAX33の予測手法の考え 方を説明する。RMAX31とRMAX33では、MRR3や 後述するRMAX24のように直接的に降水量を計算せず、最大降水量と平均降水量の比率(RATIO、最大降 水量/平均降水量)を計算し、平均降水量にこのRATIO を掛け合わせて間接的に最大降水量を予測する手法を 採用している。

最大降水量 = 平均降水量 
$$\times$$
 RATIO (4.2.1)

RMAX31では3時間平均降水量とこの3時間内の1時間最大降水量の比率(RATIO31)を、RMAX33では3時間平均降水量と3時間最大降水量の比率(RATIO33)をニューラルネットワークで計算する。このようにRATIOを用いて間接的に最大降水量を計算する理由は次の通りである。最大降水量ガイダンスの目的は、防災気象情報の対象となる強雨を予測することであり、強雨の可能性のある時は仮に予測精度が十分でなくても観測に即した頻度で強雨を予測することが望ましい。

しかし、この比率を用いた手法に変更する以前のニュー ラルネットワークで直接的に降水量を予測していた最 大降水量ガイダンス (大澤 1996) には、RMAX33 で 30 mm 以上の予測頻度が観測頻度よりかなり少ない など、防災気象情報としての利用が難しい問題があっ た (海老原 2002)。この問題は、発現頻度の少ない強雨 を統計的な手法で直接予測しようとするために発生し ていると考えられた。また、RMAX31やRMAX33を MRR3 とは独立して予測していたため、MRR3 で降 水予測が無いときにも RMAX31 や RMAX33 が弱い 降水を予測してしまうなど、両者の予測の整合性にも 問題があった。そこで海老原(2002)は、最大降水量を (4.2.1) 式のように、平均降水量と RATIO に分解して、 この RATIO を介して最大降水量を予測する手法を採 用した。ここで平均降水量に用いる MRR3 は、頻度バ イアス補正によって強い雨も予測するように補正して いるため、この RATIO が適切に予測されれば、最大 降水量も大きな値が予測できることが期待される。ま た、MRR3が0mmであれば、もちろんRMAX31や RMAX33 も 0 mm になり、常に弱い降水を予測するこ とも、MRR3との不整合が生じることもない。この手 法は、2003年に改良された最大降水量ガイダンス (海 老原 2002) から採用されており、2010年5月に予測対 象領域が二次細分区域から格子に変更となった際も同 じ手法を継続し、予測対象の変更に合わせてニューラル ネットワークを作り直して運用している (小泉 2009)。

### 作成方法

RMAX31 及び RMAX33 は以下の手順で作成している。なお、最終的な予測値を  $5~\rm km$  格子で計算する MSM-RMAX31 及び RMAX33 についても、RATIO の計算は GSM ガイダンスと同じ  $20~\rm km$  格子のニューラルネットワークで行っている。そのため、別途計算する MRR3 の分布を元にして  $20~\rm km$  格子毎の最大降水量を  $5~\rm km$  格子毎の最大降水量へ変換する。つまり、以下の手順において までは、GSM、MSM ガイダンスともに  $20~\rm km$  格子で計算する。

数値予報の GPV から説明変数を計算する。なお、MRR3 は前述の方法で計算したものを用いる。 ニューラルネットワークで RATIO を計算する。 MRR3 と RATIO を乗じて RMAX31, RMAX33 を計算する。

MSM ガイダンスでは 20 km 格子で計算した RMAX31及びRMAX33を5km格子に変換する。

まず、数値予報モデルの GPV から表 4.2.4 の RMAX31, RMAX33 列に示す、MRR3 以外の 6 個の 説明変数を計算する ( )。 MRR3 は前述の方法で別途計算した予測値を用いる。

次に、 で計算した説明変数を用いて、ニューラルネットワークで RATIO を計算する( )。このニューラルネットワークは、入力層(7ユニット)、中間層1

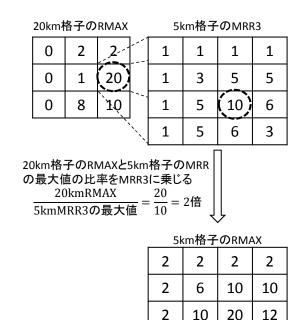


図 4.2.5 20 km 格子の RMAX から 5 km 格子の RMAX への変換方法の模式図。

2

10

12

6

層(6ユニット) 出力層(1ユニット)のシンプルなネットワークで、用いる活性化関数は入力及び中間層ではロジスティック関数を、出力層では線形関数を用いる。

で計算した RATIO と別途計算する MRR3 を乗じて、RMAX31 や RMAX33 を計算する( )。 MSM ガイダンスの場合も、GSM ガイダンスと同じ 20 km 格子の RATIO を予測するニューラルネットワークを用いるため、入力値には 5 km 格子で別途計算した MRR3を 20 km 格子の平均値に変換した値を用いる。この 20 km 格子の MRR3 と予測した RATIO から RMAXを計算し、元の 5 km 格子の MRR3 の分布に合わせて RMAXを配分して、5 km 格子の RMAX31, RMAX33を計算する( )。 MSM-RMAX31 及び RMAX33での、20 km 格子から 5 km 格子への変換手順は次の通りである(図 4.2.5)。

- 20 km 格子中に含まれる 16 個の 5 km 格子の MRR3 の最大値と 20 km 格子の RMAX との比率 を求める。MRR3 の最大値が 20 km 格子の RMAX を上回った場合には比率は 1.0 とする。
- 2. 20 km 格子中に含まれる 16 個の 5 km 格子の MRR3 に 1. で求めた比率を乗じて、5 km 格子の RMAX を計算する。

なお、RMAX31 と RMAX33 では RATIO の計算に 別々のネットワークを用いるため、それぞれの計算時 点では予測結果の大小の整合性は保証されていない。 そのため、計算の最後に RMAX31 が RMAX33 を超えないようにする整合処理を行う。

### 説明変数

RMAX31 及び RMAX33 では、RATIO の予測に表 4.2.4 の RMAX31, RMAX33 列に示した MRR3 とそ の他6つの説明変数を用いる。これらの説明変数には、 モデル改良によって予測特性が大きく変化しにくいと 考えられる変数を採用している。これは、発生頻度の低 い大雨の予測を目的とする RMAX において、その統計 関係を逐次更新することは難しいため、RATIO の計算 に用いるニューラルネットワークは重み係数を固定と しているが、同時に数値予報モデル改良時の予測精度 への影響も小さくする必要があるからである。MRR3 については、ニューラルネットワーク作成時には観測 の3時間平均降水量に置き換えている。この置き換え により、ガイダンス計算時に精度の高い、適切な予測 頻度の MRR3 を用意することができれば、対応して RMAX31 やRMAX33 の精度が高くなることが期待で きる。

続いて個々の説明変数について、その期待する効果 を説明する。MRR3 は、RATIO が平均降水量に大き く依存するため予測に有効な説明変数となる。例えば、 解析雨量の 1 km 格子が 400 個含まれる 20 km 格子に おいて、10 格子だけで50 mm/3hの降水、その他の格 子は1 mm/3hの降水の場合を考えると、平均降水量は 約 2.2 mm/3h、最大降水量は 50 mm/3h、RATIO は 約22となる。つまり、平均降水量が少ない時にはかな り大きい RATIO もあり得ることになる。一方で、平 均降水量が 30 mm/3h の場合に、RATIO が 22、つま り最大降水量が 660 mm/3h という大雨は現実には起 こり得ないだろう。このように、平均降水量を RATIO の予測に用いることは非常に重要である。D850、W850 は下層の総観規模の風の場を、SSI や 10Q4 は大気の 安定度を考慮した変数である。総観規模の風の場に対 する RATIO の反応は地域や地形により様々であると 考えられるが、格子毎に重み係数を層別化することで、 この風の場の効果を RATIO の予測に反映することが できる。また、大気の安定度が悪い場合には不安的性 の降水が生じやすく、格子内の降水が偏在して RATIO が大きくなりやすいため、大気の安定度も有効な説明 変数になると考えられる。さらに、DWL は湿潤層の 厚さを、OGR は地形の影響を考慮した変数であり、前 者は規模の大きな低気圧による降水では RATIO が小 さくなりやすいこと、後者は地形性上昇流が卓越する 場合は強雨域が限定されて RATIO が大きくなりやす いことを反映できる。

#### 層別化

RMAX31 と RMAX33 では、ニューラルネットワークに入力する平均降水量と予測対象格子について層別化を行う。まず、平均降水量による層別化について、図  $4.2.6(a)\sim(c)$  に示す、2016 年 1 年間の宮崎県、高知県、和歌山県の格子の GSM-RATIO33 の散布図を用いて

説明する。図(a)~(c)には、解析雨量から作成した観測 の RATIO33 と、説明変数の MRR3 の代わりに観測の 3 時間平均降水量を用いた予測の RATIO33 を示して いる。観測の RATIO33 は薄灰色で、予測の RATIO33 は色付きの丸でプロットしてあり、これらの各色は平 均降水量による層別化を表している(表4.2.5)。まず、 観測の RATIO33 に注目する。観測の RATIO33 は、 平均降水量が少ない場合には10倍近い値を取る一方 で、平均降水量が多くなると急激に小さくなっていき、 いずれの格子も 20 mm/3h を超えるとほぼ一定の値に なっている。このように、平均降水量の大小で RATIO の取り得る値の範囲が大きく変化をするため、これら を単一の統計関係で表現することは難しい。そこで、 RMAX31 と RMAX33 では平均降水量に対して層別 化を行い、RATIO33 の非線形な表現の予測精度を高 めている。同じ図に色付きの丸でプロットした予測の RATIO33 に注目すると、観測の RATIO33 の分布に比 べて分散がやや小さいものの、平均降水量が少ない時 には 10 倍近い RATIO33 を取りつつ、平均降水量が多 くなると急激に取りうる範囲を狭めていくといった全 体の特徴を良く表現していることが分かる。

予測対象格子による層別化は、格子毎に固定した要因である地形の影響を考慮することを目的にしている。 RATIO の大小は、不安定性降水や地形性降水の起こりやすさ、つまり格子内の平地と山地の割合といった地形の影響を大きく受けるため、予測対象格子での層別化がRATIO の予測においては非常に重要となる。

# ニューラルネットワークの作成方法

RMAX31 及び RMAX33 の RATIO を計算する二 ューラルネットワークは、前述の MRR3 や PoP6 と同 様に図4.2.4に示した青、緑、黄色の3色の格子を対象 に作成する。赤色や灰色の格子については、青色の格 子の観測値や予測値の平均値から作成したニューラル ネットワークを割り当てている。また、このニューラ ルネットワークは、「数値予報の予測が正しいとした場 合の比率」を求めると考えて、なるべく PPM (第 2.1 節)的に作成する。ここでPPM「的」と述べたのは、 ニューラルネットワーク作成時に MRR3 には観測の 3 時間平均降水量を用いるが、その他の説明変数につい ては予報時間の短い予測値を用いること、また、予測 時のMRR3には数値予報モデルの降水量でなく予測頻 度を補正したガイダンスを用いるためである。作成時 の MRR3 以外の説明変数については、モデルのスピン アップも考慮しつつ、なるべく誤差の少ない予測値を 用いるために、予報時間の短い予測値を利用する。同 時に 00, 06, 12, 18UTC 初期値の各 6 時間分の予測値 を用いることで、すべての時間帯に対応するニューラ ルネットワークを作成する。予測時の MRR3 について は、数値予報モデルの降水量予測でなくガイダンスを 用いる。モデルの降水量予測よりバイアスの軽減され

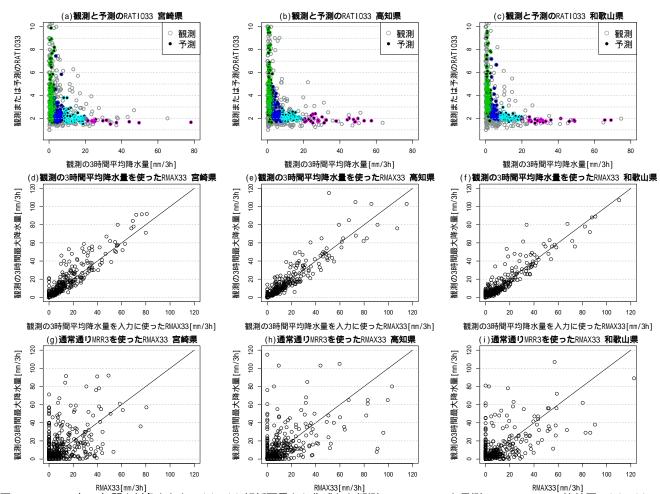


図 4.2.6 2016 年 1 年間を対象とした、 $(a)\sim(c)$  解析雨量から作成した観測の RATIO33 と予測の RATIO33 の比較図、 $(d)\sim(f)$  観測の 3 時間平均降水量を入力にして計算した RMAX33 の散布図、 $(g)\sim(i)$  通常通り MRR3 を使って計算した RMAX33 の 散布図。それぞれ、(a)、(d)、(g) は宮崎県、(b)、(e)、(h) は高知県、(c)、(f)、(i) は和歌山県の格子を対象としており、 $(a)\sim(c)$  の予測の RATIO33 の色は平均降水量による層別化を表している(表 4.2.5 を参照)。

表 4.2.5 RMAX の平均降水量による層別化範囲。降水量の単位は、RMAX31 が mm/h、RMAX33 が mm/3h、RMAX24 が mm/24h。

RMAX31	作成時	0.5 - 4	2-10	3.5 - 30	10-					
RMAX33	予測時	0.0-	3-	8.0-	20-					
RMAX24	作成時	1-5	3-8	5-12	10-22	20 - 35	30 - 50	40-100	80-200	120-
	予測時	0.0-	3-	5-	10-	20-	30-	40-	80-	120-

たガイダンスを用いることで、RATIO の予測精度を高め、RMAX の予測精度を高めている。さらに、このニューラルネットワークの学習は過去データで一括して行うが、前述の通り説明変数にはモデル改良の影響が小さいものを選んでいるため、学習に用いるデータには通常よりも長い期間のデータ  $^6$  を用いることができる。なお、現在運用している RMAX31 及び RMAX33 のニューラルネットワークの重み係数は、小泉・蟻坂 (2010) で最適化されたものである。この時は、学習に用いる十分な期間のデータを用意できなかったため、通常は乱数などを与える重み係数の初期値に最適化前の重み係数を与え、ランダムに順番を入れ替えたデー

タを繰り返し学習することで重み係数の最適化を行っている。このほかに、ニューラルネットワーク作成時とガイダンス計算時では、平均降水量による層別化の範囲を変えている(表 4.2.5)。学習時に用いるデータを平均降水量毎に分割する際に閾値付近のデータを重複して利用することで、平均降水量の層別化毎の事例数を確保して作成するニューラルネットワークの信頼性を高めるとともに、ガイダンス計算時に層が変わるタイミングで予測の不連続が生じにくいニューラルネットワークを作成する。

実際に、RMAX33のニューラルネットワークに観測の3時間平均降水量を与えてRATIO33を予測し、そのRATIO33と観測の3時間平均降水量から最大降水量

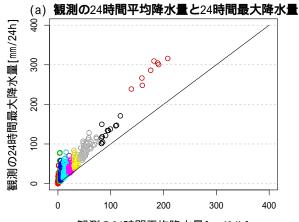
<sup>&</sup>lt;sup>6</sup> 海老原 (2002) では 4 年分のデータを用いた。

を計算した場合(その他の説明変数は数値予報モデル から与える)を図 4.2.6(d) $\sim$ (f) に、通常通りに数値予 報モデルから計算した MRR3 をニューラルネットワー クに与えて計算した場合の予測結果を図  $4.2.6(g)\sim(i)$ に示す。観測の3時間平均降水量を用いた図(d)~(f) では、強雨の予測頻度が少ないといった偏りは見られ ず、精度良くRMAX33が計算されていることが分か る。このような散布図を確認することで、RATIOを推 定するニューラルネットワークが適切に作成できるこ とが確認できる。また、この方法で計算した RMAX33 が、この RATIO を介した手法による予測精度の上限 と見ることもできる。図 (g)~(i) の通常通りに MRR3 を用いた場合では、数値予報モデルの降水予測の位置 ずれや時間ずれに起因するランダムな誤差の影響を受 けているため、予測が過大な場合や過小な場合が多く 存在するが、対角線を中心として大きな偏りのない分 布をしていることが確認できる。

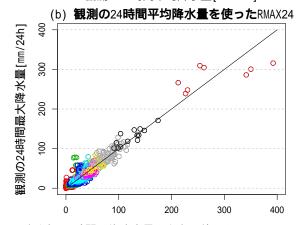
## (4) 24 時間最大降水量ガイダンス (RMAX24)

RMAX24 は、統計手法に線形重回帰を用いて、格子内最大の24時間降水量 [mm/24h] を予測するガイダンスである。現在、RMAX24も防災気象情報に記載する降水量を検討する際の予測資料として利用されている。

RMAX24 についても RMAX31, RMAX33 と同様 に、はじめに予測手法の考え方を説明する。RMAX24 は、RMAX31やRMAX33と異なり直接的に格子内の 最大降水量を予測する。以前はRMAX31やRMAX33 と同じ最大降水量と平均降水量の比率を用いた間接的な 予測手法を用いていたが、2013年3月に予測手法を線 形重回帰へ変更した $^{7}$ (蟻坂 2013)。これは、RMAX24 が予測対象とするような降水量では、平均降水量と最 大降水量の間には強い線形性があることを利用したも のである。図 4.2.7(a) は、2016 年 1 年間の熊本県内の ある格子の解析雨量から作成した観測の24時間平均 降水量と24時間最大降水量の散布図である。図より、 24 時間平均降水量が概ね 50 mm/24h 以下については、 24 時間平均降水量に対する 24 時間最大降水量のばら つきが大きいものの、概ね 50 mm/24h 以上では 24 時 間最大降水量との間に強い線形関係があることが見て 取れる。RMAX24の用途は、防災気象情報に記載する ような大雨の予測に限られているため、弱い雨の重要 度は相対的に低く、この線形関係を予測に用いても大 きな問題にはならないと考えられる。このように降水 量がある程度大きくなれば24時間平均降水量と24時 間最大降水量に強い相関があるという特徴を活かして、 実際に線形重回帰で観測の24時間平均降水量と数値予 報から計算した残りの説明変数を用いて RMAX24 を



観測の24時間平均降水量[mm/24h]



観測の24時間平均降水量を入力に使ったRMAX24[mm/24h] 図 4.2.7 (a) は観測の 24 時間平均降水量と 24 時間最大降水量の散布図、(b) は観測の 24 時間平均降水量を使った線形重回帰式による 24 時間最大降水量予測と観測の 24 時間最大降水量の散布図。プロットの各色は平均降水量による層別化を表している(表 4.2.5 を参照)。

予測した結果が図 4.2.7(b) である。ある程度の大雨については、ほぼ y=x の理想直線上に分布しており、線形重回帰による予測に大きな性能上の問題は見られず、24 時間降水量に関しては直接的に RMAX24 を計算できると考えられる。

## 作成方法

RMAX24 は以下の手順で作成している。なお、RMAX24 の予測に説明変数として用いる MRR24 の計算手順も合わせて説明する。

MRR24 を計算する。

数値予報の GPV から説明変数を計算する。 MRR24 と の説明変数の主成分上位 7 位を用い て線形重回帰式で RMAX24 を計算する。

まず、MRR24を計算する()。MRR24は、先に計算したMRR3を24時間分積算し (MRR24A)、さらに頻度バイアス補正して作成する。MRR3は既に一度、頻度バイアス補正を行った予測値であるが、第4.2.3項に示すように完全にバイアススコアを1に出来ているわけではない。そのため、単に24時間積算しただけでは MRR24Aの予測頻度が適切にならないため、再度、

<sup>7</sup> RMAX31 や RMAX33 についても、同様の線形重回帰による手法を導入する検討が行われたが、夏季の不安定降水等の「平均降水量が小さいが最大降水量が大きい」という状況に対して予測精度を確保できなかったため、本手法の導入は見送られている。

頻度バイアス補正を行っている。

次に、数値予報モデルの GPV から表 4.2.4 の RMAX24 列に示す MRR24 以外の 17 個の説明変数 を計算する()。

最後に、MRR24と の説明変数の主成分、定数項による線形重回帰式によってRMAX24を計算する( )。 17個の説明変数は、そのまま全て利用すると変数同士の相互作用によって予測性能が劣化する可能性があるため(多重共線性)、これら17個の説明変数の主成分の上位7位を予測に用いる。変数選択でなく主成分分析を用いるのは、どの説明変数を用いるのが良いかは予測式の開発に用いる期間中の事例に依存することが多く、単一の変数の選択が必ず優れているとは言えないと考えたためである。つまり、予測式は、MRR24と17個の説明変数の上位7位の主成分、そして定数項を用いる(4.2.2)式のような形となる。

RMAX24 = 
$$w_0 \times \text{MRR24} + w_1 \times$$
 主成分  $1 + \cdots$   
+ $w_7 \times$  主成分  $7 +$  定数項 (4.2.2)

ここで、 $w_i(i=0\sim7)$  は線形重回帰で求めた係数である。図 4.2.7(b) で示した通り、観測の 24 時間平均降水量を用いた予測は非常に高精度であることから、仮に適切な MRR24 を与えることが出来れば RMAX24の予測精度が高くなることは想像に難くない。ただしMRR24 は、MRR3 の計算から始まり、MRR3 を 24 時間分積算し、さらにもう一度頻度バイアス補正して計算するため、予報モデルが持つ誤差に加えて、各所の統計手法によって生じる誤差が蓄積、場合によっては増幅されて予期しない大雨が予想されてしまう場合もあり得る。そのため、RMAX24 では解析雨量から計算した気候値を用いて、極端に大きな予測値を防ぐ品質管理を行っている。

## 説明変数

で計算した MRR24 のほかに、SSI や EHQ, ESHS といった MRR3 や PoP6 と同じ説明変数や、近年、予報作業で広く参照されるようになった  $500~\mathrm{m}$  高度面の物理量 (加藤 2011) や等温位面渦位などを説明変数として利用する。なお、 $\mathrm{MRR24}$  については、予測式作成時には観測の 24 時間平均降水量を用いる。

## 層別化

RMAX24 の予測式は、予測対象格子と予測式に与える平均降水量によって層別化する(表 4.2.5)。予測対象格子による層別化は、他の予測要素と同じく格子毎に固定した要因である地形の影響などを考慮するために行う。GSM-RMAX24は、他の GSM ガイダンスと同じく20km 格子毎に予測式を持つ。一方で、MSM-RMAX24は MSM-RMAX31 や RMAX33 と異なり、GSM ガイダンスとは独立な 5km 格子毎の予測式を持つ。つまり、GSM、MSM の両ガイダンスで共通のニューラルネットワークを用いて 20km 格子の RMAX を計算した後

に 5 km 格子に変換する MSM-RMAX31 や RMAX33 と違い、MSM-RMAX24 では初めから 5 km 格子で計 算する。これは、RMAX24 を現在の手法に変更した 際に、直接 5 km 毎に予測した方が大雨の予測精度が 高かったためである。予測式は、図4.2.4に示した青、 緑、黄、赤色の4色の格子で作成し、灰色の格子は作 成に用いる観測値が存在しないため、単純に水平距離 が近い格子の予測式で代用する。また、予測式の作成 は RMAX31 や RMAX33 と同じく過去データを用い た一括学習によって行い、逐次更新は行わない。なお、 現手法の RMAX24 は 2013 年 3 月に運用を開始して おり、2014年に雨量観測点が新規に設置された島嶼部 (東京都青ヶ島や鹿児島県鹿児島郡十島村の平島や宝島 など)の観測データは、予測式の作成時に利用されて いないことに留意が必要である。このため、予測式の 作成に用いる期間を更新して最新の観測データを取り 込むことが課題の1つとなっている。

平均降水量による層別化は、図 4.2.7(a) で示したよ うな平均降水量に対する最大降水量のばらつき具合の 違いを考慮するために行う。その閾値も RMAX31 や RMAX33 と同様に、予測式作成時とガイダンス計算時 で変えており、作成時には閾値付近のデータを重複し て利用することで、作成する予測式の信頼性を高めて いる。一方、RMAX31やRMAX33には無い仕様とし て、予測式の作成時に各層に一定数以上のデータ数を 要求し、不足する場合にはその層の予測式は作成しな い点が挙げられる。これは、データ数が不足した状態 で作成された予測式は一般にその予測精度が低いと考 えられるためである。ガイダンス計算時には、作成で きている予測式を平均降水量の層別化について降順に 探索し、最初に見つかった予測式を利用する。なお、初 期時刻や予報時間による層別化は行っていない。これ は、RMAX24 がそもそも 24 時間を対象としているた めに降水の日変化を考慮する必要がないことや、予報 時間によるモデルの予測特性の変化は MRR3 の計算時 点で考慮されているものとして扱うからである。

## 4.2.3 統計検証

GSM 及び MSM 降水ガイダンスの統計的な検証結 果を示す。

GSM 及び MSM は、その数値予報モデル本体がそれ ぞれ、2017 年 5 月 25 日と 2 月 28 日に改良されている ため、本項ではこれら最新の数値予報モデルの予測結果 を使ったガイダンスの検証結果を示す。なお、GSM と MSM のモデル間の比較の場合は同じ格子サイズで比較 するために、MSM-MRR3 と MSM-POP6 は 20 km 格子中に含まれる 5 km 格子の平均値に、MSM-RMAX は 20 km 格子中に含まれる 5 km 格子の最大値に変換 してから検証する。

MRR3

初めに、GSM 及び MSM-MRR3 が入力とするモデルの降水量予測を改善しているかを確認する。

図 4.2.8 は、GSM 及び MSM-MRR3 の各々のモデル 3 時間降水量との予測精度比較で、ここでは降水量閾値 別 ETS 及び BI と予報時間別 ETS の比較を示している。検証期間はいずれも 2016 年の 1 年間で、検証に用いた初期値と予報時間は、GSM-MRR3 及び GSM3 時間降水量が全初期値 (00,06,12,18 UTC 初期値) の FT=6 から FT=27 (予報時間別 ETS は FT=84 ) MSM-MRR3 及び MSM3 時間降水量は 03,09,15,21 UTC 初期値の FT=3 から FT=24 (同 FT=39) である。検証対象格子は、図 4.2.4 で青、緑及び黄色で示した全国の陸上及 び周辺の海上格子である。

図 4.2.8(a) より、GSM-MRR3 は 80 mm/3h 以下、つまりほぼすべての閾値について GSM の降水量予測を改善していることが分かる。GSM には予測降水量が多いほど予測頻度が低くなる特性があり、GSM-MRR3 はこれを適切な頻度に修正していることから、予測精度の改善には予測頻度を観測頻度に近づける頻度バイアス補正が大きく寄与していると考えられる(図 4.2.8(b))。図 4.2.8(c) は、横軸に予報時間をとった、1.0, 10.0, 30.0 mm/3h の予報時間別 ETS の比較である。降水量閾値によって GSM からの改善幅に違いはあるものの、いずれの降水量でも予報時間全体にわたって GSM-MRR3 は GSM の降水量予測を改善していることが分かる。

一方で、MSM-MRR3 と MSM の降水量予測の間には、1 mm/3h 程度の弱い降水を除いて GSM で見られたほどの大きな予測精度の改善は見られない(図 4.2.8(d), (f))。ただし、BI を確認すると MSM の降水量予測には、GSM とは逆に予測降水量が多いほど予測頻度が過剰になる傾向があるが、MSM-MRR3 は概ね 80 mm/3hまで適切な予測頻度に補正していることが分かる(図 4.2.8(e))。 MSM-MRR3 は、MSM の予測頻度が観測頻度に比べて過大という予測特性を抑制しつつ予測精度は維持しており、予測の空振りを減らしているという点で MSM の降水量予測を改善している。

次に、夏季と冬季に分けて GSM-MRR3 と MSM-MRR3 を比較する。図 4.2.9 は、2016 年夏季(6月~8月)と 2016 年冬季(12月~翌2月)の GSM-MRR3 及び MSM-MRR3 の予測精度比較である。検証期間が夏季と冬季である点、冬季の予報時間 別 ETS 図の降水量閾値が 1.0, 5.0, 10.0 mm/3h である点以外は、検証に用いた初期値、予報時間、検証領域、夏季の降水量閾値は図 4.2.8 に同じである。なお、予報時間別 ETS の横軸には GSM-MRR3 の予報時間をとっており、MSM-MRR3 については予報対象時刻が揃うように 3 時間ずらしてプロットしている。図 4.2.9(a) より、夏季は概ね MSM-MRR3 の方が GSM-MRR3 より ETS が高く、予測精度が高いこ

とが分かる。また、図 4.2.9(c) からは、1 mm/3h の弱い雨を除いて MSM-MRR3 の方が GSM-MRR3 よりも予測精度が高いこと、予報時間半日程度までが特に高いことが分かる。前者は、MSM の方が数値予報モデルの水平格子間隔が細かく地形の表現が精緻であり強雨の表現が良いこと、後者はメソ解析においてレーダー反射強度データを同化 (幾田 2011) するなどして、初期時刻に近い予報時間の降水表現が良いことが反映されていると考えられる。これらの特徴から、実況経過と MSM の予測が大きく外れている場合を除いては、夏季の予報時間の初めころは特に MSM ガイダンスを利用するべきであろう。

一方で、冬季は夏季に比べてガイダンスの精度差が小 さく(図4.2.9(d)) また予報時間が長くなってくると GSM-MRR3 と MSM-MRR3 の予測精度が逆転する 場合が見られる(図 4.2.9(f))。特に 1 mm や 5 mm/3h では、1日程度で予測精度が逆転する傾向が明瞭であ る。1 mm/3h については、GSM と MSM の 3 時間降 水量を比較しても同様に1日程度で予測精度が逆転す る傾向が見られた(図略)ため、数値予報モデルの予 測精度が反映されたものと考えられる。5 mm/3h につ いては、GSM と MSM の 3 時間降水量の間では予測 精度が逆転する傾向は見られなかった(図略)が、図 4.2.8(a) 及び (d) の 5 mm/3h のモデルと MRR3 の精 度差からも分かるように、MSM-MRR3 と MSM はほ ぼ精度差が無い一方で、GSM-MRR3 は GSM の降水 量予測を改善しており、この影響で予報時間後半での 予測精度の逆転が発生したと考えられる。このように 冬季については、夏季ほどは MSM-MRR3 の方が優位 とは言えず、気象場に応じて適切なモデルのガイダン スを選択する必要がある。

## **RMAX**

続いて、GSM 及び MSM-RMAX31, RMAX33 及び RMAX24 の予測精度を示す。

図 4.2.10 は、2016 年 6 月~8 月の RMAX31 及び RMAX33 の予測精度比較である。予報時間別 ETS 図 の降水量閾値が、5.0, 10.0, 30.0 mm/h または mm/3h である点以外の検証条件は図4.2.9に同じである。図 4.2.10(a) 及び (d) より、RMAX31 及び RMAX33 に ついても、ほぼ全ての閾値で MSM ガイダンスの方 が GSM ガイダンスよりも予測精度が高くなっており、 MSM ガイダンスの方が短時間の強雨の予測に優れてい ると言える。しかし、予測頻度についてはGSM, MSM 両ガイダンスともに過小傾向が見られ、防災気象情報 に用いるガイダンスとしては最適な状態とは言い難い (図 4.2.10(b) 及び (e))。この原因として、MRR3 以外 の説明変数の予測特性がRMAX31,RMAX33のニュー ラルネットワーク作成時点から変わってきていること が考えられる。RMAX31 及び RMAX33 のニューラル ネットワークは、2010年の予報作業支援システムの運

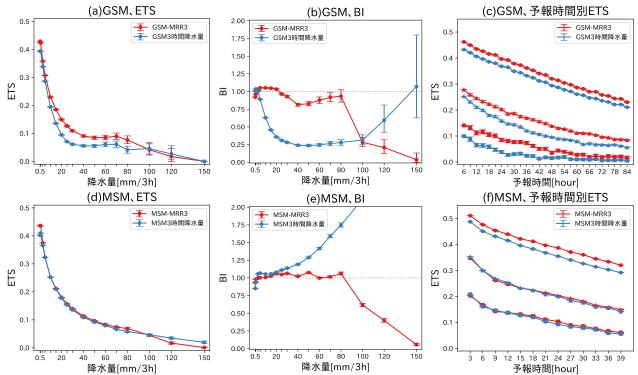


図 4.2.8 GSM-MRR3 及び MSM-MRR3 の各々のモデルの 3 時間降水量との予測精度比較。 $(a)\sim(c)$  は通年(2016 年 1 年間)の GSM-MRR3 と GSM3 時間降水量の比較、 $(d)\sim(f)$  は通年の MSM-MRR3 と MSM3 時間降水量の比較。(a)、(d) は ETS、(b), (e) は BI、(c), (f) は予報時間別 ETS で、降水量閾値は 1.0 mm/(f)3 (f)4 (f)5 (f)6 (f)7 (f)8 (f)8 (f)9 (f)

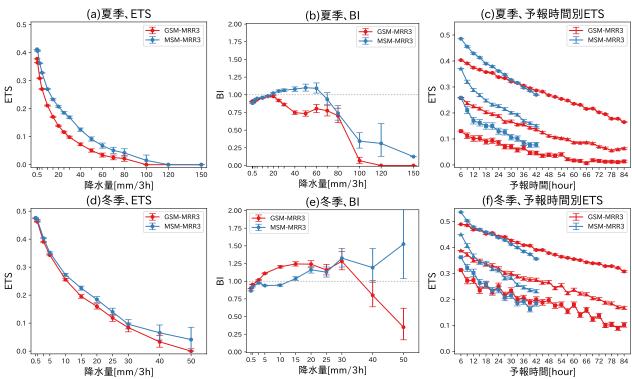


図 4.2.9 GSM-MRR3 と MSM-MRR3 の季節別の予測精度比較。 $(a)\sim(c)$  は夏季(2016 年 6 月~8 月)の GSM-MRR3 と MSM-MRR3 の比較、 $(d)\sim(f)$  は冬季(2016 年 12 月~2017 年 2 月)の比較。(c) 及び (f) 図の横軸には GSM-MRR3 の 予報時間をとり、MSM-MRR3 は予報対象時刻が揃うように 3 時間ずらしてプロットしている。また、(c) の降水量閾値は 1.0 mm/3h(丸)、10.0 mm/3h(三角)、30.0 mm/3h(四角)、(f) は降水量閾値 1.0 mm/3h(丸)、5.0 mm/3h(三角) 10.0 mm/3h(四角)である。各図ともエラーバーはブートストラップ法で求めた 95%信頼区間を、赤線は GSM-MRR3、青線は MSM-MRR3 を表す。

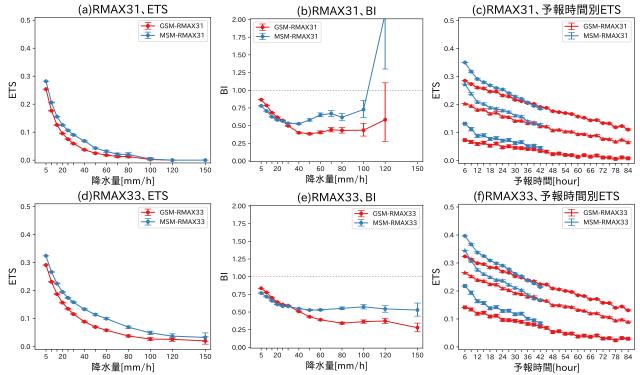


図 4.2.10 GSM 及び MSM の RMAX31, RMAX33 の予測精度比較。検証対象期間は 2016 年 6 月~8 月で、(a)~(c) は RMAX31 の ETS, BI 及び予報時間別 ETS、(d)~(f) は RMAX33 の ETS, BI 及び予報時間別 ETS。(c) 及び (f) の降水量 閾値は、5.0 mm/h または mm/3h(丸) 10.0 mm/h または mm/3h(三角) 30.0 mm/h または mm/3h(四角)で、赤線は GSM-RMAX31 または RMAX33、青線は MSM-RMAX31 または RMAX33 を表す。その他の図の見方は、図 4.2.9 に同じ。

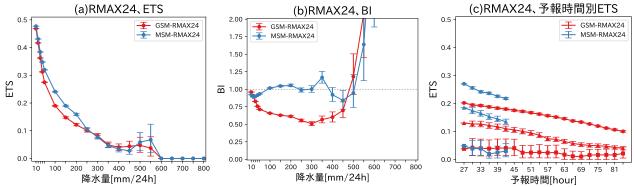
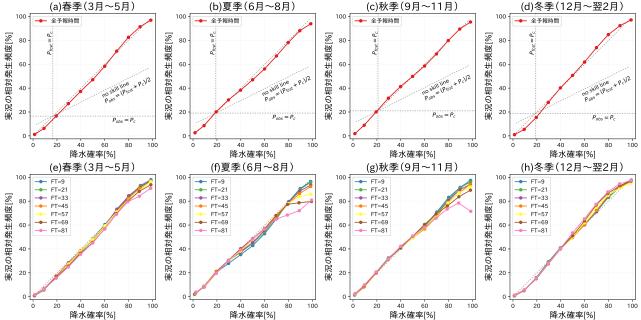


図 4.2.11 GSM-RMAX24 と MSM-RMAX24 の予測精度比較。検証対象期間は 2016 年の 1 年間で、(a) は ETS、(b) は BI、(c) は予報時間別 ETS。(c) の降水量閾値は、100 mm/24h(丸)、200 mm/24h(三角)、400 mm/24h(四角)で、赤線は GSM-RMAX24、青線は MSM-RMAX24 を表す。その他の図の見方は、図 4.2.9 に同じ。

用開始と同時に最適化されたのが最後であり、その後行われた数値予報モデルの改良には対応していない。そのため、図 4.2.8 で示したように MRR3 の予測頻度が適切であっても、RMAX31 や RMAX33 の予測頻度が適切にならないと考えられる。これら予報モデル改良による変化を取り込んだニューラルネットワークの最適化は、優先度の高い課題の1つと考えている。

MRR3 と同様に、RMAX31 やRMAX33 を予報時間 別に比較した図が図 4.2.10(c) 及び (f) である。MSM の 予報時間内は概ね MSM-RMAX31 やRMAX33 の方が GSM-RMAX31 や RMAX33 より予測精度が高い傾向 がある点は、MRR3 の場合と同じである。一方で、図 4.2.9(c) と 10 mm/3h や 30 mm/3h の閾値で比較する と、MSM-RMAX31 や RMAX33 の GSM-RMAX31 や RMAX33 からの改善幅は MRR3 の場合より小さく、MSM ガイダンスの優位性は相対的に低い。これは、RMAX31 と RMAX33 では GSM, MSM ガイダンスで同じ 20 km 格子のニューラルネットワークを RATIO の計算に用いていることが理由の 1 つに考えられる。十分な期間の開発データを準備できれば、MSM-RMAX24 のように直接 5 km 格子で作成することで予測精度の改善が期待できるため、MSM-RMAX31 や RMAX33



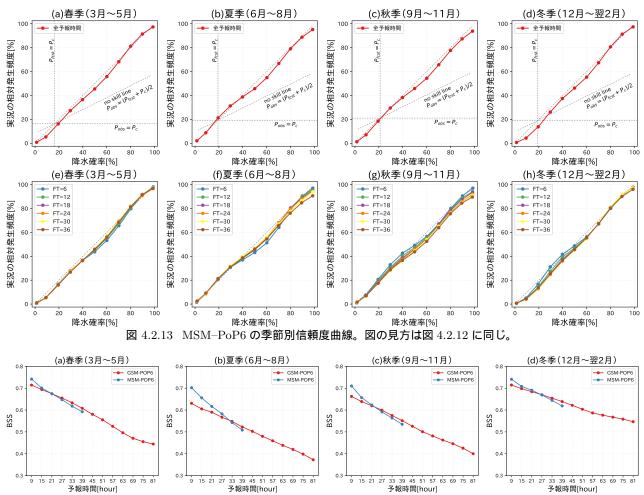


図 4.2.14 GSM-PoP6 と MSM-PoP6 の予測精度比較。(a)~(d) は季節別の GSM-PoP6 と MSM-PoP6 の BSS。季節の分け方は図 4.2.12 に同じで、赤線は GSM-PoP6、青線は MSM-PoP6 を表す。また、図の横軸には GSM-PoP6 の予報時間を用いており、MSM-PoP6 は予報対象時刻が揃うように描いている。

の作成手法の見直しも今後の課題と言えるだろう。

図 4.2.11 は、GSM 及び MSM-RMAX24 の予測精 度比較である。検証期間が2016年の1年間で予報 時間に GSM-RMAX24 では FT=27 から FT=42 (予 報時間別 ETS は FT=84) を、MSM-RMAX24 では FT=24 から FT=39 を用いた点、予報時間別 ETS の 降水量閾値が 100, 200, 400 mm/24h である点以外 は、検証条件は図 4.2.9 に同じである。RMAX24 につ いては、200 mm/24h 程度までは MSM-RMAX24 が GSM-RMAX24 よりも概ね ETS が高く、また BI も 400 mm/24h 程度まではほぼ1前後になっている(図 4.2.11(a) 及び(b) 。また、予報時間別 ETS の図から、 この傾向は MSM の予報時間内は一貫していることも 分かる (図 4.2.11(c))。 一方で、500 から 600 mm/24h では ETS が上がっていたり、BI が極端に高いなど、 個々の事例でのランダム誤差の影響を受けている可能 性や、ここまでの大雨について適切な予測式が作成で きていない可能性が考えられるなど、評価することは 難しい。ここまでの大雨の予測となると、一般に統計 的手法で精度良く予測することは非常に困難であると 考えられるため、こういった大雨の予測が出た場合に はRMAX24の予測値をそのままは使用せずに、(山本 2012) でも述べられているように大雨のポテンシャル を示す定性的な情報として扱う必要があるだろう。

## PoP6

最後に、GSM-PoP6 及び MSM-PoP6 の検証結果を 示す。図 4.2.12 及び図 4.2.13 は、季節別の PoP6 の信 頼度曲線である。ここでは、検証期間を2016年3月か ら 2017年2月とし、期間を 2016年春季(3月~5月) 夏季(6月~8月) 秋季(9月~11月) 冬季(12月 ~翌2月)の4つに分割して検証した。検証に用いた 初期値と予報時間は、GSM-PoP6 は全初期値 (00,06, 12, 18UTC 初期値) の FT=9 から FT=81 を 6 時間毎、 MSM-PoP6 は 03, 09, 15, 21UTC 初期値の FT=6 か ら FT=36 を 6 時間毎である。まず、全予報時間の図よ り、GSM-PoP6, MSM-PoP6 の両ガイダンスともに、 季節・予報時間によらず信頼度曲線は概ね y=x の理 想直線に近いが、夏季や秋季の高確率や冬季の低確率 で理想直線からやや離れる傾向が見て取れる。前者の 夏季と秋季の高確率については、予報時間別の図で詳 しく見ると、特に予報時間後半で GSM-PoP6 が理想 直線からのズレが大きいことが分かる。これは、夏季 や秋季の予報時間の長い予報の場合は、降水を精度良 く予測することが難しい影響が現れていると考えられ る。後者の冬季の低確率での過大傾向は、GSM-PoP6, MSM-PoP6で同程度に現れており、適切な降水確率予 報に資するための改良を今後も継続する必要があると 考えている。

図 4.2.14 は、GSM-PoP6 及び MSM-PoP6 の予報 時間別のブライアスキルスコア (BSS) で、横軸にとっ た GSM-PoP6 の予報時間と予報対象時刻が揃うように MSM-PoP6 を描いてある。なお、検証期間や、検証に用いた初期値は図 4.2.12 及び図 4.2.13 に同じである。この図より、いずれの季節でも予報時間が長くなると MSM-PoP6 と GSM-PoP6 の予測精度が逆転することが分かる。この特徴は、図 4.2.9(c) 及び (f) の 1.0 mm/h 閾値の検証結果に見られた特徴と合致している。特に弱い雨に関して、MSM ガイダンスの方が GSM ガイダンスより予測精度が高いのは、夏季では FT=24 程度までで、他の季節ではこれより短い(冬季ほど短い)予報時間までといった特徴には留意する必要がある。

## 4.2.4 事例検証

GSM 及び MSM 降水ガイダンスの予測特性や利用上の留意点を表す事例をいくつか紹介する。

モデルの降水量以外の情報がガイダンスに反映された例 図 4.2.15 は、平成 29 年台風第 5号 (T1705) が日本の 南海上を奄美地方に向かって西進していた8月4日9時 の日本域地上天気図、観測の3時間平均降水量、GSM3 時間降水量予測、GSM-MRR3 及び MRRA、説明変数 の1つである OGES の MRRA への寄与量である。太 平洋側の紀伊半島や四国、九州の東側斜面では、台風 との距離は離れているものの、台風の北側を回り込ん でくる暖湿気が地形の影響で強制上昇し雨雲が発生、 多い所では  $10\sim20~\text{mm/3h}$  の降水が観測されている。 この地形性の降水に対して、GSM は降水を予測して いるものの数 mm/3h しか予測しておらず過小な予測 になっている。一方で、GSM ガイダンスはモデルが 持つ降水量以外の情報を予測式に取り込むことで、こ れら東側斜面の降水予測を観測値に近づけている。図 4.2.15(f) は、この時に GSM-MRRA を増やす働きを していた、地形性上昇流と大気中の水蒸気量による降 水量への効果を見積もる OGES の寄与量である。これ ら東側斜面では OGES が GSM-MRRA に対して正の 寄与量を持っており、湿った東風が地形によって強制 上昇されるという情報が予測式に取り込まれ降水量予 測を増やしていることが分かる。このように、適切な 説明変数を設定することができれば、数値予報モデル の降水量予測が不十分な場合にガイダンスはその予測 を大きく改善できる。

## 数値予報モデルの系統誤差が修正される例

図 4.2.16 は、2017 年 8 月 25 日 9 時の日本域地上 天気図、観測の 3 時間平均降水量及び GSM-MRR3 と GSM3 時間降水量である。この日は、オホーツク 海に位置する低気圧からのびる寒冷前線が本州を南下 しており、前線上を発達しながら東進した積乱雲列に よって北陸から東北地方では未明から大雨となった。 図 4.2.16 の GSM-MRR3 と GSM の 3 時間降水量に注 目すると、GSM-MRR3 はその初期時刻に因らず一貫

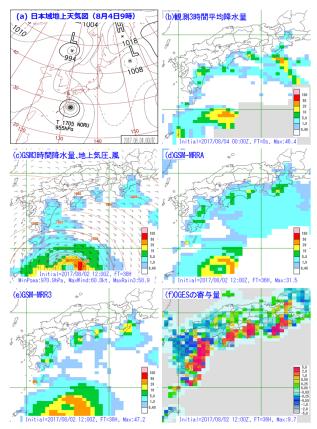


図 4.2.15 T1705 事例での 2017 年 8 月 4 日 9 時を対象とする、(a) 日本域地上天気図、(b) 観測の 3 時間平均降水量、(c) GSM の 3 時間降水量と地上気圧、風の予測、(d) GSM-MRRA、(e) GSM-MRR3、(f) OGES の寄与量。(c) $\sim$ (f) は、2017 年 8 月 2 日 12UTC 初期値の FT=36。降水量の単位はいずれも mm/3h。

して GSM の予測を上方修正して観測値に近づけている。これは、GSM の水平分解能ではこの事例で観測されたほどの降水の集中は多くの場合予想できないものの、GSM-MRR3 では主に頻度バイアス補正によって大雨の予測頻度を高めることで予測しうることを示しており、ガイダンスが数値予報モデルの系統的な誤差を修正している例と言える。

#### 夏季の短時間強雨

夏季の短時間強雨の予測は難しく、図 4.2.10 に示した通り、RMAX31 や RMAX33 の予測精度は決して高いものではない。図 4.2.17 は、2016 年 8 月 18 日の 18 時を予報対象時刻とした GSM 及び MSM-RMAX33 の予測例である。この日は北日本から東日本にかけて前線が停滞する中、上空には寒気が流入し東日本を中心に猛烈な雨となった。図 4.2.17 (c) 及び (d) は、この時の GSM 及び MSM-RMAX33 であるが、両ガイダンスともに観測値に比べて強雨の範囲が狭い。また、ピークの降水量も、MSM-RMAX33 の方が観測値に近い予測ではあるものの、観測された 150 mm/3h といった大雨には遠い予測となっている。図 4.2.10 で示した通り、RMAX33 や RMAX31 の予測頻度は観測頻度に比

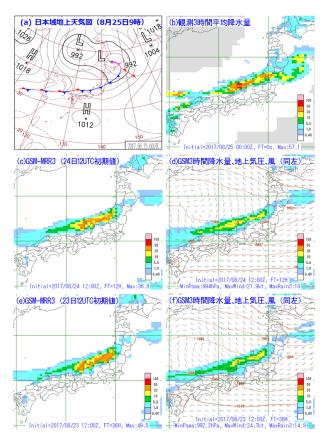


図 4.2.16 2017 年 8 月 25 日 9 時を対象とする、(a) 日本域地上天気図、(b) 観測の 3 時間平均降水量、(c) 及び (e) GSM-MRR3、(d) 及び (f) GSM の 3 時間降水量と地上気圧、風の予測。初期時刻及び予報時間は、(c) 及び (d) が 8 月 24 日 12UTC 初期値の FT=12、(e) 及び (f) は同 23 日 12UTC 初期値の FT=36。降水量の単位はいずれもmm/3h。

べて低くなっており、特に不安定性の降水ではこの事例のように GSM, MSM ガイダンスともに予測が過少傾向となることが多い。

統計検証から見える別の特徴として、予報時間のは じめ頃は予報時間後半に比べて MSM-RMAX の精度 が GSM-RMAX に比べて高いことが上げられる(図 4.2.10(c) 及び(f))。これは、図 4.2.18 に示したよう な、梅雨前線の南側で強雨域が近づいてくるような場 合の精度差を反映していると考えられ、このような場 合は GSM ガイダンスより MSM ガイダンスの方がよ い予測をする例が多い。図 4.2.18 は、梅雨前線が西日 本から東海地方に停滞する中、九州を中心に大雨となっ た事例である。この日の MSM は、初期時刻によって 前線の位置は南北に異なるものの、その南側に東西に 広がる強雨を安定して予測していた。また、前日から 前線付近では、繰り返し発生、発達する積乱雲によっ て九州では大雨が続いており、メソ解析を通してこう した観測データが取り込まれて、予報時間はじめ頃の 降水予測精度に寄与していたと考えられる。このよう に、統計検証から夏季は MSM ガイダンスの方が GSM ガイダンスよりも精度が高いが、対象とする気象場に

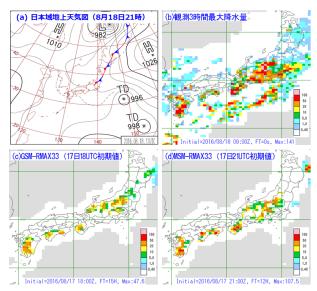


図 4.2.17 (a) 日本域地上天気図 (8月 18日 21 時)、2016 年8月 18日 18時を対象とする、(b) 観測の 3 時間最大降 水量、(c) GSM-RMAX33 (8月 17日 18UTC 初期値の FT=15)、(d) MSM-RMAX33 (8月 17日 21UTC 初期値 のFT=12)。(d) の MSM-RMAX33 は、GSM-RMAX33 との比較のために 20 km 格子に変換したものを示してい る。降水量の単位はいずれも mm/3h。

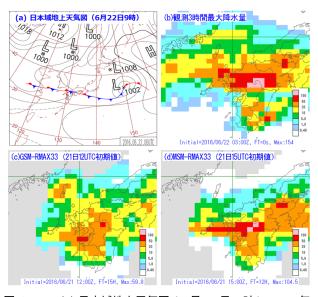


図 4.2.18 (a) 日本域地上天気図(6月22日9時)、2016年6月22日12時を対象とする、(b) 観測の3時間最大降水量、(c) GSM-RMAX33(6月21日12UTC初期値のFT=15)、(d) MSM-RMAX33(6月21日15UTC初期値のFT=12)。その他の図の見方は、図 4.2.17に同じ。

よって、そのアドバンテージには違いがあることには 留意が必要である。

#### 台風による顕著な大雨事例 (T1610)

平成 28年台風第 10号 (T1610) は、日本の南海上で複雑な動きをした後に統計開始以来初めて東北地方の太平洋側に上陸した台風である。岩手県では 300 mm/24h、北海道の十勝地方では 500 mm/24h を超える降水を観

測するなど北日本の太平洋側を中心に記録的な大雨となった。この T1610 を例に、顕著な大雨時の降水量ガイダンスの予測例とその留意点を示す。

図 4.2.19 は、T1610 が東北地方に上陸した後に日本 海へと抜けた 2016 年 8 月 31 日 3 時の GSM-RMAX24 及びモデルの地上予測などである。GSM-RMAX24で は図 4.2.19(c) に示すように、岩手県で 900 mm/24h 近 い大雨を予測するなど地域の日降水量記録を8 をはる かに上回る、やや現実離れしたガイダンスが予測され ていた。図 4.2.19(e) は、(c) と同じ初期時刻の 31 日 3 時を対象とする GSM の 3 時間降水量と GSM が予 測した T1610 の経路図 (四角印)で、T1610 のベスト トラックも合わせて赤丸印で示している。GSM が予 測した T1610 の進路は、ベストトラックに比べて 30 日日中の北北西進が遅く、31日未明にようやく岩手県 に上陸する予想だが、実際には30日21時の時点で台 風は既に日本海側に達していた。このことから、28日 12UTC 初期値の予測は、東北地方の東海上での台風 の進行速度が実況に比べて遅かったため、太平洋側の 地域は降水が実況よりも長く続き、ガイダンスも観測 値の3倍近い予測になってしまったと考えられる。比 較のために GSM の予測進路がベストトラックに大き く近づいた 30 日 00UTC 初期値の予測も確認する(図 4.2.19(b), (d) 及び(f))。31 日 12 時までの岩手県の RMAX24 は、200 mm/24h 前後と概ね実況並の予測 となっている。このように台風による大雨では台風の 進路はもちろんのこと、進行速度の違いが大雨の継続 時間に差をつけ、ガイダンスの予測を大きく左右する。 また、28 日 12UTC 初期値のような過去の観測記録か ら大きく外れたガイダンスは、統計手法で適切に補正 されていない可能性が高いため、信頼性に欠ける。こ のようなガイダンスが計算された場合には、ガイダン スの量的な予測をそのままは使用せずに、まずは大雨 のポテンシャルを示す定性的な情報として取り扱う必 要がある。

#### 4.2.5 利用上の留意点と今後の展望

本節の最後に、第 4.2.3 項及び第 4.2.4 項で示した統計検証及び事例検証を踏まえた利用上の留意点と降水ガイダンスの今後の展望をまとめる。まず、統計検証から MRR3 は概ね GSM ガイダンスよりも MSM ガイダンスの方が精度が高く、その傾向は大雨ほど強いことを示した。一方で、1 mm/3h 程度の弱い降水に関しては、予報時間後半では GSM ガイダンスの方が精度が高いこと、また、この傾向は夏季よりも冬季で強く見られることを示した。季節毎の GSM 及び MSM-MRR3 の特徴を踏まえて、気象場と実況の対比から適切なモデルを選択する必要があると言えるだろう。RMAX31 やRMAX33 についても同様に、夏季は統計的には MSM

 $<sup>^8</sup>$  岩手県の沿岸部の気象官署及びアメダスにおける 24 時間 降水量の年間極値はおよそ  $350~\mathrm{mm/24h}$  前後。

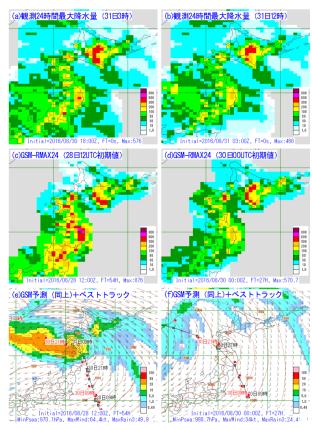


図 4.2.19 T1610 事例での、(a) 及び(b) は観測の 24 時間最大降水量 [mm/24h]、(c) 及び(d) は GSM-RMAX24 [mm/24h]、(e) 及び(f) は GSM の 3 時間降水量 [mm/3h]、地上気圧、風、T1610 の予測進路及びベストトラック。(a) は 8 月 31 日 3 時、(b) は 8 月 31 日 12 時、(c) 及び(e) は 8 月 28 日 12UTC 初期値の FT=54、(d) 及び(f) は 8 月 30 日 00UTC 初期値の FT=27。GSM による予測進路は黒四角で、ベストトラックは赤丸で 3 時間毎の中心位置を結んで示している。

ガイダンスの方が精度高いことを示した。ただし、夏季の短時間強雨で一律に MSM ガイダンスの方が良い傾向があるわけではなく、気象場によってその傾向が異なる点に留意が必要である。RMAX24 は、台風をはじめとした顕著な大雨が予測される場合に、やや現実離れした降水量を予測する場合を確認している。このような場合には、ガイダンスの量的な予測をそのままは使用せずに、大雨のポテンシャルを示す定性的な情報として取り扱う必要がある。PoP6 については、GSM, MSM 両ガイダンスともに、主に冬季の低確率の予測頻度に過大傾向が見られた。また、PoP6 の予測対象が 1 mm/6h といった弱い降水であるため、第 4.2.3 項で述べたように MRR や RMAX よりも MSM ガイダンスの GSM ガイダンスに対する優位性が低い傾向がある点に留意が必要である。

続いて、降水ガイダンスの今後の展望を述べる。まず、MRR については、説明変数の見直しやカルマンフィルタの各種パラメータなどの統計関係の再作成が必要と考えている。現在の MRR の予測に用いる説明

変数は、海老原 (1999) で改良されたのを最後に 20 年 近く同じ変数が使われている。しかし、この間、数値 予報モデルは改良や高解像度化が行われており、その 予測特性や精度は大きく変化しているため、MRR で も予測に利用する説明変数の大幅な見直しが必要と考 えられる。見直しに際しては、RMAX24で既に導入さ れている 500 m 高度面の物理量など、現在の予報現場 で利用が進んでいる各種物理量の導入も進めることで、 予測精度だけでなくガイダンスの使いやすさも含めて 改善を図っていく必要がある。RMAX については、第 4.2.3 項でも触れたが、RMAX31 及び RMAX33 に用 いるニューラルネットワークの再作成を検討している。 これは、最後にニューラルネットワークの重み係数を 最適化したのが2010年であるため、その後の数値予報 モデルの改良による予測特性の変化に RMAX を対応 させる必要があるためである。また、MSM-RMAX31 や RMAX33 については、20 km 格子での予測を介す ることなく直接 5 km 格子の RMAX を予測するよう に改良することも検討したい。これらの改良と同時に、 ディープニューラルネットワーク (第 5.2 節)で用い られる最新の技術も取り込んでいくことで、予測精度 のさらなる改善が期待できる。RMAX24についても、 2013年3月の現手法の運用開始以降、予測式の再作成 を行っていない。第4.2.2項で述べたように、予測式の 再作成によって運用開始以降に設置された島嶼部の観 測データを利用した予測式に更新することが必要であ る。PoP については、MRR と同様に説明変数の見直 しが必要な他に、現在のカルマンフィルタを使った線 形予測から、確率値の予測に適したロジスティック回 帰へと変更する必要がある。ただし、現在数値予報課 で運用しているロジスティック回帰によるガイダンス はすべて係数固定型である。ロジスティック回帰の導 入の際には、数値予報モデルの改良時に新しいモデル への適応が容易になるように、逐次学習型のガイダン スとする必要もあるだろう。

そのほかにも、メソアンサンブル予報システム (MEPS) を入力とした MEPS 降水ガイダンス (第 5.1.1 項(1)) や LFM を入力とした LFM 降水ガイダンス(第 5.1.1 項 (2) ) の開発、これらと既存の GSM 及び MSM のガイダンスを統合して単一の予測資料にした統合型 ガイダンス (第5.1.2項)の開発、さらに48,72時間最 大降水量ガイダンスなどの新しいガイダンスの開発を 計画している。MEPS 降水ガイダンスについては、既 存の MSM ガイダンスと同様の統計補正をアンサンブ ルメンバー毎に行うことで、モデル降水量から抽出す る場合よりも高精度な予測の信頼度情報やより実況の 推移に近いアンサンブルメンバーの降水量予測の提供 が可能になると考えている。LFM は、MSM よりも時 空間規模の小さい現象の表現力の向上や最新の実況を 反映した初期値による予測結果を MSM 以上に高頻度 かつ迅速に提供すること (永戸ほか 2013) を狙って運 用しており、この LFM の降水量予測を統計補正によって、より高精度にした LFM ガイダンスの提供を目指し開発を進めている。また、これら新しいガイダンスの開発によって、予測資料は現在よりも大幅に増えるものの、その予測資料の多さ故に適切な利用が難しくなることも懸念されている。そこで、予報モデルや解像度、更新頻度、予報シナリオの異なった複数のガイダンスを1つの高精度な予測資料に統合することを目指した統合型ガイダンスの開発も計画している。最後の48,72 時間最大降水量ガイダンスの開発は、予報現場からの要望に基づいた計画であり、特別警報の基準の1つとなっている48 時間降水量の予測支援や、より長い期間での総雨量の予測を支援することを目的に開発を計画している。

付録 4.2.A 降水ガイダンスの説明変数の算出方法 ここでは表 4.2.4 に示した降水ガイダンスで用いる説明変数の計算方法を解説する。本項では、計算方法が類似している場合や関連する説明変数をまとめて記述する。これらの説明変数の計算には、数値予報モデルの地上面予報値及び 1000, 925, 850, 700, 600, 500, 400, 300 hPa の計 8 層の気圧面予報値を利用しており、以下では、T は気温 [K]、P は気圧 [hPa]、RH は相対湿度 [%]、U は風の東西成分 [m/s]、V は風の南北成分 [m/s]、OMG は鉛直 P 速度 [hPa/h]、Q は比湿 [g/kg]、QS は飽和比湿 [g/kg] を表し、下付きの数値は使用する気圧面を表す。

NW85, NE85, NW50, NE50

NW85 及び NE85 は、次の式で計算する。

$$NW85 = U_{850} \times \cos \frac{315}{180} \pi + V_{850} \times \sin \frac{315}{180} \pi \quad (4.2.3)$$

NE85 = 
$$U_{850} \times \cos \frac{225}{180} \pi + V_{850} \times \sin \frac{225}{180} \pi$$
 (4.2.4)

風の東西成分  $U_{850}$  及び南北成分  $V_{850}$  に NW-SE 方向及び NE-SW 方向の単位ベクトルを乗じることで、北西-南東方向及び北東-南西方向の風速 [m/s] を計算する。 NW50 及び NE50 も、使用する気圧面が異なるだけで計算方法は同じである。

## SSI, 7SSI92

 $SSI^9$ (Showalter 1953) は、大気の成層安定度を表す指数の1つであるショワルター安定指数である。

$$SSI = T_{500} - T_{pc} (4.2.5)$$

 $T_{pc}$  は  $850~\mathrm{hPa}$  にある空気の気温と露点温度から出発し、空気塊を  $500~\mathrm{hPa}$  まで強制的に持ち上げた場合の気温  $[\mathrm{K}]$  である。その際、空気塊が飽和に達するまでは乾燥断熱減率に沿って、それ以降は湿潤断熱減率に沿って持ち上げる。また、 $7\mathrm{SSI}92$  は、大気下層の成層安定度に注目して  $925-700~\mathrm{hPa}$  間で計算した  $\mathrm{SSI}$  である。

## TPW, PCWV

可降水量 (TPW) 及び PCWV は、次の式で計算する。ただし、 ${
m OMG_{850}[hPa/h]}$  が正、つまり下降流の場合には  ${
m PCWV}=0$  とする。

$$TPW = \frac{1}{g} \sum_{n=1}^{8} (Q_n \times \delta P_n)$$
 (4.2.6)

$$\delta P_n = (P_{n-1} - P_{n+1})/2 \tag{4.2.7}$$

$$PCWV = -TPW \times WS_{850} \times OMG_{850}$$
 (4.2.8)

g は重力加速度  $[{
m ms}^{-2}]$ 、n は気圧面の層番号、 $\delta P_n [{
m hPa}]$  は積算する気圧面を挟む上下の気圧面から定義する層の厚さ、 ${
m WS}_{850}$  は  $850~{
m hPa}$  の風速  $[{
m m/s}]$  である。 $\delta P_n$  の計算には前述の  $1000~{
m hPa}$  から  $300~{
m hPa}$  までの  $8~{
m Im}$ 

<sup>&</sup>lt;sup>9</sup> Showalter Stability Index

を利用し、n=1 及び n=8 の場合には  $P_0=1000$ ,  $P_9 = 250$  として計算する。降水量を予測する際に量予 測の1つの目安となる TPW に、大気下層の上昇流と 風速を加味した変数と考えられるが、PCWV 考案時の 詳細な記録は残っていない。なお、PCWV は降水ガイ ダンスにカルマンフィルタが導入された 1996 年 3 月以 前に運用されていたガイダンスの候補変数であり、約 70 種類の候補変数の中でも多くの地点で採用されてい た変数の1つであった。

PCWV は、その定義から3変数を乗じているた め、台風や熱帯低気圧などの中心付近で極端に大き な値が計算される場合がある。そのため、PCWV を 用いる GSM-MRR3 や GSM-MRR3 を入力とする GSM-RMAX31, RMAX33 に過大な降水量が予測され る場合が確認されたため、GSM-MRR3 では (4.2.9) 式 に示す変換を行い大きな値を取りにくくした PCWVc (図4.2.20)を予測に用いることで過大な降水量予測を 抑制している <sup>10</sup>(白山 2017)。

$$PCWVc = 50 \times (\log (PCWV + 25) - \log 25)$$
 (4.2.9)

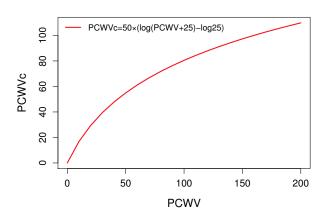


図 4.2.20 PCWV と PCWVc の関係図。

#### **QWX**

QWX は、次の式で計算する。

$$QWX = -\sum_{n=1}^{8} (OMG_n \times Q_n \times RH_n \times \delta P_n) \quad (4.2.10)$$

n 及び  $\delta P_n[\mathrm{hPa}]$  の定義は、前述の PCWV に同じであ る。また、 $\mathrm{OMG}_n[\mathrm{m/s}]$  が正、つまり下降流の層では **積算しない。上昇流のある高度に水蒸気が多いほど凝** 結が進むことを狙った変数と考えられるが、QWX 考 案時の詳細な記録は残っていない。なお、QWX は降 水ガイダンスにカルマンフィルタが導入された 1996 年 3月以前に運用されていたガイダンスの候補変数であ

り、約70種類の候補変数の中でも多くの地点で採用さ れていた変数の1つであった。

## EHQ. ESHS. DWL

はじめに、1000 hPa から 300 hPa の間で、一定基 準の相対湿度を越えている層の厚さの総和を、湿潤層 の厚さ(DWL)として定義する。ここで、湿潤とは雲 として水蒸気が凝結できる相対湿度を意味し、その基 準値は次に示す経験式で定義する。

$$f(T) = 0.01 \times (78.16 + 0.7612T - 0.01228T^{2})$$

$$(4.2.11)$$

基準湿度 f(T)[%] は気温Tの関数で、気温が低いほど 低い相対湿度で湿潤と判断される。この基準湿度を用 いて、図 4.2.21 に示す橙色の着色域の総和 DWL を計 算する。DWL は、雲の厚みを推定した指数で、降水 の有無に関係するが基準以上に湿潤な場合には違いが 現れにくいため、各層における水蒸気量などを量的に 加味した ESH 及び EHQ を次のように定義する。

$$ESH = \sum_{n=1}^{8} (DWL_n \times Qm_n)$$

$$EHQ = \sum_{n=1}^{8} (DWL_n \times \delta RH_n \times Qm_n)$$

$$(4.2.12)$$

$$EHQ = \sum_{n=1}^{8} (DWL_n \times \delta RH_n \times Qm_n)$$
 (4.2.13)

n 及び  $\delta P_n[\text{hPa}]$  の定義は、前述の PCWV に同じであ る。また、 $DWL_n$  は各気圧面での湿潤層の厚さ [hPa]、  $Qm_n[g/kg]$  と  $\delta RH_n[\%]$  は、それぞれ各湿潤層での比 湿の平均値と相対湿度の基準湿度からの超過分の平均 値である。EHQ は、DWL が正である層、つまり基準 湿度以上に湿っている層の水蒸気量が多いほど大きな 値を取る。上層が湿っている場合よりも、気温が高く 水蒸気量の多い下層が湿っている場合に値が大きくな りやすく、降水量予測に適している。

また、EHQ の特性を保ったまま、季節変化をなくし た ESHS を次のように定義する。

$$ESHS = \frac{ESH}{\sum_{n=1}^{8} QS_n}$$
 (4.2.14)

ESHS は、PoPの説明変数として EHQ を用いた場合に、 予測値に季節毎のバイアスが生じることが判明したた め改良が加えられた変数である (海老原 1999)。 (4.2.14) 式の分母は、飽和比湿 QS<sub>n</sub>[g/kg] の鉛直積算であり、 これで ESH を除すことで季節変化を取り除く。

## OGES, HOGR, OGR

OGES, HOGR 及び OGR は、地形性の上昇流に関 連した説明変数である。まず、地形データから作成し た各方位への地形勾配 (GRAD) と下層の平均風を掛け て地形性上昇流 (OGW) を計算する。地形勾配は、図 4.2.22 に示す方法で、あらかじめ GSM ガイダンスであ れば 20 km 格子毎に、MSM ガイダンスであれば 5 km

<sup>10</sup> MSM-MRR3 では、GSM-MRR3 で過大な予測が確認さ れた事例においても、問題が生じないことを確認しているた め、この処理は行わない。

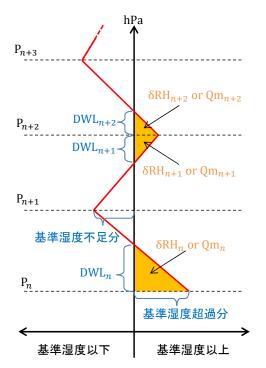


図 4.2.21 湿潤層の厚さに関する説明変数の模式図。図中の 変数の意味は本文に同じ。

格子毎に東西南北の 4 方位の勾配を計算し、定数として保持する。下層風は、1000, 850, 700 hPa 面の東西及び南北風に適当な重みを掛けた東西及び南北成分の平均風( $U_l[\mathrm{m/s}]$  及び  $V_l[\mathrm{m/s}]$ )を求め、これらから平均風速 ( $\mathrm{WS}_l[\mathrm{m/s}]$ ) と平均風向 ( $\mathrm{WD}_l$ ) を計算する。

$$U_l = (2U_{1000} + 2U_{850} + U_{700})/5 (4.2.15)$$

$$V_l = (2V_{1000} + 2V_{850} + V_{700})/5 (4.2.16)$$

$$WS_l = \sqrt{U_l^2 + V_l^2} (4.2.17)$$

$$WD_l = \arctan \frac{V_l}{U_l} \tag{4.2.18}$$

この地形性上昇流 (OGW[m/s]) に下層平均比湿  $(Q_l[g/kg])$  を掛けた値を地形性降水指数 (OGR) と定義する。なお、W は層間の距離を考慮した経験的な定数である。

$$Q_l = (Q_{1000} + 2Q_{850} + Q_{700})/W (4.2.19)$$

$$W = 1.0/(12.53 + 2.0 \times 9.63 + 6.91) \tag{4.2.20}$$

$$OGW = WS_l \times GRAD \tag{4.2.21}$$

$$OGR = OGW \times Q_l \tag{4.2.22}$$

ただし、 $\mathrm{WS}_l \leq 1.0^{-10}$  の場合には  $\mathrm{OGW} = 0$  とする。また、 $\mathrm{OGW}$  の計算では  $\mathrm{GRAD}$  は  $\mathrm{WD}_l$  のものを用いる。

PoP には、OGR のように水蒸気量を用いるより相対湿度を用いる方が対応が良いと考えられる。そこで、OGW に下層平均相対湿度  $(RH_l[\%])$  を掛けて HOGR

		10km			
	5	6	7	8	
20km	16	1	2	9	
10km	15	4	3	10	
	14	13	12	11	

格子nの標高を $H_n$ 、2点ij間の距離を $D_{i-j}$ として、格子1での西風に対する勾配を次のように定義する。

$$W_{\text{GRAD}_1} = \frac{(H_1 - H_{16}) + (H_2 - H_1)}{D_{16-2}}$$

ただし、標高差が負の場合にはOとする。 同様に、W\_GRAD<sub>2</sub>, W\_GRAD<sub>3</sub>, W\_GRAD<sub>4</sub>を求め、 総和を赤格子での西風に対する勾配W\_GRADとす る。東風、南風、北風についても同様に計算する。

図 4.2.22 各方位への地形勾配 GRAD の計算方法。図は 20 km 格子の GSM ガイダンスの場合。

を次のよう計算する。

$$RH_l = (RH_{1000} + 2RH_{850} + RH_{700})/W (4.2.23)$$

$$HOGR = OGW \times RH_l$$
 (4.2.24)

MRR には、OGR 単独よりもこれに全層的な水蒸気量を掛けた方が、OGR をトリガーとして凝結可能な水蒸気を降水量予測に反映できると考え、前述の ESH を掛けた OGES を次のように計算する。

$$OGES = OGR \times ESH \tag{4.2.25}$$

## DXQV

DXQV は、冬型降水の指数で次の式で計算する。

$$DXQV = WS_{850} \times (Q_s - Q_a) \times R(WD_{500})$$
 (4.2.26)

ただし、 $SST - T_{850} < 17$  °C の場合は、DXQV = 0 とする。ここで、 $WS_{850}$  は 850 hPa の風速 [m/s]、 $Q_s$  は海面水温 (SST[K]) での飽和混合比 [g/kg]、 $Q_a$  は 850-500 hPa の層厚温度を一般的な気温減率で海面まで下ろした時の飽和混合比 [g/kg] である。また、海面水温については、各格子から見て直近の北西象限の格子点データを用いる。 $R(WD_{500})$  は 500 hPa の風向別降水比で、冬型の気圧配置時の風向別平均値の比から格子毎に定義する。計算式のうち  $R(WD_{500})$  を除く部分は、海面からの蒸発量に 500 hPa より下層の不安定層の厚さを加味したものである。海面からの蒸発量に関係する風速には、地形の影響を受けにくくするため 850 hPa の風速を用い、 $Q_a$  に用いる気温には中層以下の不安定

層の厚さを加味するために、海面上の気温でなく850-500 hPa の層厚温度を用いる。これは下層寒気の程度が 同じでも上空にかけて寒気層が厚いほど 850-500 hPa の層厚温度は低くなり、 $Q_a$  は小さく DXQV は大きく なることを考慮したためである。また、同程度の寒気の 場合でも、西谷や東谷などの 500 hPa の高度場の違い で異なる降水特性を示す地域もあるため、 $R(WD_{500})$ を用いてこのような高度場の違いを取り込む。なお、 海面水温 Q。だけは、予測作成格子から最も近い西側 海上の水温を用いる。

#### 10Q4

10Q4 は、1000 hPa の比湿 Q<sub>1000</sub>[g/kg] と 400 hPa の飽和比湿  $QS_{400}[g/kg]$  の差で、大気の成層安定度を 表す。

$$10Q4 = Q_{1000} - QS_{400} (4.2.27)$$

#### QFLXZ5, EPTZ5, 7VSHZ5

QFLXZ5, EPTZ5 及び 7VSHZ5 は、加藤 (2011) で 提唱されている  $500 \mathrm{~m}$  高度面に関する物理量で、 $500 \mathrm{~m}$ 高度の水蒸気フラックス、相当温位及び500m高度と 700 hPa の風速鉛直シアーである。

$$QFLXZ5 = Q_{500m} \times WS_{500m}$$

$$(4.2.28)$$

$$CPTZ5 \qquad T \qquad (1000) \frac{R_d}{C_p} \qquad UW_s$$

EPTZ5 = 
$$T_{500\text{m}} \times \left(\frac{1000}{P_{500\text{m}}}\right)^{\frac{R_d}{C_p}} \times \exp\left(\frac{LW_s}{C_p \text{TD}_{500\text{m}}}\right)$$
(4.2.29)

$$7VSHZ5 = \frac{WS_{700} - WS_{500m}}{Z_{700} - Z_{500m}}$$
(4.2.30)

ここで、 $T_{500\mathrm{m}}[\mathrm{K}]$  や  $Q_{500\mathrm{m}}[\mathrm{g/kg}]$ 、 $P_{500\mathrm{m}}[\mathrm{hPa}]$  のよう に、下付きで500 m と書いた物理量は気圧面予報値か ら $500 \mathrm{~m}$  高度に鉛直内挿して計算した物理量、 $\mathrm{WS}_{500\mathrm{m}}$ は同じく  $500 \mathrm{\ m}$  高度に鉛直内挿した U と V から計算 した 500 m 高度の風速 [m/s] である。なお、500 m 高 度  $Z_{500\mathrm{m}}$  は、数値予報モデルの標高が  $300\mathrm{\ m}$  以上の場 合には、数値予報モデルの標高+200 m としている。ま た、 $R_d$  は乾燥空気の気体定数  $[\mathrm{J} \ \mathrm{kg}^{-1} \mathrm{K}^{-1}]$ 、 $C_p$  は定 圧比熱  $[J kg^{-1}K^{-1}]$ 、L は蒸発の潜熱 [J/kg]、 $W_s$  は持 ち上げ凝結高度における飽和混合比 [kg/kg]、TD<sub>500m</sub> は 500 m 高度での露点温度 [K] である。

#### IPV305, IPV345, IPV355

IPV305, IPV345 及び IPV355 は、それぞれ 305, 345 及び355 K 等温位面上の渦位で次の式で表される。

$$IPV_{\theta} = (\zeta_{\theta} + f) \left( -g \frac{\partial \theta}{\partial p} \right) \tag{4.2.31}$$

 $\theta$  は温位 [K]、 $\zeta_{\theta}$  は等温位面での相対渦度 [ $10^{-6}$ s $^{-1}$ ]、 f はコリオリパラメーター、g は重力加速度  $[m s^{-2}]$ 、 *p* は気圧 [hPa] を表す。

### FRR3

FRR3 は、数値予報モデルの 3 時間降水量 [mm] で ある。数値予報モデルからガイダンス予測対象格子へ の変換方法が GSM ガイダンスと MSM ガイダンスで は異なっており、GSM ガイダンスでは 20 km 間隔の 格子点を囲む4格子の降水量の線形内挿値を、MSMガ イダンスでは 5 km 間隔の格子点を囲む 4 格子の最大 値を用いる (木村 2004)。

#### **CFRR**

CFRR(海老原 1999) は、PoP の予測において数値予 報モデルの降水量 FRR[mm] の代わりに用いる。PoP の値は 0%から 100%に限定されるが、FRR はこれに 線形な対応をしていない。例えば、FRRが8mm程度 で PoP が 100%に近づくとすると、それ以上は 10 mm でも 100 mm でも PoP は同じである。そのため、FRR が PoP と線形な対応になるように、以下のような経験 的な式で CFRR に変換する (図 4.2.23)。

$$CFRR = FRR^2/(FRR^2 + 2)$$
 (4.2.32)

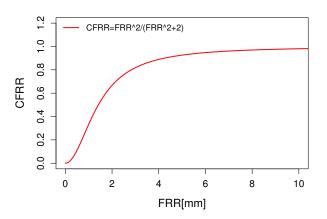


図 4.2.23 FRR と CFRR の関係図。

#### D850, W850

D850 及び W850 は、850 hPa の風向及び風速 [m/s] で、850 hPa の風の東西成分 U<sub>850</sub>[m/s] 及び南北成分  $V_{850} [\mathrm{m/s}]$  から次の式で計算する。なお、風向は計算後 に北が 0°になるように回転する。

$$D850 = \arctan \frac{V_{850}}{U_{850}} \tag{4.2.33}$$

$$W850 = \sqrt{V_{850}^2 + U_{850}^2} \tag{4.2.34}$$

## PT500, EPT850

PT500 及び EPT850 は、500 hPa の温位 [K] 及び 850 hPa の相当温位 [K] であり、次の式で表される。な お、850 hPa の相当温位は、Bolton (1980) の (38) 式 に基づいて計算する。

$$PT500 = T_{500} \times \left(\frac{1000}{P_{500}}\right)^{\frac{R_d}{C_p}} \tag{4.2.35}$$

EPT850 = 
$$T_{850} \times \left(\frac{1000}{P_{850}}\right)^{\frac{R_d}{C_p}} \times \exp\left(\frac{LW_s}{C_p \text{TD}_{850}}\right)$$
(4.2.36)

なお、 $R_d$  は乾燥空気の気体定数  $[J kg^{-1}K^{-1}]$ 、 $C_p$  は定圧比熱  $[J kg^{-1}K^{-1}]$ 、L は蒸発の潜熱 [J/kg]、 $W_s$  は持ち上げ凝結高度における飽和混合比 [kg/kg]、 $TD_{850}$  は 850~hPaの露点温度 [K] である。

#### Z500, VOR500, RH85, OMG850

Z500, VOR500, RH85 及び OMG850 は、500~hPa の高度 [m] と渦度  $[10^{-6}s^{-1}]$ 、850~hPa の相対湿度 [%] 及び鉛直 P 速度 [hPa/h] である。これらの変数には、数値予報モデルの気圧面予報値をガイダンスの予測対象格子を囲む 4 格子で線形内挿した値を用いる。

#### 参考文献

- 蟻坂隼史, 2013: 24 時間最大降水量ガイダンスの改良. 平成 25 年度数値予報研修テキスト, 気象庁予報部, 42-48
- Bolton, D., 1980: The Computation of Equivalent Potential Temperature. *Mon. Wea. Rev.*, **108**, 1046–1053.
- 海老原智, 1999: 降水ガイダンスの改良と検証. 平成 11 年度数値予報研修テキスト, 気象庁予報部, 23-33.
- 海老原智, 2002: 最大降水量ガイダンス. 平成 14 年度 数値予報研修テキスト, 気象庁予報部, 35-39.
- 永戸久喜, 原旅人, 倉橋永, 2013: 日本域拡張・高頻度 化された局地モデルの特性. 平成 25 年度数値予報研 修テキスト, 気象庁予報部, 18-41.
- 幾田泰酵, 2011: メソ解析におけるレーダー反射強度 データの同化. 平成 23 年度数値予報研修テキスト, 気象庁予報部, 9-12.
- 加藤輝之, 2011: 大雨発生で着目すべき下層水蒸気場の 高度. 平成 22 年度予報技術研修テキスト, 気象庁予 報部, 71-88.
- 木村陽一, 2004: 最大降水量ガイダンス. 平成 16 年度 数値予報研修テキスト, 気象庁予報部, 48-50.
- 気象庁予報部, 1991: 天気予報ガイダンスの解説. 予報 技術資料第 41 号.
- 小泉友延, 2009: 最大降水量ガイダンス. 平成 21 年度 数値予報研修テキスト, 気象庁予報部, 21-26.
- 小泉友延, 蟻坂隼史, 2010: 降水ガイダンスの改良. 平成 22年度数値予報研修テキスト, 気象庁予報部, 71-77.
- 大澤和裕, 1996: 大雨ガイダンスの解説. 平成8年度量 的予報研修テキスト, 気象庁予報部, 30-36.
- 瀬上哲秀, 大林正典, 国次雅司, 藤田司, 1995: カルマンフィルター. 平成7年度数値予報研修テキスト, 気象庁予報部, 66-78.
- 白山洋平, 2017: 降水ガイダンスの改良. 平成 29 年度 数値予報研修テキスト, 気象庁予報部, 86-93.

- Showalter, A. K., 1953: A stability index for thunderstorm forecasting. *Bull. Amer. Meteor. Soc.*, **34**, 250–252.
- 山本佳緒里, 2012: 事例で見る数値予報の特性. 平成 24 年度数値予報研修テキスト, 気象庁予報部, 61-67.
- 予報部予報課, 2011: 市町村を対象とした警報・注意報 発表のための新しい予報作業支援システムの整備に ついて. 測候時報, **78**, 1–18.

## 4.3 降雪ガイダンス<sup>1</sup>

## 4.3.1 はじめに

降雪ガイダンスは、降雪に関する予報を支援するガ イダンスであり、降雪量を予測する最大降雪量ガイダ ンスと降雪量地点ガイダンス及び降水の種別(「雨」、 「雨か雪」、「雪か雨」、「雪」の4カテゴリ)を予測する 降水種別ガイダンスがある。以降、単に降雪ガイダン スという場合はこれら3つのガイダンスの総称である。 降水種別ガイダンスと最大降雪量ガイダンスは、2010 年5月の予報作業支援システム (YSS) の更新 (予報部 予報課 2011) にあわせて運用を開始したガイダンス (古 市・松澤 2009) で、降水の種別と面的な降雪量の予報 を支援することを目的としている。また、降水種別ガ イダンスの予測値は、天気ガイダンス (第4.6節)の作 成にも用いている。これは、天気ガイダンスでも「雨」 や「雪」といった降水の種別に関連する予測を行って いるため、ユーザーが利用しやすいように予測の整合 を図るためである。降雪量地点ガイダンスは、積雪深 計が設置されているアメダス地点<sup>2</sup>を対象に降雪量を 予測するガイダンスである。1997 年 11 月に RSM を入 力として試験運用を開始し(第1.3節)、航空予報等で の利用のため 2008 年 11 月に GSM を入力としたガイ ダンスに変更され正式運用を開始した(古市 2008)。ま た、予測精度の向上と予報のさらなる支援を目的とし て、ガイダンスの改良と予測対象地点の拡充及び MSM 降雪量地点ガイダンスの開発 (白山 2016) が行われて いる。降雪量地点ガイダンスは統計手法で直接降雪量 を求めるため、雪水比と降水量から降雪量を求める最 大降雪量ガイダンスと比べて予測精度が高い(第4.3.3 項)。一方で、YSS では面的な予測資料である最大降雪 量ガイダンスが必要であることから、2種類の降雪量 ガイダンスを並行運用している。

本節ではまず、第 4.3.2 項で降雪ガイダンスの仕様と作成方法を述べ、次に第 4.3.3 項と第 4.3.4 項で統計検証と事例検証から予測特性について説明し、最後に第 4.3.5 項で利用上の留意点と今後の展望について述べる。

## 4.3.2 仕様

本項では降雪ガイダンスについて、降水種別ガイダンス (PTYP)、最大降雪量ガイダンス (MAXS)、降雪量地点ガイダンス (SNOW)の順に、その仕様と作成方法を述べる。また、最大降雪量ガイダンス作成で用いる雪水比 (降水量から降雪量への変換比率 (降雪量[cm]/降水量 [mm]))の作成方法についても解説する。

#### (1) 降水種別ガイダンス

PTYP の仕様を表 4.3.1 に示す。PTYP は、診断法 (第2.8節)を用いて全国の5 km 格子毎に前3時間降 水種別 (PTYP3) を予測する格子形式のガイダンスで あり、入力値に GSM 及び GSM ガイダンスを用いる GSM-PTYP と MSM 及び MSM ガイダンスを用いる MSM-PTYP を作成している。PTYP は MAXS や天 気ガイダンスの入力とするために面的な予測値を作成 する必要があるが、降水種別のうち特に雪の観測には 地方による不均一性があり、統計手法では観測の少な い地域についてある程度の精度を持ったガイダンスを 作ることが難しい。そのため PTYP では、全国一様な 予測精度を確保するために診断法を採用している。ま た、予測式の作成は説明変数に実況(地上気温及び地 上相対湿度)を使う PPM 方式 (第 2.2.2 項) とし、全 国で単一の予測式を用いている。予測時には格子形式 気温ガイダンス (第4.4節) とモデルの地上相対湿度を 用いる。地上相対湿度についても、モデルのバイアス が取り除かれたガイダンスを用いるのが好ましいが、 2018年現在では格子形式のガイダンスが作成されてい ないため、モデルの予測値を用いている。なお、降水 種別ガイダンスは降水の有無に関係なく全格子で降水 種別を予測する。つまり、降水があった場合の降水種 別を予測するガイダンスである。

#### 作成方法

PTYP の作成は次の順に行う。

格子形式気温ガイダンスおよびモデルの地上相対 湿度の予測値を用いて1時間降水種別を判別する。 モデルの850,800,700 hPaの気温予測値および 標高を用いて1時間降水種別を補正する。

を用いて PTYP3 を作成する。

では気温と相対湿度の値から、図 4.3.1 の通り降水種別を判別する。図中の赤線で示す境界線はそれぞれ以下の式で表される。

「雨」と「雨か雪」を分ける境界線

$$RH = -\frac{100}{9} \times (T - 10.5)$$

「雨か雪」と「雪か雨」を分ける境界線

$$RH = -\frac{100}{9} \times (T - 9.75)$$

「雪か雨」と「雪」を分ける境界線

$$RH = -\frac{100}{9} \times (T - 8.75)$$

ここで  $\mathrm{RH}$  は地上相対湿度 [%]、T は地上気温  $[^{\circ}\mathrm{C}]$  である。

この境界線は次のように決定している。まず、地上 気象官署の降水種別、地上気温、地上相対湿度の観測<sup>3</sup>

<sup>1</sup> 土田 尚侑

<sup>&</sup>lt;sup>2</sup> 試験運用開始時は北海道・東北・北陸地方の地点が予測対象であった (中務ほか 1997)。その後、段階的に対象地点を増やしていき、白山 (2016) からは積雪深計を設置している全地点を予測対象としている。

<sup>3 2004</sup> 年から 2008 年の 5 年間の冬季 (12 月~翌 3 月) の観測を用いた。ただし、特別地域気象観測所は除く。

表 4.3.1 降水種別ガイダンスの仕様一覧

作成対象	5 km 格子		
作成方法	モデルや格子形式気温ガイダンスの予測値を用いた診断的方法		
予報対象時間単位	3時間		
予報期間と間隔	GSM: FT=6 から FT=84 まで 3 時間間隔 MSM: FT=3 から FT=39 まで 3 時間間隔		
逐次学習の有無	なし		
説明変数	格子形式気温ガイダンス、地上相対湿度、850,800,700 hPa 気温		
層別化処理の対象	なし		
備考	・降水の有無に関わらず降水種別を予測する。 ・850, 800, 700 hPa 気温、標高、補正前の種別に応じて補正(図 4.3.2 参照)。		

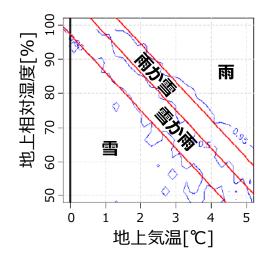


図 4.3.1 観測値から作成した雨の等出現確率線 (青線) および降水種別の境界線 (赤線)。青線は左から雨の出現確率が0.05、0.5、0.95 となる境界線を示し、赤線は左から「雪」と「雪か雨」、「雪か雨」と「雨か雪」、「雨か雪」と「雨」の境界線を示す。

表 4.3.2 降水種別ガイダンスにおける降水種別の定義

降水種別	雨の出現確率
雨	0.95 以上
雨か雪	0.5 以上 0.95 未満
雪か雨	0.05 以上 0.5 未満
雪	0.05 未満

から、雨の出現確率と地上気温・相対湿度との関係を調べた (図 4.3.1 の青線)。ここで雨の出現確率とは、地上気温を 0.2 °C 毎、地上相対湿度を 2.5%毎に区切った領域内に含まれる降水事例数に対する雨の割合である。そして 4 つの降水種別を雨の出現確率に対して表 4.3.2 のように定義し、柳野 (1995) を参考に図 4.3.1 の赤線で示す境界線を決定した。

では、 で判別した降水種別を図 4.3.2 のフローチャートを通して補正する。これは、 の方法では、地上付近に逆転層が形成されて地上付近だけが低温になっ

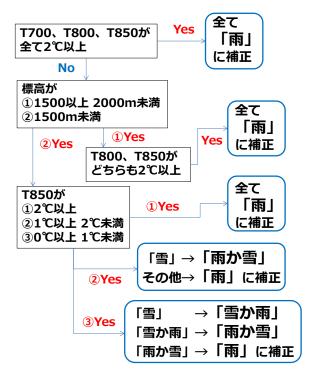


図 4.3.2 1 時間降水種別を補正するフローチャート。T700, T800, T850 はそれぞれ 700, 800, 850 hPa 面の気温、標高はガイダンス用地形データの標高。

ている場合に、「雨」を「雪」などと誤って判別する可能性があるためである。フローチャートでは、 $700,800,850~\mathrm{hPa}$ の気温を用いて補正を行う。標高が  $1500~\mathrm{m}$  未満の場所では、 $850~\mathrm{hPa}$ の気温を用いて降水種別が段階的に変化するように設定している。また、標高  $1500~\mathrm{m}$  以上で不自然に「雪」の格子が孤立して存在するような場合があるため、 $700~\mathrm{hPa}$  や  $800~\mathrm{hPa}$  の気温も用いて補正する。ここで用いる閾値は、各高度の気温と降水種別の出現分布や PTYP の面的な分布等から主観的に決定している。

では、 で求めた1時間毎の降水種別に対して雨 の出現確率に相当する以下の値を割り当てる。

「雨」:1

表 4.3.3 最大降雪量ガイダンスの仕様一覧

作成対象	5 km 格子
作成方法	平均降水量ガイダンスに雪水比をかけて算出する。雪水比は、ロジスティック関数による非線形回帰で決定した回帰式に、格子形式気温ガイダンスを入力して予測する。 1時間毎の降雪量(1時間に線形内挿した3時間平均降水量ガイダンス×1時間雪水比)を計算し、それを積算して3,6,12,24時間降雪量を算出。
予報対象時間単位	3, 6, 12, 24 時間
予報期間と間隔	GSM: 3 時間: FT=6 から FT=84 まで 3 時間間隔 6 時間: FT=9 から FT=84 まで 3 時間間隔 12 時間: FT=15 から FT=84 まで 3 時間間隔 24 時間: FT=27 から FT=84 まで 3 時間間隔 MSM: 3 時間: FT=3 から FT=39 まで 3 時間間隔 6 時間: FT=6 から FT=39 まで 3 時間間隔 12 時間: FT=12 から FT=39 まで 3 時間間隔 24 時間: FT=24 から FT=39 まで 3 時間間隔
逐次学習の有無	なし(入力としている平均降水量ガイダンスはあり)
説明変数	平均降水量ガイダンス、雪水比(格子形式気温ガイダンス)
層別化処理の対象	雪水比(降水量)
備考	降水種別が「雨」の場合または、格子形式気温ガイダンスの地上気温が $+2^{\circ}\mathrm{C}$ 以上の場合には降雪量を $0~\mathrm{cm}$ とする。

「雨か雪」: 0.75 「雪か雨」: 0.25

「雪」:0

これを 3 時間で平均し、平均値が  $0.95\sim1$  なら「雨」  $0.5\sim0.95$  なら「雨か雪」、  $0.05\sim0.5$  なら「雪か雨」  $0\sim0.05$  なら「雪」とする。

#### (2) 最大降雪量ガイダンス

MAXS の仕様を表 4.3.3 に示す。MAXS は、全国の 5 km 格子毎を対象に降雪量を予測する格子形式のガ イダンスであり、入力値に GSM 及び GSM ガイダン スを用いる GSM-MAXS と MSM 及び MSM ガイダン スを用いる MSM-MAXS を作成している。予報要素 は前3時間最大降雪量 (MAXS3)、前6時間最大降雪 量 (MAXS6)、前 12 時間最大降雪量 (MAXS12)、前 24 時間最大降雪量 (MAXS24) である。また、空港地 点向けに前3時間降雪量(空港周辺4格子のMAXS3 の内挿値)と最大前3時間降雪量(空港周辺16格子の MAXS3の最大値とその格子位置)を作成している。降 雪量の予測には、予想される降水量に気温に応じた経 験的な雪水比を掛けて降雪量を見積もる雪水変換法を 使用する。この降水量には平均降水量ガイダンス (第 4.2 節) を、雪水比の作成には格子形式気温ガイダンス を用いている。格子形式気温ガイダンスは 5 km 格子 で予測を行うことから、格子間隔約 20km の GSM に 対しても地形等による詳細な気温分布を考慮した予測 を可能にしている。また、雪水比の計算では降水量に よる層別化を行う。

### 雪水比の作成方法

MAXSの作成方法の説明に入る前に、降雪量の計算に用いる雪水比について説明する。観測値から雪水比を正確に算出するためには、正確な降水量及び降雪量

が必要となる。しかし、降雪量及び降水種別が雪の場 合の降水量の観測は、雪の吹き払いの問題、圧密・沈降 の問題、雨量計の捕捉率の問題等により正確に行うこ とが難しく、算出した雪水比は誤差が大きくなりやす い。MAXS 運用開始前に降雪量予測のために作成して いた雪水比ガイダンスは、こうした誤差が含まれやす い雪水比を目的変数にした MOS 方式 (第 2.2.2 項) の ガイダンスであったため、降雪量の予測精度にも影響 していた。そこで MAXS では、雪水比と関係が深い気 温の観測値から予め回帰式を作成しておき、その回帰 式を用いて予測を行う PPM 方式により雪水比を作成 することにした。予測時の入力値はモデルバイアスを 取り除くために、気温には格子形式気温ガイダンスを、 降水量には平均降水量ガイダンスを利用する。回帰式 は Byun et al. (2008) の方法を採用してロジスティック 関数とし、地上気象官署の観測データ4を用いて3時 間降水量及び降雪量の観測から求めた雪水比を目的変 数、地上気温の観測の前3時間平均値を説明変数とす る非線形回帰により決定した。回帰式は以下のとおり で、SWR は雪水比 [cm/mm]、a, b, c は回帰係数、Tは気温 [°C] である。

$$SWR = \frac{a}{1 + \exp\{(T - b)/c\}}$$
 (4.3.1)

図 4.3.3 に地上気温と雪水比の散布図および当てはめた回帰式を、表 4.3.4 に回帰係数の一覧を示す。雪水比と降水量、地上気温との関係を確認したところ、雪水比は降水量に強く依存していることがわかったため、回帰式は 3 時間降水量による層別化を行っている。3 つの回帰係数は、回帰係数 a は気温がおおよそ -2  $^{\circ}$ C 以下の

 $<sup>^4</sup>$  2004 年から 2008 年の特別地域気象観測所を含む地上気象官署 147 地点の観測データから、天気が雪かつ気温が  $2.5~^{\circ}\mathrm{C}$  以下の事例を抽出した。

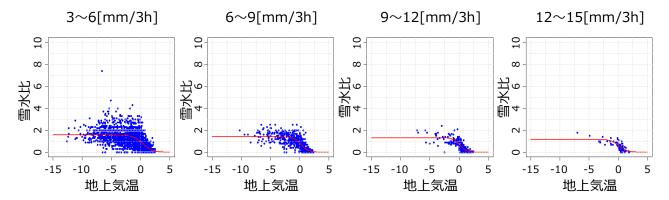


図 4.3.3 地上気温 [°C] と雪水比 [cm/mm] の散布図。左から 3 時間降水量が  $3\sim6$  mm/3h、 $6\sim9$  mm/3h、 $9\sim12$  mm/3h、 $12\sim15$  mm/3h における関係を示す。赤線は当てはめた回帰式を示す。

雪水比、回帰係数bは雪水比がa/2になるときの気温、 回帰係数 c は回帰式の 0°C 前後の傾きを表す。散布図 を見ると、降水量が少ない場合には雪水比のバラつき が大きく、予測時も雪水比の推定誤差が大きくなるこ とが予想される。また、地上気温が0°C前後では回帰 関数の傾きが大きく、地上気温の変化に対する雪水比 の変化が大きい。そのため、0°C 前後での降雪量予測 には地上気温の予測精度が特に重要となる。3時間降水 量が3 mm/3h以下の回帰係数b, cは $3 \sim 6 \text{ mm}/3\text{h}$ の 係数を用いている。これは降水量の測定間隔が 0.5 mm 毎なのに対し、降雪量の測定間隔が1 cm 毎になって いるため、例えば、降水量が $0.5~\mathrm{mm/3h}$ の場合は雪水 比は(0を除けば)必ず2以上の値となり、雪水比を過 大に評価する可能性があることや、回帰係数 b 及び cの値が重要となる地上気温が  $0\sim2$  °C の範囲では雪 水比が 0 に近く、より雪水比を過大に評価する可能性 があることを考慮したためである。

表 4.3.4 回帰式に用いる回帰係数

3 時間降水量 (mm)	<i>a</i>	b	c
$0 \sim 1$	2.250	0.413	0.814
$1 \sim 3$	2.042	0.413	0.814
$3 \sim 6$	1.571	0.413	0.814
$6 \sim 9$	1.396	0.242	0.603
$9 \sim 12$	1.319	0.191	0.571
$12 \sim 15$	1.178	0.157	0.730
$15 \sim$	1.126	0.127	0.495

## 最大降雪量ガイダンスの作成方法

MAXS の作成は次の順に行う。

平均降水量ガイダンスから 1 時間平均降水量を作成する。

の1時間平均降水量および格子形式気温ガイダンスの予測値から1時間雪水比を作成する。

で求めた1時間雪水比と の1時間平均降水量

から 1 時間最大降雪量を作成する。 1 時間最大降雪量を積算することで、MAXS3, MAXS6, MAXS12, MAXS24 を作成する。

では、3 時間平均降水量ガイダンス (MRR3) から 5 km 格子の 1 時間平均降水量を作成する。GSM 最大 降雪量ガイダンスでは、20 km 格子の MRR3 を時間 方向に 3 等分した後に、5 km 格子に線形内挿を行う。 MSM 最大降雪量ガイダンスでは、5 km 格子の MRR3 を時間方向に 3 等分した後に、格子ごとの予測のばら つきを抑えるために周囲 9 格子の平均値とする。

では、 で求めた 1 時間平均降水量と 5 km 格子の格子形式気温ガイダンスの 1 時間値を入力として、雪水比の作成方法で述べた (4.3.1) 式を用いて 5 km 格子の 1 時間雪水比を作成する。回帰式は 3 時間降水量により層別化されているため、1 時間平均降水量を 3 倍した値に対応する回帰式を選択して用いる。なお、降水種別ガイダンスとの整合をとるため、降水種別が「雨」の格子では雪水比を 0 とする。また、気温が $+2^{\circ}$ C 以上の場合も雪水比を 0 とする。

では、 で求めた 1 時間雪水比に で求めた 1 時間平均降水量を乗じて、 $5~\rm km$  格子の 1 時間最大降雪量を作成する。なお、降雪量が  $0.1~\rm cm/h$  以下のような弱い降雪域の広がりを防ぐために、 $1~\rm th$ 間平均降水量が  $0.02~\rm mm/h$  未満の場合には  $1~\rm th$ 間最大降雪量を  $0~\rm cm$  とする。

最後に で、 で作成した1時間最大降雪量を積算して、MAXS3, MAXS6, MAXS12, MAXS24を作成する。

#### (3) 降雪量地点ガイダンス

SNOW の仕様を表 4.3.5 に示す。SNOW は、統計手法にニューラルネットワークと頻度バイアス補正を用いて、積雪深計が設置されたアメダス地点を対象に降雪量を予測する地点形式のガイダンスである。モデル入力値に GSM を用いる GSM-SNOW と MSM を用い

表 4.3.5 降雪量地点ガイダンスの仕様一覧

作成対象	積雪深計が設置されているアメダス地点
作成方法	ニューラルネットワーク (3層:中間、出力とも活性化関数にロジスティック関数を使用)による予測に頻度バイアス補正を行う。
予報対象時間単位	6, 12, 24 時間
予報期間と間隔	GSM:6時間:FT=9からFT=84まで3時間間隔 12時間:FT=15からFT=84まで3時間間隔 24時間:FT=27からFT=84まで3時間間隔 MSM:6時間:FT=6からFT=39まで3時間間隔 12時間:FT=12からFT=39まで3時間間隔 24時間:FT=24からFT=39まで3時間間隔
逐次学習の有無	あり(ニューラルネットワーク、頻度バイアス補正)
説明変数	地上気温、925, 850, 700, 500 hPa の気温、地上相対湿度、925, 850, 700 hPa の相対湿度、700, 500 hPa の高度、925, 850, 700, 500 hPa の風向、925, 850, 700, 500 hPa の風速、「海面水温 – 925 hPa の気温」、925, 850, 700 hPa の上昇流、「地上~850 hPa・925~700 hPa の SSI」、地形性降水指数、降水量、地上気圧、「気温で層別化した雪水比にモデル降水量を乗じた降雪量」
目的変数	アメダスの積雪深計で観測された $6,12$ 時間降雪量
層別化処理の対象	作成対象地点、初期時刻、予報時間(3時間毎1日分)
備考	頻度バイアス補正の閾値は $6$ 時間降雪量が $1,5,15,30~\mathrm{cm}/6\mathrm{h}$ 、 $12$ 時間降雪量が $3,10,20,40~\mathrm{cm}/12\mathrm{h}$ を使用。 $24$ 時間降雪量は $12$ 時間降雪量を積算して算出。

る MSM-SNOW を作成している。予報要素は前 6 時間降雪量 (SNOW6)、前 12 時間降雪量 (SNOW12)、前 24 時間降雪量 (SNOW24)である。ニューラルネットワークと頻度バイアス補正はともに逐次学習を行う。また、予測対象地点、初期時刻、予報時間により層別化を行っている。以下に作成方法、ニューラルネットワークの説明変数と層別化について説明する。

## 作成方法

SNOW の作成は次の順に行う。

表 4.3.5 に示した説明変数を入力としたニューラルネットワークにより、6 時間降雪量および 12 時間降雪量を予測する。

で求めた予測値に頻度バイアス補正を行う。

で補正を行った予測値に対し、降水量と気温の予測を用いた整合処理を行って SNOW6, SNOW12 とする。

で求めた SNOW12 を積算することで、SNOW24 を作成する。

では、数値予報モデルの予測値を入力としてニューラルネットワークによる予測を行う。ネットワークの構成は、入力層 29 ユニット、中間層 16 ユニット、出力層 1 ユニットの 3 層順伝播型ニューラルネットワークである。表 4.3.5 に示した説明変数を入力とし、中間層および出力層の活性化関数にはロジスティック関数を用いている。出力層の要素である目的変数はアメダスの積雪深計で観測された 6 時間降雪量および 12 時間降雪量である。

では、 で得られた予測値に対して頻度バイアス 補正を行う。ニューラルネットワークは統計手法のた め、発生頻度の高い現象に最適化されやすい。そのため、発生頻度の低い大雪の予測頻度は低くなり、精度も十分でない。頻度バイアス補正はこれを解消することを目的としている (第 2.9 節)。頻度バイアス補正の閾値は、6 時間降雪量では 1,5,15,30 cm/6h、12 時間降雪量では 3,10,20,40 cm/12h を設定している。

では、モデルの降水量と気温ガイダンスの予測値を用いて降水量が非常に少ない場合( $0.05~\mathrm{mm}$  以下)や気温が高い場合( $5.5~\mathrm{C}$  以上)の降雪量を $0~\mathrm{cm}$  とする整合処理を行う。少ない降水量に対しては北海道東部及び冬型時の関東地方に対する空振りを、高い気温に対しては春先の予測降水量の多い事例での空振りを減らす効果がある。気温による処理では、 $5.5~\mathrm{C}$  前後で突然予測が変化することを防ぐため、 $4.5~\mathrm{C}\sim5.5~\mathrm{C}$  の間で $4.5~\mathrm{C}$  で の予測値、 $5.5~\mathrm{C}$  で  $0~\mathrm{cm}$  となるような線形内挿を行っている。

#### 説明変数

説明変数は、予測地点の周囲 4 格子のモデル予測値を内挿した値を用いる。ただし、「海面水温 -925 hPaの気温」は、予報地点から西、北方向にそれぞれ 2 度ずつずらした格子点を使用し、日本海側における冬型の気圧配置時の海面からの顕熱輸送を考慮する 5。説明変数のうち、SSI (ショワルターの安定指数)は大気の成層安定度を、地形性降水指数は、925 hPaの風向を8方位で表したときにその風向に対応する地形の傾き(風下に向かって標高が高くなっているのか、低くなっているのか)を表す。また、降雪量は雪水比とモデルの降水量から計算する。この雪水比はモデルの地

<sup>&</sup>lt;sup>5</sup> 利用する格子点が陸地の場合には、海面水温を内挿した値 を用いる。

上気温で層別化した定数で、MAXS に用いる雪水比と は異なる。

#### 層別化

ニューラルネットワークの重み係数は、予測対象地 点、初期時刻及び予報時間で層別化している。予報時間 による層別化は、6時間降雪量はGSM-SNOW(MSM-SNOW) では  $FT=9 \sim 30 \ (6 \sim 27)$  を 3 時間毎に 1 日 分、12 時間降雪量は GSM-SNOW (MSM-SNOW) で は  $FT=15\sim36~(12\sim33)$  を 3 時間毎に 1 日分を対象 に行い、それ以降の FT は対象時刻毎にこれらの係数 を繰り返し使用する。これは第2.2.2項で述べられた pseudo-PP と呼ばれる手法で、先の予測時刻に対して も目先24時間と同程度のメリハリを持った降雪量予測 を利用することができる。ただし第2.2.2項でも述べら れている通り、予報時間によってモデルの系統誤差が 異なる場合、先の予測ではニューラルネットワークに よる系統誤差の修正が十分でない可能性がある。目先 24 時間以降の予測値は各地域におけるモデルの系統誤 差を把握した上で利用することが望ましい。

#### 4.3.3 統計検証

本項では、GSM, MSM ガイダンスの統計的な検証結果を示す。検証期間は 2016 年 12 月 ~ 2017 年 2 月である。数値予報モデルには、2018 年 3 月現在で最新のGSM (米原 2017) 及び MSM (原 2017) を用いた。

## (1) 降水種別ガイダンス

ここでは、全国の地上気象官署 (特別地域気象観測所 を含む) における天気の観測を降水種別とした PTYP3 の検証結果を示す。地上気温の観測値が $-3\sim5$  °C か つ降水(みぞれ・あられ等は除く)が観測されている場合 (天気が霧雨、地雨、しゅう雨、地雪、しゅう雪の場合) のみを検証対象とし、降水種別を「雨」と「雪」の2カ テゴリに分類して実施した。実況の降水種別は、天気が 霧雨、地雨、しゅう雨を「雨」、地雪、しゅう雪を「雪」 とし、予報の降水種別は、「雨か雪」を「雨」、「雪か雨」 を「雪」として扱う。検証地点の予報値には最近接格 子の値を用いた。また、検証に用いた初期値および予 報時間は、GSM-PTYP (MSM-PTYP) が00,06,12, 18UTC (03, 09, 15, 21UTC) 初期値の FT=15  $\sim 36$  $(12 \sim 33)$  である。  $2 \times 3.4$  および $2 \times 3.5$  に PTYP3 の 地域別の雪予報のエクイタブルスレットスコア (ETS) およびバイアススコア (BI) を示す。ETS を比較する と、概ね MSM-PTYP が GSM-PTYP を上回ってお り、MSM-PTYP の方が予測精度が高いことがわかる。 また、BI についても、MSM-PTYP は GSM-PTYP と 比較して全国的に BI=1 に近く、適切な頻度で雪を予 測していることがわかる。一方で GSM-PTYP は西日 本を中心にBIが1より低く、雪の見逃しが多い。この 原因の一つには、西日本を中心に GSM のモデル湿度 に正バイアスがあることが考えられる (後藤 2017a)。

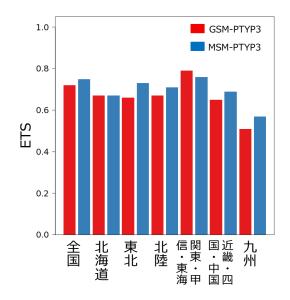


図 4.3.4 PTYP3 の各モデルの地域別 ETS の比較。検証期間は 2016 年 12 月~2017 年 2 月。赤は GSM-PTYP、青は MSM-PTYP で、それぞれ左から全国、北海道、東北、北陸、関東・甲信・東海、近畿・中国・四国、九州の順に地域ごとの検証結果を示す。

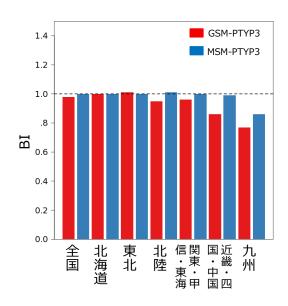


図 4.3.5 PTYP3 の各モデルの地域別 BI の比較。検証期間 や図の設定は図 4.3.4 と同じ。

## (2) 最大降雪量ガイダンス

ここでは、全国のアメダス地点で観測された積雪深を観測値とした MAXS6, MAXS12, MAXS24 の検証結果を示す。検証地点の予測値は、周囲の 4 格子点を線形内挿することにより求めた。検証対象の初期値はGSM-MAXS (MSM-MAXS) が 00,06,12,18UTC (03,09,15,21UTC) である。予報時間は、MAXS6 はGSM-MAXS (MSM-MAXS) が  $FT=9\sim30$  (6  $\sim$ 

27)、MAXS12はGSM-MAXS (MSM-MAXS)が  ${
m FT}{=}15\sim36$  (  $12\sim33$  ) 、 MAXS24 は GSM-MAXS (MSM-MAXS) が  $FT=27\sim36$  (24  $\sim33$ ) を対象と した。図 4.3.6 に MAXS6, MAXS12, MAXS24 の ETS および BI を示す。まず GSM 及び MSM-MAXS を全 体的に比較すると、ETS の値は両ガイダンスの間に大 きな差はないが、全予報要素でわずかに MSM-MAXS の方が大きな値となっており、予測精度が高いことがわ かる。BI は MSM-MAXS の方が高い値となっており、 どの閾値に対しても GSM-MAXS よりも予測頻度は多 い。次に閾値毎に確認する。両ガイダンスとも MAXS6 で 10 cm/6h、MAXS12 で 20 cm/12h 程度以下の降雪 ではBIは1に近く、適切な頻度で予測できている。し かし、MAXS6 で 20 cm/6h、MAXS12 で 40 cm/12h を超えるような大雪は BI が 1 を下回っており、予測頻 度は過小である。また、ETS の値は小さく、予測精度 も低くなっている。 一方で MAXS24 は 20 cm/24h 以 上の閾値で BI が 1 より大きくなっており、大雪の予測 頻度が多くなっている。ただし、MAXS24 は予測対象 とする期間が長いため、圧密等による新積雪の沈降に よる影響も大きい (第 4.3.4 項 (4))。

## (3) 降雪量地点ガイダンス

ここでは、全国のアメダス地点で観測された積雪深に よる SNOW6, SNOW12, SNOW24 の検証結果を示す。 検証対象の初期値、予報時間は GSM (MSM) -SNOW は GSM (MSM) -MAXS とそれぞれ同様である。図 4.3.7 に SNOW6, SNOW12, SNOW24 の ETS および BI を示す。また、MAXS との比較を行うため、図 4.3.8 に GSM-MAXS と GSM-SNOW の検証結果の比較図 を、図 4.3.9 に MSM-MAXS と MSM-SNOW の検証 結果の比較図を示した。ETS の比較結果から、GSM 及び MSM-SNOW の予測精度に大きな差はないが、 15 cm/12h から 25 cm/12h 程度の閾値で見た場合は わずかに GSM-SNOW の方が予測精度が高いことが わかる。次に BI を確認する。頻度バイアス補正を行っ ている 6 時間降雪量と 12 時間降雪量は 10 cm/12h 以 下の降雪から 40 cm/12h を超える大雪まで、BI が 1 に近い値に保たれており、適切な頻度で予測できてい る。MAXS との比較図で確認すると、MAXS では予測 過小の傾向がある大雪についても SNOW では適切な 頻度の予測が可能であることがわかる。また、ETS は 全閾値で MAXS を上回っており、積雪深計が設置され たアメダス地点での予測には SNOW を採用したほう が精度が良いことがわかる。一方でSNOW24のBIは MAXS24 と同様に大きな降雪量閾値ほど1よりも大き な値となっており、大雪の予測頻度が過大な傾向が見 られる。ただし、MAXS24 と比較すると ETS の値は 大きく、予測精度は高い。SNOW24 は SNOW12 の積 算値であるため、頻度バイアス補正の効果は限定的で ある。MAXS24と同様に新積雪の沈降の影響があると 考えられる。また、開発時の検証では SNOW24 の予測頻度が過大な傾向は見られず、SNOW6 や SNOW12 で予測過小な傾向があった (白山 2016)。モデル改良あるいは経年変動により予測頻度が変化している可能性も考えられる。

## 各降雪量ガイダンスの予測特性の違い

ここで、SNOW と MAXS の検証結果を地点毎に比 較することで、両ガイダンスの地域ごとの予測特性に ついて述べる。図 4.3.10 に GSM-SNOW 及び MAXS の 20 cm/24h に対する ETS と BI のスコアマップ、図 4.3.11 に MSM-SNOW 及び MAXS の 20 cm/24h に対 する ETS と BI のスコアマップを示す。まず ETS を確 認すると、GSM ガイダンスでは SNOW, MAXS とも に北海道東部と東北南部から山陰にかけての日本海側 の山間部での予測精度が高く、さらに SNOW では日本 海側の沿岸部でも予測精度が高くなっている。差分図 からは全国的に SNOW の方が予測精度が高いことが わかる。MSM ガイダンスについても GSM ガイダンス と同様の傾向である。次に BI を確認すると、SNOW は GSM、MSM ともに北海道、東北北部や北陸の沿岸 部で予測頻度が高い。一方、MAXS は GSM、MSM と もに北海道から北陸にかけての日本海側で予測頻度が 高いが、そのほかの地域では予測頻度が低い地点が目 立つ。このように、SNOW、MAXS ともに、全地点の 検証結果では24時間降雪量は大雪の予測頻度が高い 結果であったが、その分布は両者の間で異なり、また、 MAXS は予測頻度が過小な地域もあることに注意が必 要である。第4.3.4項(2)では、山陰地方の沿岸部での 予測過小事例を示している。

## 4.3.4 事例検証

本項では、いくつかの事例を通して降雪ガイダンス の予測特性や利用上の留意点を示す。

## (1) 冬型の予測事例

2017年12月12日から13日にかけて、日本の東に 向かって発達しながら進んだ低気圧の影響で日本付近 は強い冬型の気圧配置となり、北陸から東北地方を中 心に大雪となった。図 4.3.12 は、2017 年 12 月 12 日 12UTC を対象とした 24 時間降雪量の予測である。ま ず、SNOW の予測を見ると、GSM, MSM の両ガイダ ンスとも量的に予測がやや過大となっている地点もあ るものの、30~40 cm/24h を超える大雪の分布をよく 予測できている。ただし、実況で $50\sim80 \text{ cm}/24\text{h}$ を観 測した新潟県から福島県にかけては、MSM-SNOW は 比較的よく予測できているが、GSM-SNOW は新潟県 側で予測過大な地点が多く、福島県側では予測過小な 地点が多い。これは、新潟県から福島県にかけてのシ アライン上の降水を MSM がよく予測していた一方で、 GSM の降水予測は新潟県側に偏っていたためと考えら れる。次に MAXS であるが、GSM, MSM の両ガイダ

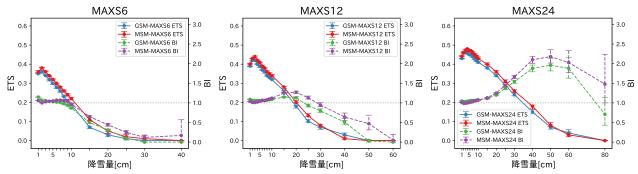


図 4.3.6 MAXS の降雪量閾値別の ETS および BI。検証期間は 2016 年 12 月 ~ 2017 年 2 月。左が 6 時間降雪量、中が 12 時間降雪量、右が 24 時間降雪量を対象とした検証で、青線が GSM-MAXS の ETS、赤線が MSM-MAXS の ETS で、緑線が GSM-MAXS の BI、紫線が MSM-MAXS の BI である。エラーバーは 95%信頼区間を表す。

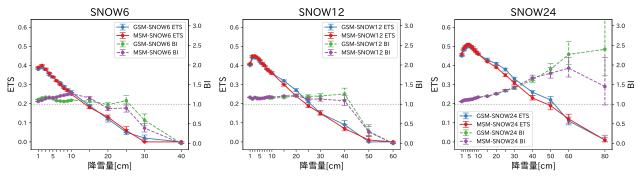


図  $4.3.7~{
m SNOW}$  の降雪量閾値別の ETS および BI。検証期間や図の設定は図  $4.3.6~{
m と同じであるが、}$  それぞれ SNOW を対象 としている。

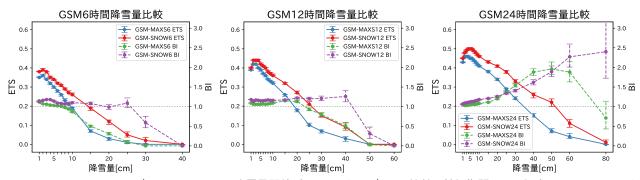


図 4.3.8 GSM-MAXS 及び GSM-SNOW の降雪量閾値別の ETS および BI の比較。検証期間や図の設定は図 4.3.6 と同じだが、青線が GSM-MAXS の ETS、赤線が GSM-SNOW の ETS で、緑線が GSM-MAXS の BI、紫線が GSM-SNOW の BI である。

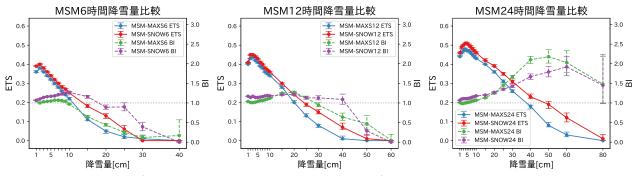


図 4.3.9~ MSM-MAXS 及び MSM-SNOW の降雪量閾値別の ETS および BI の比較。検証期間や図の設定は図 4.3.8~と同じであるが、それぞれ MSM ガイダンスを対象としている。

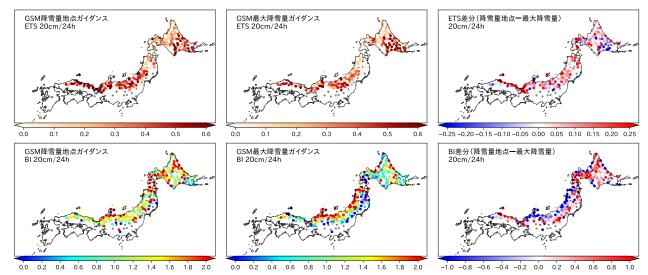


図 4.3.10 GSM-SNOW 及び GSM-MAXS の ETS および BI のスコアマップ。24 時間降雪量の閾値 20 cm に対する検証結果を、左から SNOW24、MAXS24、両者の差分図の順に示す。上段が ETS、下段が BI である。灰色の地点は、検証期間中に予報、観測ともに閾値を超える事例がない地点を表す。

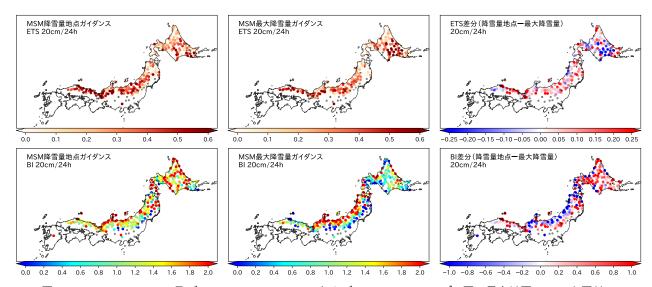


図 4.3.11 MSM-SNOW 及び MSM-MAXS の ETS および BI のスコアマップ。図の見方は図 4.3.10 と同じ。

ンスとも実況よりも大雪の予測範囲が広く、また、量的にも SNOW よりも予測過大となっている。その一方で、福島県の若松など日本海側以外の地域では SNOW と比較して予測過小となっている。統計検証で示した通り、SNOW と MAXS はともに日本海側で予測過大な傾向があるものの、量的な予測精度は SNOW の方が高い。また、MAXS は日本海側以外では大雪に対して予測過小となっているなど、予測傾向が一貫していない。量的な予測では、より精度の高い SNOW を用いて MAXS の降雪量を修正することが有効である。

## (2) 山陰地方の大雪事例

2017 年 2 月 10 日から 11 日にかけて、西日本の日本海側で大雪となった事例を紹介する。この事例では、

山陰地方に向かって JPCZ<sup>6</sup> が形成され、鳥取では 10日の日降雪量が 65 cm に達するなど、西日本の日本海側で大雪となった。図 4.3.13 に、2017年2月10日12UTC を対象とした24時間降雪量の予測を示す。山陰地方について、SNOW はアメダスによる観測と同程度の大雪を予測できていることがわかる。ただし、実際よりも大雪の中心を西寄りに予測しており、松江や米子では予測過大となっている。これは、モデルの降水予測が実際よりも西寄りに予測されていたためと考えられる。一方、MAXS は山間部を中心とした予測となっており、沿岸部の地点の降雪量予測は実際よりも少ない。山陰地方の沿岸部は北陸や東北と比べて海面

 $<sup>^6\,</sup>$  日本海寒帯気団収束帯 (Japan sea Polar air mass Convergence Zone)。

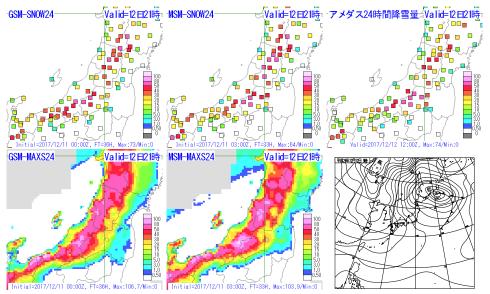


図 4.3.12 2017 年 12 月 12 日 12UTC を対象とした 24 時間降雪量の予測事例。上段左から GSM-SNOW、MSM-SNOW、アメダスによる前 1 時間積雪深差の 24 時間積算値、下段左から GSM-MAXS、MSM-MAXS、12 日 00UTC の速報天気図。降雪量の単位は [cm/24h]。

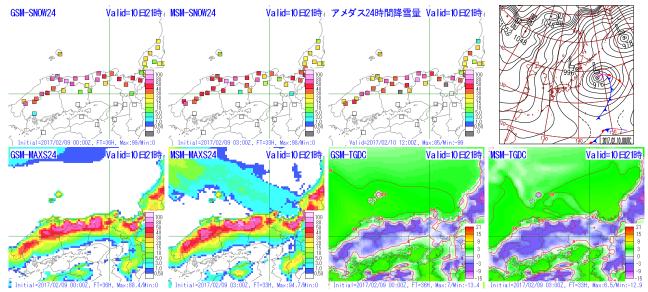


図 4.3.13 2017 年 2 月 10 日 12UTC を対象とした 24 時間降雪量の予測事例。上段左から GSM-SNOW、MSM-SNOW、アメダスによる前 1 時間積雪深差の 24 時間積算値(灰色の地点は期間中に積雪深に欠測を含むことを表す。)、10 日 00UTC の日本域地上天気図、下段左から GSM-MAXS、MSM-MAXS、GSM 格子形式気温ガイダンス、MSM 格子形式気温ガイダンス。降雪量の単位は [cm/24h]、気温の単位は  $[^{\circ}C]$ 。

水温が高く、モデルは沿岸の気温を十分に解像できていないことから地上気温予測も高めの値となる。地上気温が0°C付近では雪水比は急激に減少するため、山陰地方の沿岸部では降雪量が少なく予測されたと考えられる。

#### (3) 南岸低気圧による大雪事例

図 4.3.14 に 2016 年 1 月 18 日 00UTC を対象とした 12 時間降雪量の予測事例を示す。この事例では、本州の南岸を低気圧が通過し、東京でも 18 日 06JST に日

最深積雪となる積雪深  $6~\mathrm{cm}$  を観測するなど関東甲信で大雪となった。このような事例で降雪を精度よく予測するには、気温の予測精度が重要となる。特に、雪水比の計算に診断的手法を用いている MAXS の予測は、気温の予測に大きく左右される。実況では東京で $6~\mathrm{cm}$ 、横浜で $5~\mathrm{cm}$  が観測されたが、MAXS は降雪を予測していない。茨城から神奈川の沿岸部では、予測対象期間の PTYP は「雨」が予測されており、また、気温ガイダンスの予測も $2~\mathrm{cm}$  以上が予測されていたため、雪水比が $0~\mathrm{cm}$  となり、MAXS は沿岸部の降雪を予

測できなかった。一方で、SNOW は沿岸部の地点にも降雪を予測している。SNOW は、NN で地上気温以外にも多くの説明変数を用いて予測を行うため、地上気温以外の説明変数を用いて降雪の可能性を示すことができたと考えられる。

## (4) 大雪が続いた事例

図 4.3.15 は、2017 年 1 月 11 日から 16 日頃にかけて、 北陸地方を中心に大雪が続いた事例である。図の上段 には降雪期間の前半である 2017 年 1 月 13 日 18UTC を対象とした24時間降雪量の予測、下段には降雪期 間の後半である 2017 年 1 月 16 日 12UTC を対象とし た前24時間降雪量の予測事例を示す。どちらの対象時 刻に対しても、降雪量地点ガイダンス、最大降雪量ガ イダンスはともに 50 cm/24h を超える降雪を予測して いる。一方、アメダス 24 時間降雪量は 13 日 18UTC では新潟県を中心に50 cm 以上が観測されているもの の、16 日 12UTC では多くの地点で 20 cm 程度で、16 日 12UTC に対する予測はかなり過大であるように見 える。ここで、図 4.3.16 に本事例の期間中のアメダス 十日町における降水量と積雪深観測値を示すが、期間 中は継続して日降水量 40 mm/24h 程度以上が観測さ れており、実際には比較的多くの降雪が継続していた と考えられる。ところが、日最深積雪は14日以降増加 していない。多量の降雪が継続したときに積雪深が増 加しなくなる理由として、積雪が自身の重さで圧縮さ れる圧密の効果があげられる。降り積もった新雪は圧 密により圧縮され、積雪後数日間は急速に積雪深を減 少させる。アメダス十日町の場合、14日以降は圧密の 効果により積雪深の増加が抑えられていたと考えられ る。したがって、図 4.3.15 の 16 日 12UTC の積雪深計 による24時間降雪量は実際に降った雪の量と比べると かなり過少であると考えられる。

大雪が数日間継続するような事例では、事象の後半になると降雪初期に降り積もった雪が圧密効果により 圧縮されていくため、新雪が降り積もっても積雪深は あまり増えなくなる。このため、1時間毎の積雪深差 として見積もっている実況の降雪量は実際に降った雪 よりも少なくなり、ガイダンスの降雪量との差が大き くなることに留意が必要である。

## 4.3.5 利用上の留意点と今後の展望

本項では、降雪ガイダンスの利用上の留意点と今後 の展望を紹介する。まず、利用上の注意点を述べる。

PTYP は統計検証で GSM-PTYP よりも MSM-PTYP の方が予測精度が高いことを示した。特に西日本では GSM-PTYP は雪の予測頻度が低く、見逃しが多い。第 4.3.4 項 (2) で示した通り西日本の沿岸部は予測気温が高いため、気温の僅かな予測誤差で降水種別が「雨か雪」や「雨」と判別されてしまい、予測が難しい。西日本の沿岸部ではこのような特性に注意していただきたい。

SNOW 及び MAXS は統計検証で GSM ガイダンス と MSM ガイダンスの間で予測精度には大きな差がないことを示した。ただし、事例検証で示したように、モデルの降水や気温の予測精度が降雪量予測に影響するため、実況を踏まえてより適切な降水・気温の予測を行っているモデルを選択して利用いただきたい。

SNOW と MAXS 相互に比較した統計検証では、SNOW の方が予測精度が高いことを示した。面的な予測を MAXS で把握して、量的な予測は SNOW で修正するといった利用方法が有効であると考えられる。ただし、MAXS は降水量ガイダンスと格子形式気温ガイダンスとの整合が取れた予測となっているのに対し、SNOW は多くの要素を説明変数とした統計手法を用いているためモデル予測との対応が分かり難いこともある。また、予測対象地点では SNOW の方が予測精度が高いが、その近傍でアメダス地点との降雪量に違いがあるような地域では MAXS の予測が適切な場合もあると考えられるため、両ガイダンスの予測特性や予測対象地点の地域特性を把握した上で利用していただきたい。

24 時間降雪量の BI は SNOW, MAXS 共に 1 より大きいことを統計検証で示した。これは第 4.3.4 項 (4) で述べた通り実況の降雪量が実際に降った雪よりも少なくなること、そして第 4.3.3 項で述べたようにガイダンスが新積雪の沈降を考慮していないことが原因である。防災の観点からは降った雪の総量が重要であると考えており、現時点では降雪ガイダンスで新積雪の沈降について考慮する予定はない。このため、今後の統計検証でも 24 時間降雪量の BI は 1 より大きい状況が続くことに留意していただきたい。

最後に今後の展望を述べる。第 4.3.2 項でも述べたように、降水種別ガイダンスは地上気温と地上相対湿度から作成するが、地上気温の入力には格子形式の気温ガイダンスを用いているものの、地上相対湿度はモデルの予測値を入力としている。統計的な手法でモデルのバイアスを低減した格子形式の湿度ガイダンスを開発すれば、降水種別ガイダンスの精度向上が見込める。また、ここまで述べた通り北陸地方や山陰地方、太平洋側でそれぞれ降雪の特性は異なると考えられる。雪水比の計算に地域層別を導入できれば地域特性に応じた降雪量予測が可能となり、精度向上が見込める。さらに、本運用に向けた準備が進んでいるメソアンサンブルシステム (MEPS) を入力としたガイダンスの開発についても検討する必要がある。

## 参考文献

Byun, K.-Y., J. Yang, and T.-Y. Lee, 2008: A Snow-Ratio Equation and Its Application to Numerical Snowfall Prediction. *Wea. Forecasting*, **23**, 644–658.

後藤尚親, 2017: GSM 改良のガイダンスへの影響. 平成

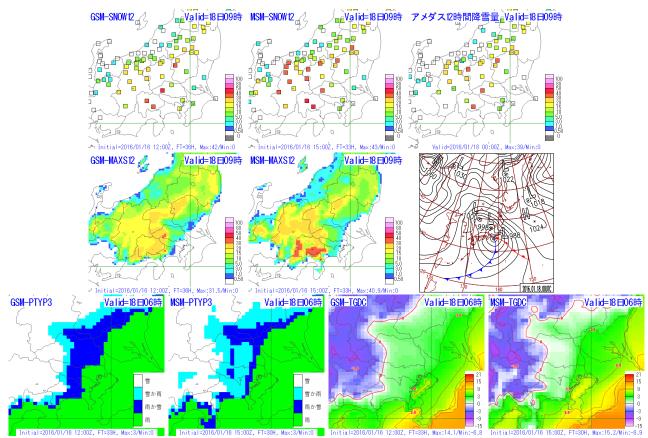


図 4.3.14 2016 年 1 月 18 日 00UTC を対象とした 12 時間降雪量の予測事例。上段左から GSM–SNOW、MSM–SNOW、アメダスによる前 1 時間積雪深差の 12 時間積算値、中段左から GSM–MAXS、MSM–MAXS、18 日 00UTC の日本域地上天気図、下段左から GSM–PTYP、MSM–PTYP、GSM 格子形式気温ガイダンス、MSM 格子形式気温ガイダンス(それぞれ前 3 時間値)。降雪量の単位は [cm/12h]、気温の単位は  $[^{\circ}C]$ 。

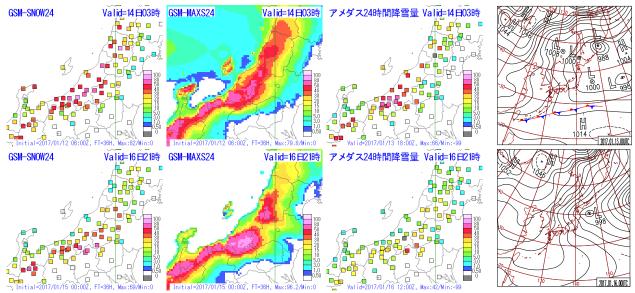


図 4.3.15 上段は 2017 年 1 月 13 日 18UTC、下段は 2017 年 1 月 16 日 12UTC を対象とした 24 時間降雪量の予測事例。左から GSM-SNOW、GSM-MAXS、アメダスによる前 1 時間積雪深差の 24 時間積算値(灰色の地点は期間中に積雪深に欠測を含むことを表す。) 日本域地上天気図。日本域地上天気図は上段が 13 日 00UTC、下段が 16 日 00UTC。降雪量の単位は  $[\mathrm{cm}/24\mathrm{h}]$ 。

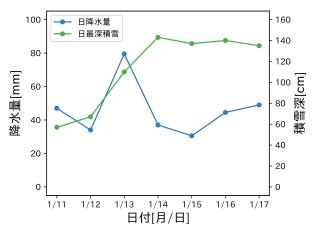


図 4.3.16 アメダス十日町における 2017 年 1 月 11 日から 17 日までの日降水量および日最深積雪。青線は日降水量 [mm]、緑線は日最深積雪 [cm]。

29 年度数値予報研修テキスト, 気象庁予報部, 31-34. 原旅人, 2017: メソ数値予報システムの改良の概要. 平成29 年度数値予報研修テキスト, 気象庁予報部, 42-47. 古市豊, 2008: 降雪量ガイダンス. 平成20 年度数値予報研修テキスト, 気象庁予報部, 73-76.

古市豊, 松澤直也, 2009: 最大降雪量ガイダンス. 平成 21年度数値予報研修テキスト, 気象庁予報部, 27-38. 中務信一, 成井昭夫, 原昌弘, 1997: 降雪量予報に向け ての検討. 平成 9年度量的予報研修テキスト, 気象庁 予報部, 28-38.

白山洋平, 2016: 降雪量地点ガイダンスの改良と開発. 平成 28 年度数値予報研修テキスト, 気象庁予報部,

柳野健, 1995: ニューラルネットによるガイダンス. 平成7年度量的予報研修テキスト, 気象庁予報部, 54-69. 予報部予報課, 2011: 市町村を対象とした警報・注意報発表のための新しい予報作業支援システムの整備について. 測候時報, 78, 1-18.

米原仁, 2017: 全球数値予報システムの改良の概要. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 1-7.

#### 4.4 気温ガイダンス 1

## 4.4.1 はじめに

気温ガイダンスは地点形式と格子形式で地上気温を 予測するガイダンスである。地点形式の気温ガイダン スにはアメダス地点と空港地点を対象とした最高・最 低気温ガイダンスと時系列気温ガイダンスがあり、天 気予報と飛行場予報の作成に利用されている。格子形 式の気温ガイダンスは 5 km 格子を予測対象とし、最 大降雪量ガイダンスと降水種別ガイダンスの入力デー タとして利用されている。

天気予報用の気温ガイダンスは1984年に運用を開始 した (気象庁予報部 1986)。 運用開始当初は全国 60 地 点の最高・最低気温を予測対象とし、作成には線形重 回帰を用いていた。説明変数は下層から中層の気温や 高度、風などであった。その後、使用する説明変数の見 直し、予測対象地点の増加などの改良や、時系列気温 予測の開発・運用を実施してきた。1996年3月にはモ デル更新への柔軟な対応を目指して、カルマンフィル タの利用を開始した (瀬上ほか 1995)。 カルマンフィル タ導入当初の説明変数は、地上気温と平年値の差、中 下層雲量、東西風、南北風であったが、2003, 2014年 に説明変数の見直しを行った(松本・海老原 2003;後 藤 2014)。入力となる数値予報モデルとしては、1996 年以降、領域モデル (RSM) を利用していたが、2001 年からは GSM に基づく気温ガイダンスの運用も開始 した (海老原ほか 2000)。2007 年 11 月には RSM の廃 止と GSM の格子間隔 20 km への高解像度化に伴い、 RSM ガイダンスが廃止となり、GSM ガイダンスに一 本化された。また、2007 年 5 月の MSM の予報時間延 長に伴い、MSM 気温ガイダンスが運用開始となった (阿部 2007)。

空港の気温を予測する航空気温ガイダンスは、1993年に最高・最低気温を予測対象として運用開始し (田崎・此田 1994)、1996年に天気予報用ガイダンスと同様に作成方法を線形重回帰からカルマンフィルタに変更した。2006年には、時系列気温を予測対象に追加するとともに説明変数の見直しや係数の変化速度を決定するパラメータの調整などを行った (新美 2005)。

格子形式気温ガイダンスは、降雪予測においてモデルバイアスを除去した面的な気温予測が必要であったことから、2009年に導入され(古市・松澤 2009)、2010, 2016年に計算方法を見直した(古市 2010; 黒木 2017)。

本節ではまず第 4.4.2 項で気温ガイダンスの仕様と作成方法を述べ、続いて第 4.4.3 項と第 4.4.4 項で統計 および事例検証を通じて得られた精度や特性を述べる。最後に第 4.4.5 項で利用上の留意点と今後の展望を述べる。なお本節において各気温ガイダンスの名称は、GSM (MSM) 時系列気温ガイダンスなどと各気温ガイ

ダンスの名称の前に数値予報モデル名を付加して表記する。航空気温ガイダンスは、利用する数値予報モデルや作成方法が MSM による天気予報用ガイダンスとほぼ同様であり、両者の仕様の違いは目的変数における観測種別(アメダスか METAR 等か)以外にない。予測精度や予測特性についても大きな違いはないため、航空気温ガイダンスの解説は割愛し、MSM 気温ガイダンスで代用する。

## 4.4.2 仕様

表 4.4.1 に最高・最低気温ガイダンス、表 4.4.2 に時系列気温ガイダンス、表 4.4.3 に格子形式気温ガイダンスの仕様を示す。なお、本項における仕様の説明では GSM, MSM をまとめて「モデル」と記載する。

#### (1) 最高、最低および時系列気温ガイダンス

最高気温ガイダンスは  $00\sim09$ UTC における最高気温、最低気温ガイダンスは  $15\sim00$ UTC における最低気温、時系列気温ガイダンスは毎正時の気温を予測するガイダンスで、GSM, MSM を入力としている。最高、最低および時系列気温ガイダンスにおいて、目的変数は観測値とモデル地上気温予測値の差である。用いる観測値は、天気予報用の気温ガイダンスの場合はアメダス地点の観測値(1分値)、航空気温ガイダンスの場合は METAR, METAR AUTO および SPECI<sup>2</sup> である。

## 説明変数

最高、最低および時系列気温ガイダンスの説明変数は数値予報モデルの、地上気温、風速、東西南北の風向4成分、気温減率、前3時間降水量、前日との気温差、中・下層雲量である。

地上気温は、モデルの系統誤差が地上気温の高低で異なることに対応するために用いている。利用の際にはモデル出力値に  $40^{\circ}$ C を加算している。これは、 $0^{\circ}$ C を境に説明変数である地上気温の符号が反転することを防ぐことを目的としている。説明変数の正負の反転は寄与量(係数と説明変数の積)が正負に反転することを意味し、精度悪化の一因となる。モデル地上気温は  $40^{\circ}$ C を加算すればほぼ正となるため、この処理を行うことで寄与量の正負反転は起こらなくなる。

東西南北の風向 4 成分は、地形等の影響により各風向によってモデルの系統誤差が変わることに対応するために用いている。各風向成分および風速には地点ごとに上限値を設けており、さらに上限値で除算し  $0\sim1$  に規格化して利用している。風向は 4 方位に分けて利用しており、予測された風向と逆となる成分は 0 としている。風向が逆の場合の、モデルの系統誤差の正負は必ずしも逆ではない。逆方向の成分を 0 とすることで、寄与量の正負反転を防止することができ、風向ごとの系統誤差を適切に軽減できる。

<sup>&</sup>lt;sup>2</sup> それぞれ、航空気象定時・自動・特別観測気象報。

<sup>1</sup> 三戸 洋介

表 4.4.1 最高・最低気温ガイダンスの仕様

ルロウナナ 在			心気温カイタン人	O) 11 13K			
作成対象	アメダス、空港						
作成方法	カルマンフィルタ						
作成対象とするモデル	アメダス:GSM, MSM 空港:MSM						
	GSM:アメ	ダス					
	初期値	当日	翌日	翌々日	3 日後		
	00UTC	最高	最高・最低	最高・最低	最低		
	06UTC		最高・最低	最高・最低	最高・最低		
	12UTC		最高・最低	最高・最低	最高・最低		
	18UTC	最高	最高・最低	最高・最低	最低		
7 to the color to	MSM:アメ	ダス、空港					
予報期間(対象要素)	初期値	当日	翌日	翌々日			
	00UTC	(最高)	最高・最低				
	03UTC		最高・最低				
	06UTC		最高・最低				
	09UTC		最高・最低	(最低)			
	12UTC		最高・最低	(最低)			
	15UTC	最高(・最低)	(最高・)最低				
	18UTC	最高	(最高・)最低				
	21UTC	最高	(最高・)最低				
	「翌日」等は、日本時間で初期時刻から見た日付を示す、()内はアメダスのみ						
逐次学習の有無	あり						
説明変数	モデル予測値(地上の東・西・南・北風成分、地上風速、地上気温、中・下層雲量、						
<b>元明复数</b>	気温減率、降水量、前日との気温差)						
	アメダス:アメダスの気温観測 (1分値)から算出した最高・最低気温						
目的変数	空港:METAR, METAR AUTO, SPECI から算出した最高・最低気温 当日・翌日の、最高気温は 00~09UTC の最高気温、最低気温は 15~00UTC の最低気温。						
			節(暖候期(4~9)	月 ) 寒候期 (1	0~3月)) 予報	対象要素	
層別化処理の対象		最低気温)予報		た行う			
	咳・巻帙期(	の派数切管削に上	か月間の事前学習	を打つ。			

## 表 4.4.2 時系列気温ガイダンスの仕様

作成対象	アメダス、空港
作成方法	カルマンフィルタ
作成対象とするモデル	アメダス:GSM, MSM 空港:MSM
予報対象時間	毎正時
予報期間と間隔	GSM: FT=3 から FT=84 まで 1 時間間隔 MSM: FT=1 から FT=39 まで 1 時間間隔
逐次学習の有無	あり
説明変数	モデル予測値(地上の東・西・南・北風成分、地上風速、地上気温、中・下層雲量、気温 減率、降水量、前日との気温差)
目的変数	アメダス:アメダスで観測された毎正時の気温 空港:METAR, METAR AUTO の毎正時の気温
層別化処理の対象	作成対象地点、初期時刻、予報時間、季節(暖候期(4~9月) 寒候期(10~3月)) 暖・寒候期の係数切替前に1か月間の事前学習を行う。

## 表 4.4.3 格子形式気温ガイダンスの仕様

作成対象	5 km 格子
作成方法	各格子のモデル地上気温を時系列気温ガイダンスとモデル海陸比で補正
作成対象とするモデル	GSM, MSM
予報対象時間	毎正時
予報期間と間隔	GSM: FT=3 から FT=84 まで 1 時間間隔 MSM: FT=1 から FT=39 まで 1 時間間隔

気温減率は、モデル格子点の標高と現実地形の標高 の差による系統誤差を適切に補正するために利用して いる。標高差による系統誤差は主に定数項で補正され るが、気温減率の違いによる補正量を適切に評価する ために導入している。

降水量は、降水時にモデル地上気温が低下するような場合に、この低下が実況と比べて不十分でガイダンスも正バイアスとなる傾向があったため、3 mm/3h を上限として導入している。

前日との気温差は、雨天などによる天気の一時的な変化によるモデルの誤差特性の変化を考慮するために導入している。晴天時と雨天時ではモデルの誤差特性は異なると考えられるが、梅雨期や冬季の日本海側の地方を除けば雨天が続くことは少ない。そのため雨天時の誤差特性の変化は一時的なものとなる。この一時的な変化にカルマンフィルタの学習が過敏に反応し、その後の予測に悪影響を与えることがあった。前日との気温差を利用することで、このような一時的な変化を評価して過敏な学習を防ぐことが期待できるため、説明変数に用いている。

説明変数の計算は、時系列気温では対象時刻の前3時間平均としている。最高・最低気温はモデルの最高・最低気温発現時3の前3時間平均としている。利用する格子点については、気温、風、降水量に関しては、予測対象地点の近接4格子から線形内挿した値を使用している。雲量については、2014年の改良までは最近接格子の値を用いていたが、雲が広がるような状況で雲量予測が過少となる傾向が見られたため、近接4格子の最大値を用いている。

## 層別化

予測式は、地点、初期時刻、予報時間(最高・最低 気温ガイダンスは予測対象日)および季節を対象に層 別化している。気温はモデル標高と実標高の違い、海 陸分布の違い、風の影響度合いの違いなど、地点ごと にモデルの誤差特性が異なることから、予測対象地点 で層別化している。初期時刻と予報時間で層別化する ことによって、予報対象時刻で異なるモデルの誤差特 性の変化に対応できる。第4.4.3項で後述するように、 モデルの平均誤差は予報時間に対して概ね $0\sim-1$ °Cの 間で一日周期の増減をしており、予報対象時刻ごとに 異なった誤差特性がある。特にモデルの海格子にある 地点では、モデル気温はほとんど日変化しないが、時 刻で層別化することによって、ガイダンスは日変化を 表現できる (第1.2節)。季節で層別化している理由 は、学習の程度を調整する観測ノイズの分散などを暖 候期と寒候期で層別化し、季節別の誤差特性に適合し た学習を行うためである (後藤 2014)。

カルマンフィルタのパラメータ設定

気温ガイダンスではシステムノイズの分散、観測ノ イズの分散を層別化の対象ごとに設定している。それ ぞれの値は、過去の一定期間のデータを用いて予測を 繰り返し行い、その予測の二乗誤差から観測ノイズの 分散を、予測式の係数の変動からシステムノイズの分 散を決定している(第2.7.8項)。システムノイズの分 散は設定後は固定としている。観測ノイズの分散も通 常固定としているが、ガイダンスの予測が大きく外れ た時に一時的に大きな値を用いている。観測ノイズの分 散が大きいと係数の修正量が小さくなるため(第 2.7.6 項(4)) 気温の予測を大外しした場合に係数の大きな 変動を抑えることができる。ランダムな要因によるモ デルの予測の大外しをそのまま学習して係数を更新し た場合、逆に精度が悪化する可能性がある。そこで上 記のように係数の変動を抑える仕組みを導入して、そ の後の精度低下を防止している。具体的には、ガイダ ンスの予測が 4°C 以上外れた場合を、モデルが気象場 の予測を外したとみなして、一時的に観測ノイズの分 散を通常の3倍にして係数の変化を小さくしている。

#### (2) 格子形式気温ガイダンス

格子形式気温ガイダンス(以下、格子気温ガイダンス)は、格子形式で毎正時の気温を予測するガイダンスである。GSM, MSM を入力として作成し、いずれも 5 km 格子を予測対象としている。格子気温ガイダンスは最大降雪量ガイダンスの雪水比計算と降水種別ガイダンスの雨雪判別に利用される。

格子気温ガイダンスは、各アメダス地点における時系列気温ガイダンスのモデル気温修正量から、各格子におけるモデル気温修正量を推定し、それを各格子のモデル地上気温に加えることで作成される。時系列気温ガイダンスのモデル気温修正量は、大きく以下の2種類に分類することができる。

- 1. 各説明変数から計算される修正量。
- 2. モデルと現実の標高差や、海陸分布の違いなどに 起因する修正量。アメダス地点や格子毎にある程 度値が決まっており、主に定数項に含まれる。

格子気温ガイダンスでは、これら 2 種類の修正量を別々に扱う。修正量 1 については、空間内挿可能な修正量として扱い、各アメダス地点の値の重み付け平均により各格子のモデル気温修正量を求める。修正量 2 については、基本的に空間内挿に適さない修正量であるので、格子ごとに値を見積もる。格子ごとに算出した修正量 1、修正量 2 をモデル地上気温に加えることで、格子気温ガイダンスを作成する(図 4.4.1)。具体的な作成手順は以下の通りである。なお、数式に用いられる添字は、i はアメダス地点、grid は予測対象格子点の値であることを表す。

各格子点において、大気下層の気温とジオポテンシャル高度から気温減率を計算する。

 $<sup>^3</sup>$  入力データの関係で、GSM は 1 時間ごと、MSM は 30 分ごとで期間内の発現時刻を判定している。

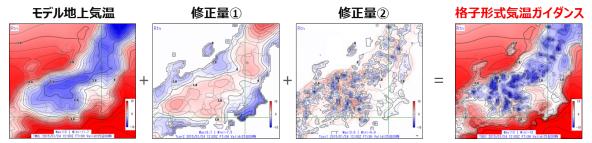


図 4.4.1 格子形式気温ガイダンスの作成の流れの概念図。モデル地上気温に、修正量 1 (格子毎に周囲のアメダス地点の重み付け平均で推定)と修正量 2 (格子毎にモデル標高から 5 km 平均標高への標高補正で推定)を加えることで、格子形式気温ガイダンスは作成される。図中の赤系統の色は正、青系統の色は負値を示す。

で求めた気温減率を用いて、モデル地上気温をモデル標高から  $5~{
m km}$  平均標高  $^4$  へ高度補正する  $(T_2)$  。

各アメダス地点で  $T_2$  と時系列気温ガイダンスと の差をとる  $(T_1)$ 。差を取る際の  $T_2$  はアメダス地 点周辺の近接 4 格子点を用いて線形内挿した値を 用いる。

予測対象格子点を中心とした半径 R [km] 以内に含まれるアメダス地点の  $T_1$  を重み付け平均して  $T_2$  に加算する  $(T_{bcd})$  。

Contour-Following Smoother (CFS) (Glahn et al. 2009) を用いて  $T_{bcd}$  を平滑化し、格子気温ガイダンスとする。

前述した 2 種類の修正量で考えると、 と で各格子点の修正量 2 の見積もり、 で各アメダス地点の修正量 1 の見積もり、 で修正量 1 の各格子点への分配を行っていることになる。

の処理で用いる各アメダス地点の重み  $(w_i)$  には、 予測対象格子点とアメダス地点間の距離  $(d_i)$  について (Cressman 1959) を参考に、またモデルの海陸比も考慮した、以下の式で定義されるものを利用している。

$$w_i = \{1 - |SL_i - SL_{grid}|\} \frac{R^2 - d_i^2}{R^2 + d_i^2}$$
 (4.4.1)

 $SL_{grid}$  は予測対象格子点の、 $SL_i$  はアメダス地点のモデル海陸比  $^5$  で、 $SL_i$  については各アメダス地点の周囲  $^4$  格子のモデル海陸比を内挿した値である。右辺の $\{\}$  内が予測対象格子点と各アメダス地点との海陸の性質の違いによる重みで、R と  $d_i$  による分数が距離による重みである。予測対象格子点と修正に用いられるアメダス地点とのモデル海陸比や距離が近いほど  $w_i$  は大きくなる。なお、距離による重みは 2016 年の改良までは距離の -2 乗を重みとしていたが、予測対象格子点がアメダス地点付近に位置すると、予測値がアメダス気温に寄りすぎてしまい精度が悪化した事例があっ

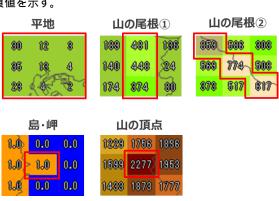


図 4.4.2 Contour-Following Smoother での場合分けの一例。上段および下段右図の数字は格子の標高 [m] を表しており、下段左図の数字は海陸(0:海格子、1:陸格子)を表している。中央の格子の平滑化を行う場合、地形情報(標高、海陸)を元に周囲の格子から地形的な性質の近い格子(図の赤枠内)を選出し、選出された格子のみ用いて平滑化を行う。周囲の格子の選出パターンは13種類あるが、ここでは例として5種類だけを示している。

たため、距離に対する重みが緩やかに変化するようにしている。これらの値から予測対象格子点の $T_{bcd}$  は以下のように求められる。

$$T_{bcd} = T_2 + \frac{\sum_{i=1}^{N} T_{1,i} w_i}{\sum_{i=1}^{N} w_i}$$
 (4.4.2)

と の処理はRを変えて二度行う。一度目はRを  $75~\rm km$  として  $T_{bcd}$  を計算する。二度目は、 $T_{bcd}$  を  $T_2$  と置き直し、R を  $35~\rm km$  として 、 の処理を行う。一度目の処理では広範囲のアメダス地点の情報を利用する。これによって予測対象格子点の近くにアメダス地点が少ない場合でも、複数のアメダス地点の情報を利用して気温の修正を行うことができる。ただし、観測地点が密に存在する地域ではむやみに遠い地点の情報が入り、修正量の推定精度が悪化する可能性がある。このため、二度目の処理では範囲を限定することで、観測が密な地域に対しても適切に修正が行われるようにしている。

までの処理では、地形や気象場と整合しないノイズのような不自然な値が残ることがある。これを除去するため、で CFS により平滑化を行う。 CFS では、まず対象格子点と周辺格子の地形を海陸・標高のデータを元に山の尾根や谷筋、島嶼部など 13 の場合に分

 $<sup>^4</sup>$  国土数値情報を元に作成された  $5~\rm km$  格子の格子平均標高データ。モデル標高はモデルの計算安定性のために平滑化を行っているので、同じ解像度のモデル標高と  $5~\rm km$  平均標高を比較した場合は、一般に後者の方が現実に近く急峻である。  $5~\rm tr$  モデル格子の海陸の割合。 $0\sim1~\rm tr$  の値を取り、完全に海なら $0~\rm tr$  の、完全に陸なら  $1~\rm tr$  となる。

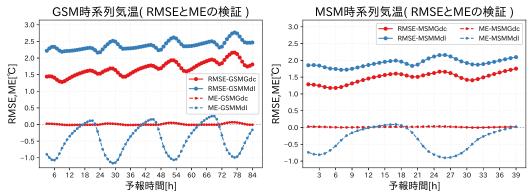


図 4.4.3 時系列気温ガイダンスとモデルの予報時間ごとの RMSE, ME。左が GSM、右が MSM。期間は 2016 年 3 月 ~ 2017 年 2 月。検証に用いた初期値は GSM が 00UTC、MSM が 03UTC。縦軸は RMSE, ME[°C] で、横軸は予報時間 [h]。青線 がモデル、赤線がガイダンスで、実線が RMSE、破線が ME を示す。

類する(図 4.4.2)。分類後、13 の場合それぞれに応じた設定(平滑化に用いる周辺格子や平滑化の強さなど)で平滑処理を行う。この CFS によって、一様な平滑化を行った場合に比べ、地形に由来する特性を保持したまま平滑化できる。

## 4.4.3 統計検証

ここでは気温ガイダンスの統計検証結果を示す。まず、通年で見た時系列気温ガイダンスの予報時間ごとの検証結果から、モデルの予報時間別の誤差とガイダンスの改善の程度を示す。次に、夏季の最高気温ガイダンスと冬季の最低気温ガイダンスの地点別検証から、地域ごとの誤差特性を示す。その後、夏季の最高気温ガイダンスと冬季の最低気温ガイダンスの、実況気温の高低に関する誤差特性を示す。GSM ガイダンスは2017年5月に変更されたGSM の試験データから作成し、MSM ガイダンスは2017年2月に変更されたMSMの試験データから作成した。使用した初期値は、17時発表の天気予報で利用される、GSM の00UTC 初期値とMSM の03UTC 初期値である。

(1) モデルと時系列気温ガイダンスの予報時間別検証 図 4.4.3 に GSM と MSM、およびそれらの時系列気温ガイダンスの検証結果を二乗平均平方根誤差 (RMSE) と平均誤差 (ME) で示す。検証対象は全アメダス地点、検証期間は 2016 年 3 月 1 日から 2017 年 2 月 28 日までの一年間で、検証に用いた予報時間は GSM が  $FT=3\sim84$ , MSM が  $FT=1\sim39$  である。モデル気温は、説明変数や目的変数に利用する時と同様に、各アメダス地点の周辺 4 格子による線形内挿のみを行っており、高度補正は行っていない。

GSM の ME は概ね  $0\sim-1^\circ\mathrm{C}$  の間で一日周期の増減をしている。この周期は、 $\mathrm{GSM}$  の予測誤差が予報対象時刻にも依存しているためである。これに対して  $\mathrm{GSM}$  時系列気温ガイダンスでは、予報時間によらず ME はほぼ  $0^\circ\mathrm{C}$  になっている。これは初期値ごとに予報時間で層別化しているためである。 $\mathrm{RMSE}$  については、 $\mathrm{GSM}$ 

時系列気温ガイダンスは全予報時間で GSM を約  $0.7^{\circ}$  C 改善している。 MSM の場合も GSM と同様で、 MSM 時系列気温ガイダンスの ME は  $0^{\circ}$  C 付近まで改善しており、 RMSE も全予報時間で約  $0.4^{\circ}$  C 改善している。

## (2) 夏の最高気温、冬の最低気温ガイダンスの地点別検証

図 4.4.4 に夏季 ( 2016 年 6 月  $\sim$  8 月 ) の GSM・MSM 最高気温ガイダンスによる翌日予測の ME と RMSE を示す。 GSM 最高気温ガイダンスは、夏は東北地方以南の主に内陸部で負バイアスの傾向、北海道、東北地方北部の太平洋側では正バイアスの傾向がある。 ME の大きさとしては、概ね  $\pm 0.6$  °C 以内となっている。 RMSE を見ると、北日本と東日本の内陸部で誤差が大きい傾向がある。

MSM 最高気温ガイダンスでもバイアスの正負の地域分布に GSM 最高気温ガイダンスと同様の傾向が見られる。ただし、GSM 最高気温ガイダンスに比べると東北地方で正バイアスになっている地点がやや多い。

ME の差分図からは、GSM 最高気温ガイダンスは MSM 最高気温ガイダンスより低く予測する傾向がある地点が多いことがわかる。RMSE の差分図からは、MSM 最高気温ガイダンスの方が GSM 最高気温ガイダンスより精度の良い地点が多いが、北海道南東部、東北地方南部、中国地方西部では MSM 最高気温ガイダンスの方が精度がやや悪い地点があることがわかる。

次に図 4.4.5 に冬季 ( 2016 年 12 月 ~ 2017 年 2 月 ) の GSM・MSM 最低気温ガイダンスによる翌日予測の ME と RMSE を示す。 GSM 最低気温ガイダンスは全 国的に負バイアスの地点が多く、特に北海道の内陸部 に顕著な負バイアスが見られる。これは 2016 年 3 月 に変更された GSM から見られる特徴で、放射冷却による低温を予測したが、実況ではそれほど気温が低下しない事例が多かったことが影響している。これにより、北日本、東日本の内陸部で RMSE が大きい。

MSM 最低気温ガイダンスは、北海道内陸部に正バイアスの大きい地点が見られる。これは、放射冷却による

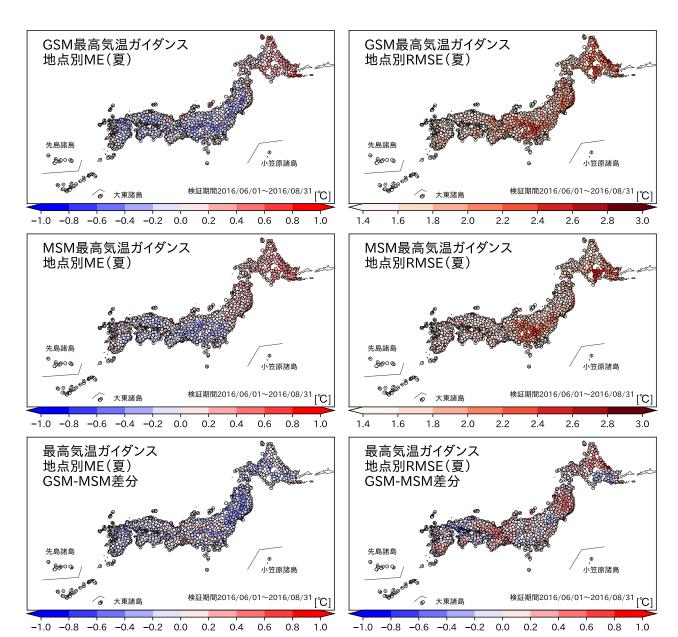


図 4.4.4 夏季の  $GSM \cdot MSM$  最高気温ガイダンスによる翌日予測の ME (左) RMSE (右)。上段が GSM 最高気温ガイダンス、 中段が MSM 最高気温ガイダンス、下段が  $GSM \cdot MSM$  最高気温ガイダンスの差分 (GSM-MSM)。単位は全て [ $^{\circ}$ C]。 検証期間は 2016 年 6 月  $^{\circ}$ 8 月。使用初期値は GSM が 00UTC、MSM が 03UTC。

低温を十分に予測できなかった影響と思われる。MSM 最低気温ガイダンスの RMSE が大きい地点は GSM 最 低気温ガイダンスと同様の分布となっている。

GSM および MSM 最低気温ガイダンスの RMSE の 差分を見ると、MSM 最低気温ガイダンスの方が GSM 最低気温ガイダンスより精度の良い地点が多いが、北 海道北西部、東北地方太平洋側、近畿地方、四国地方には MSM 最低気温ガイダンスの方が精度の悪い地点がある。

# (3) 夏の最高気温、冬の最低気温ガイダンスの実況気温別検証

モデル気温の予測特性は、例えば夏季の顕著な高温 時に負バイアス、冬季の顕著な低温時に正バイアスが あるといった、気温帯ごとに異なる系統誤差を持っていることが知られている。この点から、気温ガイダンスを適切に利用するためには、気温帯ごとの誤差特性を把握することが重要となる。図 4.4.6 に実況気温を3°C ごとに区分けした、GSM および MSM とその最高・最低気温ガイダンスの ME および RMSE、実況気温の事例数を示す。なお、検証対象は全アメダス地点である。

図 4.4.6 (a) の GSM の夏季の最高気温の検証からは、実況が  $36^{\circ}$ C を超える猛暑日において、GSM 最高気温ガイダンスは GSM を大きく改善しているものの、やや負バイアスがあり実況より低めに予測していることがわかる。一方、実況の最高気温が  $21^{\circ}$ C を下回る低い

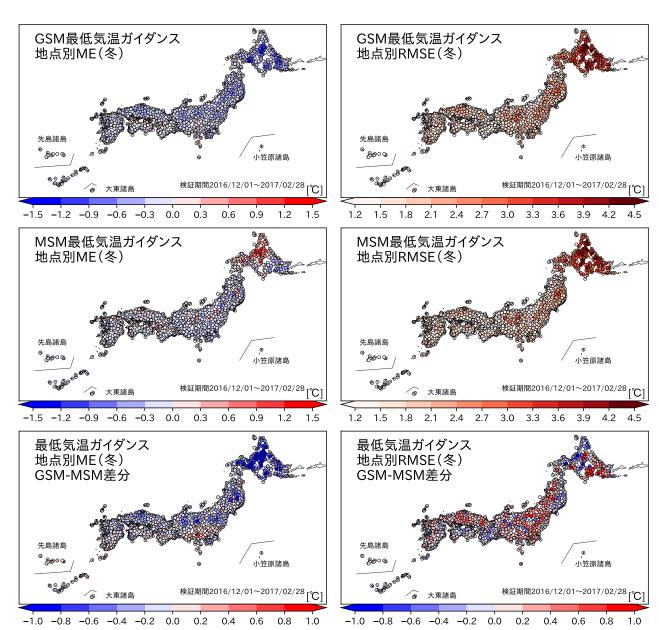


図 4.4.5 冬季の GSM・MSM 最低気温ガイダンスによる翌日予測の ME, RMSE。検証期間は 2016 年 12 月 ~ 2017 年 2 月。 図の並びや単位は図 4.4.4 に同じ。

最高気温の予測については、正バイアスがあり、実況より高めに予測していることがわかる。このような事例の少ない猛暑日や低い最高気温の場合には、極端な予測を出しにくい傾向があり、誤差も大きくなる傾向があることに留意が必要である。図 4.4.6 (b) の GSM の冬季の最低気温の検証からは、実況が $-15^{\circ}$ C を下回る最低気温において、最低気温ガイダンスはモデルを大きく改善しているが、正バイアスがある。地点別検証では冬季の GSM 最低気温ガイダンスが北海道内陸部などに負バイアスの地点が多いことを述べたが、このような極端な低温を予測できているかどうかの検証からは、ガイダンスはその低温を予測できているわけではないことに留意が必要である。

図 4.4.6 (c) の MSM の夏季の最高気温の検証からは、

実況が $36^{\circ}$ C を超える猛暑日において、MSM 最高気温ガイダンスの ME はゼロに近く、GSM 最高気温ガイダンスと傾向が異なっている。また、RMSE も GSM 最高気温ガイダンスより小さく、猛暑日の予測に対しては MSM 最高気温ガイダンスの方が精度が良いことがわかる。一方、実況の最高気温が $21^{\circ}$ C を下回る低い最高気温の予測については、GSM 最高気温ガイダンスと同様に正バイアスがあり、高めに予測していることがわかる。GSM 最高気温ガイダンスよりは RMSE、正バイアスとも小さいため、やはり MSM 最高気温ガイダンスの方が精度が良いが、このような低い最高気温の場合には MSM 最高気温ガイダンスにも正バイアスがあることに留意が必要である。図4.4.6 (d) の MSM 最低気温ガイダンスの冬季の検証からは、実況が $-15^{\circ}$ C

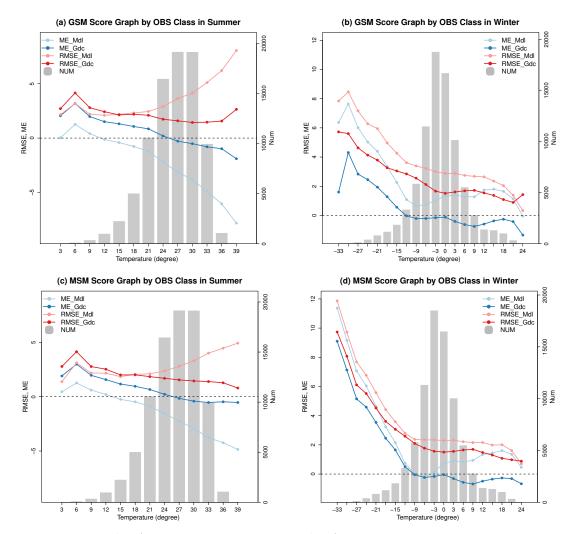


図 4.4.6 夏季の最高気温ガイダンス(左列)と冬季の最低気温ガイダンス(右列)の実況気温別の RMSE, ME。上段が GSM 気温ガイダンス、下段が MSM 気温ガイダンス。どの図も濃赤線がガイダンスの RMSE、濃青線がガイダンスの ME、淡赤線がモデルの RMSE、淡青線がモデルの ME、灰色のヒストグラムは実況の事例数を示す。横軸は実況気温 [°C]。左縦軸は RMSE, ME[°C]。右縦軸は事例数。対象は全アメダス地点。検証期間は、夏季が 2016 年 6 月 8 月、冬季が 2016 年 12 月 2017 年 2 月。使用初期値は GSM が 00UTC、MSM が 03UTC。スコアは 3 °C ごとに計算しており、例えば 18 °C のスコアは 18 °C 以上 21 °C 未満のスコアを示す。

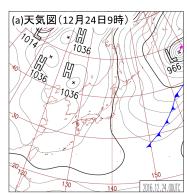
を下回る最低気温において、強い正バイアスがあり、RMSE も大きく、GSM 最低気温ガイダンスより精度が低いことがわかる。GSM 最低気温ガイダンスも極端な最低気温を予測しにくい傾向はあるが、MSM 最低気温ガイダンスはさらにその傾向が強く、誤差が大きくなることに留意が必要である。

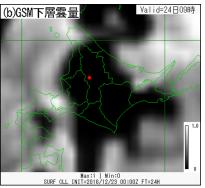
## 4.4.4 事例検証

前項で確認したように、気温ガイダンスは数値予報 モデルが持つ系統誤差を大きく改善し、より精度の高 い予測値を得ることができる。一方で、個々の事例で 見ると、気温ガイダンスで適切な補正を行うことがで きず、気温予測が大きく外れた事例は少なからず存在 する。例えば、数値予報モデルが予測する気象場(雲 の表現や寒気・暖気移流など)が大きく外れた場合に は、その予測誤差を補正できず、気温ガイダンスの予測値も大きく外れることが多い。また、数値予報モデル内である程度系統的なバイアスとして現れているが、現在の気温ガイダンスの仕様では十分に補正しきれないような事例(降水時の気温低下など)もある。気温ガイダンスを適切に利用する上では、これらのような気温ガイダンスの予測が外れやすい現象について十分に理解・留意し、適宜、GSM・MSM 気温ガイダンスの使い分けや修正を行うことが重要となる。以下本項では、気温ガイダンスの予測が大きく外れた3つの事例を紹介し、気温ガイダンスが苦手とするパターンと、それぞれにおける留意点について説明する。

# 冬季北海道内部の放射冷却の事例

地点別検証で、冬季の北海道内陸部では GSM 最低 気温ガイダンスに顕著な負バイアスがあることを述べ





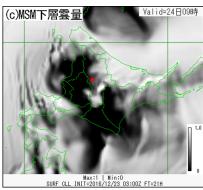


図 4.4.7 2016 年 12 月 24 日 9 時の、 (a) 日本域地上天気図、 (b) 23 日 00UTC 初期値の GSM の下層雲量予測、 (c) 23 日 03UTC 初期値の MSM の下層雲量予測。 (b) と (c) の中央の赤丸は旭川の位置を示す。

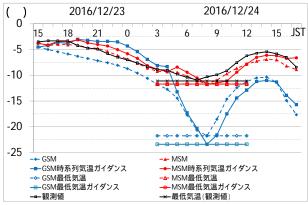


図 4.4.8 2016 年 12 月 23 日 15 時から 24 日 17 時までの旭川の、毎正時の気温と最低気温の観測値(黒線)、23 日 00UTC 初期値の GSM のモデル(破線)とガイダンス(実線)の時系列気温と最低気温の予測値(青線)、23 日 03UTC 初期値の MSM のモデル(破線)とガイダンス(実線)の時系列気温と最低気温の予測値(赤線)。

た。ここでは、モデルで放射冷却が予測された場合の 最低気温予測について特徴的な事例を示す。

図 4.4.7 に 2016 年 12 月 24 日 9 時を対象とした日本 域地上天気図と GSM の下層雲量予測 (23 日 00UTC 初期値)、MSM の下層雲量予測(23 日 03UTC 初期 値)を示す。当時の気圧配置は、北海道の東海上に低 気圧、中国東北区に高気圧があり、西高東低の気圧配 置となっていた(図4.4.7(a))。 旭川の地上気象観測に よると、23日夜から24日未明にかけて断続的に降雪が 観測されたものの、24日2時前に降雪はやみ、3時の 天気は晴れ、6時は快晴であった。しかしその後雲が広 がり、9時の観測では下層雲量は7割(10分量)であっ た。風の実況は、24日0時は8.3 m/s で、時間の経過 とともに次第に弱まったが、24 日 6 時では 4.8 m/s と 弱風とは言えない状況となっており、放射冷却がやや 効きづらい気象場となっていた。これに対し GSM の 雲量予測は、24日朝にかけて雲が次第に消散し、9時 には旭川付近の下層雲はほとんど予測されていなかっ た(図 4.4.7 (b))。 MSM も 24 日朝にかけて旭川付近 の雲量が減少する予測であったが、どの時間帯も GSM よりは雲を多く予測していた ( 図 4.4.7 (c) )。 風速は  $GSM \cdot MSM$  ともに、 $24 \ominus 3$  時から  $24 \ominus 9$  時にかけては 3 m/s 以下の予測であった。

図 4.4.8 に 2016 年 12 月 24 日の、旭川における地 上気温の観測値、GSM・MSM とそれらのガイダンス による時系列気温と最低気温の予測を示す。GSM は 2016年12月23日00UTC、MSM は同日03UTC 初期 値の予測結果である。12月24日朝の旭川の最低気温 は −11.1°C であったが、これに対し、GSM 最低気温 ガイダンスは -23.4°C と予測し、大外れとなった。-方、MSM 最低気温ガイダンスは -11.7°C と概ね正確 に予測していた。時系列気温を見ると、各モデル、ガ イダンスとも日の出後の24日8時まで気温が下がり続 ける予測となっているが、特に GSM と GSM 時系列 気温ガイダンスは MSM に比べて 24 日 5 時以降大きく 気温が下がる予測になっている。 GSM の予測は放射冷 却によるものだが、実際は、6 時過ぎまでは  $4\sim6$  m/s 程度の風が吹いており、また9時にかけては下層雲が 広がったため、観測値は GSM の予測のような大幅な 下がり方とはならなかった。MSM は、9 時の予想で図 4.4.7 (c) にあるように下層雲が予測されていたことで、 放射冷却が効かず、実況と同程度の予測となったと思 われる。

GSM は 2016 年 3 月の改良以降、冬季夜間の放射冷却の予測が外れた場合、実況とは異なる顕著な気温の低下を予測することがある。これはガイダンスにとってランダム誤差となり、このような場合には補正ができずに大外しとなってしまう。これに対し MSM は、放射冷却による低温を予測して外した場合にも GSM ほどの大外しが起こることは少ない。ただし、冬季の MSM 最低気温ガイダンスの ME で見られるように、北海道内陸部では正バイアスがあることに留意が必要である。放射冷却時の気温の下がり方は、雲や風などの気象場により大きく異なる。現状のモデルではこれらを正確に予測することは難しい。モデルによる気象場の予測の妥当性を考慮して、GSM と MSM ガイダンスの両方を参照しながら利用していただきたい。

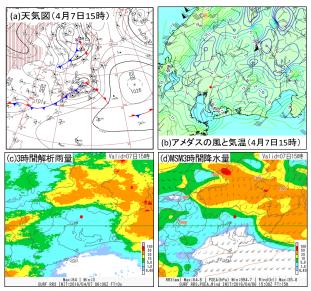


図 4.4.9 2016 年 4 月 7 日 15 時を対象とした、(a) アジア 太平洋地上天気図の日本周辺の拡大図、(b) アメダスの 風向風速と気温の等値線(アメダス観測値を基に作成)、(c) 3 時間積算解析雨量、(d) 6 日 15UTC 初期値の MSM の 3 時間降水量と地上風の予測。(b) から(d) の赤丸は 名古屋の位置を示す。

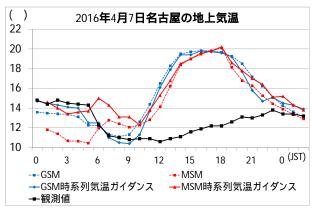


図 4.4.10 2016 年 4 月 7 日 0 時から 8 日 2 時までの名古屋の毎正時の気温の、観測値(黒線) 6 日 12UTC 初期値のGSM のモデル(青破線)と時系列気温ガイダンス(青実線)、6 日 15UTC 初期値のMSM のモデル(赤破線)と時系列気温ガイダンス(赤実線)。

## 冷気層により昇温しなかった事例

放射冷却以外の要因でも、モデルの気温予測が大きく外れた場合は、気温ガイダンスの予測も大きく外れることが多い。ここでは、放射冷却以外の要因で気温予測値が大きく外れた例として、モデルが下層の冷気層を適切に予測できなかった事例を示す。図 4.4.9 に2016 年 4 月 7 日 15 時のアジア太平洋地上天気図の日本周辺の拡大図、アメダスによる風と気温の実況(気温はアメダス観測値を基に作成した等値線) 3 時間積算解析雨量、MSM3 時間降水量予測(6 日 15UTC 初期値)を示す。2016 年 4 月 7 日 15 時には能登半島付近に低気圧があり、東海地方は低気圧の暖域側に位置していた。これにより、東海地方では広い範囲で降水

が観測されていた。一方、名古屋周辺の風は地上気圧配置から想定される南風ではなく、概ね北よりの風が観測されていた。これに対し、15 時の MSM の予測は、降水量は同時刻の解析雨量と比較して概ね良い予測となっていたが、風については南風で暖気が流入すると予測しており、実況とは異なっていた。また、GSM の予測についても概ね MSM と同様に暖気の流入を予測していた(図略)。

図 4.4.10 に 2016 年 4 月 7 日の名古屋の毎時の地上 気温の観測値、GSM (6日 12UTC 初期値)・MSM (6 日 15UTC 初期値)およびそれらの時系列気温ガイダ ンスの予測値を示す。実況では北寄りの風で、南から の暖気が流入せず、日中の昇温は小さい。これに対し、 GSM・MSM は暖気による気温の上昇を予測し、GSM・ MSM 時系列気温ガイダンスもほぼ同様に昇温を予測 して大外れとなった。7日15時のアメダスの気温(図 4.4.9 (b) ) や 7 日 06UTC のメソ解析 (図略) によれ ば、濃尾平野では降水に伴って下層に薄い冷気層が形 成されていた。GSM・MSM ではこの冷気層を予測でき ず、結果としてガイダンスでも大外れとなった。この例 に限らず、現状ではモデルの降水予測が適切であって も、下層の薄い冷気層をモデルで正確に予測すること は難しい。ガイダンスもこれを補正できないため、大 外れとなる場合があることに留意する必要がある。同 様に、モデルによる正確な予測が難しい現象では、位 置や発生のタイミングなどの予測にずれが生じやすく、 ガイダンスが大外れとなる場合があることに留意して いただきたい。

### 降水時の気温低下の事例

第4.4.2 項で説明変数にモデル降水量を導入して、降水時にモデル地上気温の低下が不足する系統誤差を補正していることを述べた。ここでは降水時における気温ガイダンスの予測事例を示し、降水時における気温ガイダンスの特性について述べる。

図 4.4.11 に 2016 年 11 月 19 日 9 時の日本域地上天 気図と、同日 12 時の岐阜周辺の 3 時間積算解析雨量、GSM (18 日 00UTC 初期値 ) MSM (18 日 03UTC 初期値 )の 3 時間降水量予測を示す。2016 年 11 月 19 日 9 時には山陰沖に前線を伴った低気圧があり、東北地方から九州まで広い範囲で降水が観測された。岐阜でも19 日 2 時から 19 時頃まで断続的に降水が観測されていた。モデルは、強雨域の位置ずれがあるものの、解析雨量に見られる岐阜県から長野県にかかる降水域を概ね予測していた。また、降水の時間帯も概ね良く予測できていた(図略)。

図 4.4.12 に岐阜における地上気温の観測値、18 日 00UTC 初期値の GSM 時系列気温ガイダンス、18 日 03UTC 初期値の MSM 時系列気温ガイダンスの予測値を示す。実況は、降水による影響で、日中にかけて 3°C 程度しか昇温していない。一方 GSM・MSM ガイダン

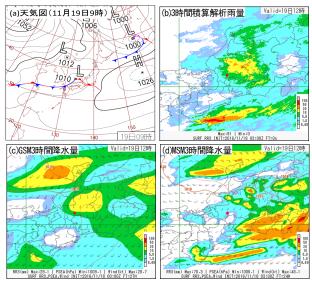


図 4.4.11 (a) 2016 年 11 月 19 日 9 時の日本域地上天気図、2016 年 11 月 19 日 12 時を対象とする、(b) 3 時間積算解析雨量、(c) 18 日 00UTC 初期値の GSM3 時間降水量予測、(d) 18 日 03UTC 初期値の MSM3 時間降水量予測。(b) から(d) の赤丸は岐阜の位置を示す。

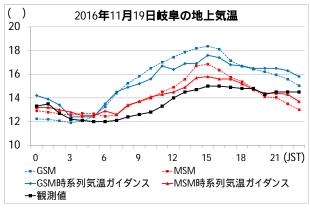


図 4.4.12 2016 年 11 月 19 日 0 時から 23 時までの岐阜の 毎正時の気温の、観測値(黒線) 18 日 00UTC 初期値の GSM のモデル(青破線)と時系列気温ガイダンス(青実 線) 18 日 03UTC 初期値の MSM のモデル(赤破線)と 時系列気温ガイダンス(赤実線)。

スはモデルの予測値を下方修正しているが、修正量は不十分で、5 時以降は実況より高く予測している。この事例では、降水量による寄与は GSM ガイダンスで約-1°C、MSM で約-0.5°C となっており、それぞれのモデルの正バイアスを軽減していたが、その大きさは十分ではなかった。気温ガイダンスでは、ランダム誤差を含むモデルの予測値を説明変数とするため、予測のランダム誤差の大きい要素(降水の有無など)に対しては説明変数による修正が十分ではないことが多い。

このように、モデルの降水予測が適切であっても、 気温ガイダンスによる下方修正は不十分な場合がある。 そのためモデルの降水量予測が適中すると考えた場合 には、気温ガイダンスを下方修正することが有効な場 合もあるが、その際には予測が適中すると考えた根拠 を再確認するなど、モデルの降水予測が外れる可能性に常に留意願いたい。ここでは降水量による説明変数について取り上げたが、他の説明変数についても同様である。例えば風についても、南風の際にモデル地上気温より昇温する地点があっても、モデルの南風予測のランダム誤差を考慮した修正となるため、ガイダンスの上方修正は大きくなく、実況とは違いが生じやすいことに留意が必要である。

## 4.4.5 利用上の留意点と今後の展望

本節のまとめとして、前項までの検証を踏まえた利用上の留意点と今後の展望を述べる。

地点別の最高・最低気温ガイダンスの検証では、MSM ガイダンスの方が GSM ガイダンスより精度の良い地点が多いが、冬季北海道内陸部では GSM ガイダンスの方が精度の良い地点もあることを示した。基本的には MSM ガイダンスの利用を推奨するが、モデルの予測する気象場の妥当性を判断して、適切なガイダンスを選択していただきたい。

実況気温ごとの最高・最低気温ガイダンスの検証からは、GSM・MSM 気温ガイダンスは、夏季に極端な高温や低温を、冬季に極端な低温を予測しづらい傾向があることに留意が必要である。GSM と MSM で比較した場合に、夏季は MSM 最高気温ガイダンスの方が GSM 最高気温ガイダンスより精度が良いが、冬季の極端な低温については GSM 最低気温ガイダンスの方が MSM 最低気温ガイダンスより精度が良いことにも留意していただきたい。

事例検証では、まず、気温ガイダンスの予測が大きく外れた事例の顕著なものとして、冬季の北日本内陸部で GSM が放射冷却の予測を外した時に、GSM 最低気温ガイダンスが低すぎる値を予測した事例を示した。 GSM が冬季に放射冷却による顕著な低温を予測した場合には、放射冷却の予測が妥当かを判断して利用していただきたい。これは放射冷却に限った話ではなく、GSM と MSM の予測する気温が大きく異なる場合には、それぞれのモデルが予測する気象場を確認してガイダンスを利用することが重要である。

また、モデルのランダム誤差を補正できずに気温ガイダンスの予測が大きく外れた事例として、下層の冷気層により暖気が流入せず、GSM・MSM ガイダンス共に大外れする事例を示した。下層の冷気層は両モデルとも苦手とする現象であり、予測の位置ずれなどが発生しやすく、ガイダンスも予測を大きく外すことがある。このような両モデル・ガイダンスとも予測が難しい現象として、他に海風の入るタイミング、局所的なフェーン現象などがある。

更に、説明変数で十分な補正ができなかった事例として降水時の気温低下の例をあげ、降水量の説明変数の効果は限定的なものであり、モデルの地上気温を大きく下方修正できるわけではないことを示した。その

気象管理官, 39-49.

他の説明変数についても同様で、モデルの予測(説明変数の値)が外れる場合もあるため、ガイダンスは大きくモデルの気温予測を修正せず、実況とは違いが生じやすいことに留意が必要である。

最後に気温ガイダンスの今後の展望について述べる。 夏季の最高気温ガイダンス、冬季の最低気温ガイダン スの実況気温に関連したバイアス傾向に対して、2018 年3月現在は説明変数にモデル地上気温を入れてこの 傾向を軽減しているが、モデル地上気温による予測式 の層別化を行うとさらに軽減できる可能性がある。層 別化や説明変数の見直しも含めて検討したい。また、 事例検証で示した GSM が冬季北日本内陸部で放射冷 却の予測を外した時のガイダンスの大外しを緩和する ために、どのような手法が有効であるかも調査を行い たい。ほかにも、メソアンサンブル予報 (MEPS) を入 力とした気温ガイダンスの開発や、複数の気温ガイダンスを統合した統合型ガイダンス(第5.1節)の開発 を行う予定である。

# 参考文献

- 阿部世史之, 2007: ガイダンス類の概説. 平成 19 年度 数値予報研修テキスト, 気象庁予報部, 48-49.
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.
- 海老原智, 酒井喜敏, 山田芳則, 2000: 天気予報ガイダンス. 平成 12 年度数値予報研修テキスト, 気象庁予報部, 28-35.
- 古市豊, 2010: 降雪ガイダンスの改良と利用. 平成 22 年度数値予報研修テキスト, 気象庁予報部, 78-81.
- 古市豊, 松澤直也, 2009: 最大降雪量ガイダンス. 平成 21年度数値予報研修テキスト, 気象庁予報部, 27-38.
- Glahn, B., K. Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, 2009: The gridding of MOS. Wea. Forecasting, 24, 520–529.
- 後藤尚親, 2014: 気温ガイダンスの改良. 平成 26 年度 数値予報研修テキスト, 気象庁予報部, 72-79.
- 気象庁予報部, 1986: 短期間予報用ガイダンスの解説. 予報技術資料第 35 号.
- 黒木志洸, 2017: 格子形式気温ガイダンスの改良、及び 最大降雪量・天気ガイダンスの特性変化. 平成 29 年 度数値予報研修テキスト, 気象庁予報部, 94-101.
- 松本逸平, 海老原智, 2003: 気温ガイダンスの改善. 平成 15 年度数値予報研修テキスト, 気象庁予報部, 47-52.
- 新美和造, 2005: 航空気温ガイダンス. 平成 17 年度数値予報研修テキスト, 気象庁予報部, 60-61.
- 瀬上哲秀, 大林正典, 國次雅司, 藤田司, 1995: カルマンフィルター. 平成7年度数値予報研修テキスト, 気象庁予報部, 66-78.
- 田崎允一, 此田進, 1994: 航空気象予報ガイダンスの利用例について. 航空気象 ノート第48号, 気象庁航空

## 4.5 風ガイダンス<sup>1</sup>

## 4.5.1 はじめに

風ガイダンスは国内のアメダス観測地点および空港地点の地上風を予測するガイダンスである。2018年現在運用中の風ガイダンスには、正時の風向・風速を予測する定時風ガイダンス、前3時間または前1時間内の最大風速・風向を予測する最大風速ガイダンス、前3時間内にガストが発生する確率を予測するガスト発生確率ガイダンス、前1時間内のガスト風速を予測するガスト風速ガイダンスと、前1時間内の最大瞬間風速を予測する最大瞬間風速ガイダンスがある。本節ではこれらのガイダンスを総称して「風ガイダンス」と記述する。風ガイダンスは天気予報、防災気象情報、航空気象情報の作成のための基礎資料として利用されている。

風ガイダンスは、1978年に地上風の日最大風速を予 測するガイダンスとして運用を開始し (気象庁予報部 予報課 1980)、1993 年までに予測対象を 6 時間最大風 速にする変更、最大風速が基準値を超える確率を予測 する風確率ガイダンスの追加、航空用の最大風速およ び風確率ガイダンスの追加などが行われた。これらの ガイダンスにはいずれも線形重回帰が用いられている。 1996 年から 1997 年までには、それまでの最大風速・ 風確率ガイダンスを廃止し、カルマンフィルタによる 定時風ガイダンスの運用を開始した(國次 1997)。その 後、1999年に頻度バイアス補正を導入し(木村 1998)、 2002年にカルマンフィルタによる最大風速ガイダンス の運用を開始した(松本 2003)。また、2011 年にロジ スティック回帰を用いたガスト発生確率と、カルマン フィルタによるガスト風速ガイダンス、最大瞬間風速 ガイダンスの開発を行い (小泉 2011)、2012 年に運用 を開始した。

本節では、まず第 4.5.2 項で風ガイダンスの仕様と作成方法を述べる。続いて第 4.5.3 項と第 4.5.4 項で統計検証の結果と特徴的な事例を示した後、最後に第 4.5.5 項で利用上の留意点と今後の展望を述べる。

# 4.5.2 仕様

以下に、定時風、最大風速、ガスト発生確率、ガスト風速、最大瞬間風速の各ガイダンスの仕様と作成手法を順に示す。

# (1) 定時風ガイダンス

定時風ガイダンスは、アメダスおよび空港地点の正時の前 10 分間平均風速とその風向を予測するガイダンスである。定時風ガイダンスの仕様を表 4.5.1 に示す。予測手法にはカルマンフィルタと頻度バイアス補正を用いており、予測式と頻度バイアス補正の閾値は、地点、初期時刻、予報対象時刻と風向で層別化している。

実況はアメダスまたは METAR, METAR AUTO の風を利用し、GSM では 3 時間毎の、MSM では毎時の予測を行う。目的変数は実況と数値予報モデルの地上風の東西・南北成分の差で、説明変数には数値予報モデルの最近接格子における地上風の東西・南北成分を用いている。

#### 作成方法

定時風ガイダンスの作成手順は以下の通りである。

カルマンフィルタで係数の更新と予測を行う。 で求めた予測値に数値予報モデルの地上風速の 東西・南北成分を加算する。

で求めた風速成分から風向・風速を算出する。 で求めた風速に対して頻度バイアス補正を行う。 の風向と の風速をガイダンス値として出力す る。

カルマンフィルタのパラメータのうち、観測ノイズとシステムノイズの分散は全地点で同一の値を用いており、逐次更新はしない。また、以下の条件に該当する場合にはカルマンフィルタの係数と頻度バイアス補正の閾値は更新しない。

- (i) アメダスの風速が 0 m/s の場合 (METAR, METAR AUTO では 0 ノット<sup>2</sup> (静穏) の場合)
- (ii) METAR, METAR AUTO で風向が VRB と報じ られた場合
- (iii) アメダスの風速が 100 m/s より大きい場合
- (iv) METAR, METAR AUTO の風速が 100 ノットより大きい場合
- (v) モデルの地上風向と観測の風向の差が GSM では 157.5° より大きい場合、MSM では 90° より大きい場合

定時風ガイダンスは風向で層別化しているため、風向が定まらない(i)と(ii)については学習を行わない。(iii)は強すぎる風を学習しないこととしたためで、(iv)はMETARやMETAR AUTOでは100ノット以上は値が報じられない("P99"と報じられる)ためである。(v)は台風中心付近などで風向が大きく外れた場合に不自然な学習をしないようにするために導入している(小泉2013)。MSMと比べて更新しない条件をGSMで緩く設定しているのは、MSMより粗い地形表現に起因して内陸部で風向がより大きく外れる場合があり、厳しく設定した場合に更新機会が少なくなり、予測精度が低下するのを防ぐためである。

# 層別化

風は地形の影響を強く受けるため、地点ごとに数値 予報モデルの誤差特性が異なると考えられることから、 予測対象地点で層別化する。また、数値予報モデルの 地上風は、日中は実況よりも弱く夜間は強いといった

表 4.5.1 定時風ガイダンスの仕様一覧

作成対象	アメダス、空港
作成方法	カルマンフィルタによる予測に風速の頻度バイアス補正を行う。
予報対象とするモデル	アメダス:GSM, MSM 空港:MSM
予報対象時間	3 時間毎の正時 (GSM) 毎正時 (MSM)
予報期間と間隔	アメダス (GSM): FT=3 から FT=84 まで 3 時間間隔 アメダス (MSM): FT=1 から FT=39 まで 1 時間間隔 空港 (MSM): FT=2 から FT=39 まで 1 時間間隔
逐次学習の有無	あり
説明変数	モデル地上風の東西・南北成分
目的変数	実況とモデルの地上風の東西・南北成分との差 実況:アメダスは 1,3 時間毎の正時の風の東西・南北成分、空港は METAR, METAR AUTO (毎正時)の風の東西・南北成分
層別化処理の対象	作成対象地点、初期時刻、予報対象時刻( $1$ 時間每 $1$ 日分 ) 風向( $4$ 方位:北東、南東、南西、北西)
頻度バイアス補正の閾値	アメダスには 2.5, 5.5, 9.5, 13.0 m/s、空港には 5.0, 11.0, 18.0, 24.0 ノット。

### 表 4.5.2 最大風速ガイダンスの仕様一覧

作成対象	アメダス、空港
作成方法	カルマンフィルタによる予測に風速の頻度バイアス補正を行う。
予報対象とするモデル	アメダス:GSM, MSM 空港:MSM
予報対象時間単位	3 時間 (アメダス) 1 時間 (空港)
予報時間と間隔	アメダス (GSM): FT=3 から FT=84 まで 3 時間間隔 アメダス (MSM): FT=3 から FT=39 まで 3 時間間隔 空港 (MSM): FT=2 から FT=39 まで 1 時間間隔
逐次学習の有無	あり
説明変数	モデル地上風の東西・南北成分
目的変数	実況とモデルの地上風の東西・南北成分との差 実況:アメダスは前3時間の最大風速時(10分毎の観測から算出)の風の東西・南北 成分、空港は前1時間の最大風速時(METAR, SPECI, METAR AUTO から算出)の 風の東西・南北成分
層別化処理の対象	作成対象地点、初期時刻、予報対象時刻( $1$ 時間每 $1$ 日分 $)$ 風向( $4$ 方位:北東、南東、南西、北西)
頻度バイアス補正の閾値	アメダスには 3.0, 7.0, 11.0, 15.0 m/s、空港には 6.0, 13.0, 20.0, 25.0 ノット。

ように時刻毎に異なる誤差特性を持つため、初期時刻と予報対象時刻で層別化する。予報対象時刻の層別化は FT=24まで行い、FT=24以降は同じ対象時刻の係数を繰り返して利用している。このようにすることで、例えば FT=36の係数は FT=12 の係数と等しくなり、36 時間後の実況を待たなくても係数が更新されるという利点がある。

風向での層別化は小泉 (2013) で導入した。風向で層別化することにより、風向毎の数値予報モデルの系統誤差を補正できるというメリットがある。その一方で、前線の位置ずれ・台風の進路・速度のずれなどによって数値予報モデルの風向が外れた場合、他の風向の予測式を使用するため、ガイダンスの予測が大きく外れる可能性もある。定時風ガイダンスでは4つの風向(北東、南東、南西、北西)で層別化するが、風向が各方位の境界をまたいで時間変化する場合に時系列でみて予測値が急変しないように、各方位の境界から22.5°まで隣接する方位の風がモデルで予測された場合には、

両方の風向に属する係数と頻度バイアス補正の閾値を 更新する $^3$ 。

## (2) 最大風速ガイダンス

最大風速ガイダンスは、アメダスおよび空港地点の最大風速とその風向を予測するガイダンスである。アメダスを対象としたガイダンスは前 3 時間内の 10 分間平均風速の最大値を、空港地点を対象としたガイダンスは前 1 時間内の 10 分間平均風速の最大値を予測対象とする。最大風速ガイダンスの仕様を表 4.5.2 に示す。予測手法、層別化、説明変数は定時風ガイダンスと同様であるが、頻度バイアス補正の観測側の閾値は定時風よりも大きい値を用いており、これにより注意報基準である  $12\sim15$  m/s での予測精度向上を図っている。

 $<sup>^3</sup>$  例えばモデルの地上風向が  $190^\circ$  であった場合には南西と南東の係数と閾値を同時に更新する。

表 4.5.3 ガスト発生確率ガイダンスの仕様一覧

作成対象	空港
作成方法	ロジスティック回帰
作成対象とするモデル	MSM
予報対象時間単位	3 時間
予報時間と間隔	FT=3 から FT=39 まで 3 時間間隔
逐次学習の有無	なし
説明変数	表 4.5.4 にまとめる。
目的变数	空港における前 3 時間のガスト通報の有無 ( METAR, SPECI, SCAN, METAR AUTO から算出 )
層別化処理の対象	作成対象地点、風向 (8方位:北、北東、東、南東、南、南西、西、北西)

表 4.5.4 ガスト発生確率ガイダンスの説明変数

説明変数	内容
	モデルが予測した地上 $10~\mathrm{m}$ の風速(地上風速)と乱流による風速の変動の前 $1$ 時間最大値の和。乱流による地上 $10~\mathrm{m}$ 風速の変動とは、次の式で定義される平均風速の揺らぎ $C_{ugn}\sigma_u$ のことである。
地上風速最大値(地 上風速の前1時間最 大値)	$\sigma_u = 2.29u_* \cdot \max\left[\left(1 - \frac{1}{24} \frac{z_i}{L}\right)^{\frac{1}{3}}, 1.0\right]$
	ここで、 $C_{ugn}=7.71$ 、 $z_i=1000$ m、 $L$ は Monin-Obukov 長、 $u_*$ は摩擦速度 $u_*=\left(\overline{u'w'}^2+\overline{v'w'}^2\right)^{\frac{1}{4}}$ である (Bechtold and Bidlot 2009)。
境界層最大風速	境界層に含まれる気圧面風速の最大値。境界層を「地上から上層までの大気が混合し得る領域」と定義し、地上から等温位または不安定である層を判定して境界層の高さと算出する。気圧の上限は 500 hPa とする。
水平風の鉛直シアー	地上風と境界層最大風との差分ベクトルの風速を $1000~{ m ft}$ 当たりに換算した値。地上と最大風の気圧面との高度差が $1000~{ m ft}$ 未満の場合は、差分ベクトルの値とする。
SSI	ショワルターの安定指数 ( 850 hPa-500 hPa )。
$\omega 925$	925 hPa <b>面の鉛直</b> p 速度。

#### (3) ガスト発生確率ガイダンス

ガスト発生確率ガイダンスは、空港地点でガストが発生する確率を予測するガイダンスである。ガスト発生確率ガイダンスの仕様を表 4.5.3 に示す。予測手法にはロジスティック回帰を用いており、予測式は地点と風向(8方位)で層別化している。実況は航空観測気象報 (METAR, SPECI, SCAN, METAR AUTO)で前3時間にガストが報じられた場合に実況あり、報じられなかった場合に実況なしとする。

ガストは境界層内での風の乱れによって生じる。境界層内での乱れは、風が強いこと(地形や建物により乱流が強くなる)や成層が不安定であることによって生じる。また、水平風の鉛直シアーが強い場合にはガストも強くなりやすいことから、ガスト発生確率ではこれらの特徴を考慮した説明変数を用いている。ガスト発生確率に用いる説明変数を表 4.5.4 に示す。この表に示した 5 つの説明変数のうち、地上風速最大値、境界層最大風速、 $\omega 925$  の 3 つは必ず利用する。水平風速の鉛直シアーと SSI については、係数の符号が正(鉛直シアー)と負(SSI)の場合のみ説明変数に用いることとしている。このようにした理由は、鉛直シアーの係数が負または SSI の係数が正となる場合、鉛直シアー

が大きいほど、大気の成層が不安定なほどガスト発生 確率が低くなってしまうからである。これらの説明変 数は数値予報モデルの最近接格子から算出している。

# (4) ガスト風速ガイダンス

ガスト風速ガイダンスは、空港地点の前1時間内のガストの最大値を予測するガイダンスである。ガスト風速ガイダンスの仕様を表4.5.5に示す。ガスト風速ガイダンスは、以下の2つのガイダンスから成る。

- モデルの地上風速に乱流の変動を加えた地上風速 最大値を補正する「ガスト風速ガイダンス A」で、 説明変数は地上風速、地上風速最大値、境界層最 大風速(表 4.5.4 を参照)を用いる。
- モデルの地上風速を補正する「ガスト風速ガイダンスB」で、説明変数は地上風を用いる。

モデルの地上風速が 10 m/s 未満の場合はガスト風速ガイダンス A が、10 m/s 以上の場合はガスト風速ガイダンス B が用いられる。ただし、ガスト風速ガイダンス A と B の差が 20 m/s 以上の場合には常にガスト風速ガイダンス B を利用する。このように 2 つのガイダンスを使い分けている理由は、概ねガスト風速ガイダンス A の方が予測精度は良いが、モデルの地上風

表 4.5.5 ガスト風速、最大瞬間風速ガイダンスの仕様一覧

作成対象	ガスト風速:空港 最大瞬間風速:アメダス
作成方法	ガスト風速 $A$ : カルマンフィルタ ガスト風速 $B$ : カルマンフィルタによる予測に風速の頻度パイアス補正を行う。 モデルの地上風速が $10~\mathrm{m/s}$ 未満の場合はガスト風速 $A$ を、 $10~\mathrm{m/s}$ 以上の場合はガスト風速 $B$ をガスト風速ガイダンスの予測値とする。
IFPX/J/A	最大瞬間風速 $A$ 、最大瞬間風速 $B$ : カルマンフィルタによる予測に風速の頻度バイアス補正を行う。 モデルの地上風速が $10~m/s$ 未満の場合は最大瞬間風速 $A$ を、 $10~m/s$ 以上の場合は最大瞬間風速 $B$ を最大瞬間風速ガイダンスの予測値とする。
作成対象とするモデル	MSM
予報対象時間単位	1 時間
予報時間と間隔	ガスト風速 $A$ 、ガスト風速 $B:FT=2$ から $FT=39$ まで $1$ 時間間隔最大瞬間風速 $A$ 、最大瞬間風速 $B:FT=1$ から $FT=39$ まで $1$ 時間間隔
逐次学習の有無	あり
説明変数	ガスト風速 A、最大瞬間風速 A:モデル予測値(地上風速最大値、地上風速、境界層最大風速)(表 4.5.4 を参照) ガスト風速 B、最大瞬間風速 B:モデル予測値(地上風速の東西・南北成分)
目的変数	ガスト風速 A、最大瞬間風速 A:実況と地上風速最大値との差ガスト風速 B、最大瞬間風速 B:実況とモデルの地上風速の東西・南北成分との差実況:ガスト風速は空港における前3時間のガストの最大値(METAR, SPECI, SCAN, METAR AUTO でガストの通報があった場合に限る)、最大瞬間風速はアメダスで観測された1時間毎の最大瞬間風速の東西・南北成分
層別化処理の対象	ガスト風速 A:作成対象地点、初期時刻、予報対象時刻(3時間毎1日分) ガスト風速 B:作成対象地点、初期時刻、予報対象時刻(3時間毎1日分) 風向(4方位:北東、南東、南西、北西) 最大瞬間風速 A、最大瞬間風速 B:作成対象地点、初期時刻、予報対象時刻(1時間毎1日分)
頻度バイアス補正の閾値	ガスト風速 $B$ には、 $1.0,25.0,35.0$ ノット。 最大瞬間風速 $A$ と最大瞬間風速 $B$ には、 $13.0,18.0,23.0~\mathrm{m/s}$ 。

速が 15 m/s 以上ではガスト風速ガイダンス B の方が 精度が良いからである。ガスト風速ガイダンス A はカルマンフィルタを、ガスト風速ガイダンス B はカルマンフィルタと頻度バイアス補正を用いて作成している。予測式は、地点、初期時刻、予報対象時刻で層別化し、ガスト風速ガイダンス B はこれらに加えて風向でも層別化している。予報対象時刻の層別化は 3 時間毎であり、3 時間毎に同じ係数を用いて予測を行っている 4。

目的変数には、ガスト風速ガイダンス A が実況と地上風速最大値との差で、ガスト風速ガイダンス B が実況とモデルの地上風の東西・南北成分との差を用いる。実況は、航空観測気象報 (METAR, SPECI, SCAN, METAR AUTO) でガストの通報があった場合のガストの最大値を用いる。航空観測気象報ではガストの風向は報じられないため、平均風速の風向で代用する。ガイダンスの係数はガストが観測された場合のみ逐次学習するが、予測は全ての初期値と FT で行う。このため、ガスト風速ガイダンスは実際のガストの有無に関わらず、ガストがあったとした場合にどれくらいの風速になるか、という値を予測するガイダンスである

ことに注意が必要である。ガストの有無はガスト発生 確率などを用いて予測し、ガストが予測される場合に その風速をガスト風速ガイダンスで予測するといった 方法での利用を想定している。

# (5) 最大瞬間風速ガイダンス

最大瞬間風速ガイダンスは、アメダスの前1時間内の最大瞬間風速を予測するガイダンスで、ガスト風速ガイダンスと同様の手法で作成している。最大瞬間風速ガイダンスの仕様を表4.5.5に示す。最大瞬間風速ガイダンスとガスト風速ガイダンスの主な相違点は以下の4点である。

- 目的変数がアメダスで前1時間内に観測された最 大瞬間風速であること
- 頻度バイアス補正の観測側の閾値が 13.0, 18.0, 23.0 m/s であること
- 最大瞬間風速ガイダンス A に頻度バイアス補正が 用いられていること
- 予報時間方向には1時間毎に層別化していること

ガスト風速ガイダンスではガストが通報された場合の み学習することから、サンプル数を確保するために3時 間毎の予報時間で層別化している。これに対して最大 瞬間風速の実況は毎時通報されるため、最大瞬間風速

 $<sup>^4</sup>$  例えば  $01,02,03{
m UTC}$  を対象とした予測は同じ係数を用いて行い、係数の学習は 3 つの時刻の中で最大のガストが観測された時刻の実況と説明変数を用いて行う。

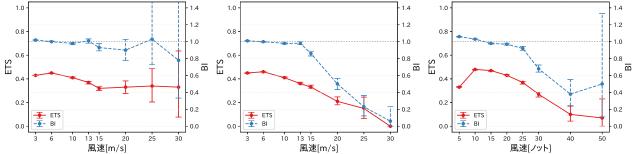


図 4.5.1 定時風ガイダンスの風速の閾値別スコア。(左) GSM ガイダンス、(中) MSM ガイダンス、(右) 航空ガイダンス。 赤実線: ETS、青破線: BI。エラーバーは 95% 信頼区間を表す。横軸は風速の閾値で、単位は GSM, MSM ガイダンスが [m/s]、航空ガイダンスが [Jット] である。

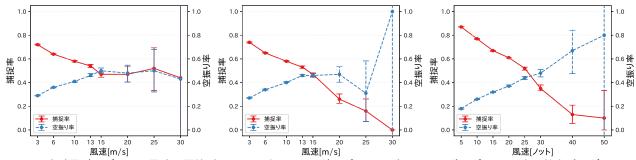


図 4.5.2 定時風ガイダンスの風速の閾値別スコア。(左) GSM ガイダンス、(中) MSM ガイダンス、(右) 航空ガイダンス。赤実線:捕捉率、青破線:空振り率。エラーバーは 95% 信頼区間を表す。横軸は風速の閾値で、単位は GSM, MSM ガイダンスが [m/s]、航空ガイダンスが [Jット] である。

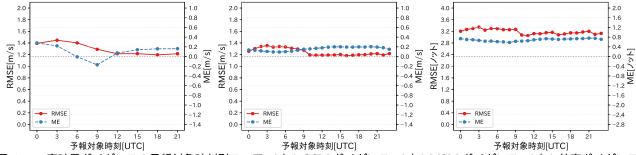


図 4.5.3 定時風ガイダンスの予報対象時刻別スコア。(左) GSM ガイダンス、(中) MSM ガイダンス、(右) 航空ガイダンス。 赤実線: RMSE、青破線: ME。

ガイダンスでは1時間毎の予報時間で層別化している。

## 4.5.3 統計検証

ここでは、GSM, MSM 及び航空の風ガイダンス(以下、本項ではそれぞれ GSM/MSM/航空ガイダンスと書く)について、統計検証と各風ガイダンスが持つ統計的特性を示す。検証期間は 2016 年 1 月から 12 月までの 1 年間である。GSM ガイダンスは 2017 年 5 月に変更された GSM の試験データから作成し、MSM ガイダンスと航空ガイダンスは 2017 年 2 月に変更された MSM の試験データから作成した。GSM ガイダンスは 2017 年 2 月に変更された 2017 年 2017 年

例えば 00UTC 初期値の GSM ガイダンスと 03UTC 初期値の MSM ガイダンスの対象時刻が一致するように選択している。ここでは全対象地点(GSM, MSM ガイダンスは全アメダス、航空ガイダンスは全空港)で検証した結果を示す。よって、GSM, MSM ガイダンスと航空ガイダンスでは対象地点が異なる。風速の閾値別スコアは、ガイダンスの風速が閾値以上の場合に「予測あり」観測の風速が閾値以上の場合に「実況あり」とした分割表から算出する。

## (1) 定時風ガイダンス

図 4.5.1~ 図 4.5.3 に GSM, MSM および航空の定時 風ガイダンスの検証結果を示す。図 4.5.1 の BI と ETS を見ると、GSM ガイダンスでは風速 13~m/s 程度までは BI が 1 に近く、25~m/s 以上でも概ね 1 に近い結果となった(図 4.5.1 左)。一方、MSM(航空)ガイダンスでは、風速 13~m/s(25~Jット)までは BI が 1 に

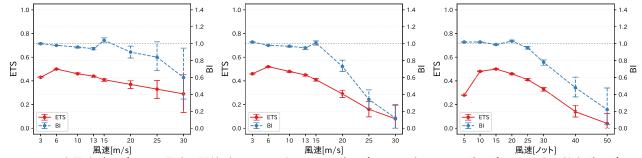


図 4.5.4 最大風速ガイダンスの風速の閾値別スコア。(左) GSM ガイダンス、(中) MSM ガイダンス、(右) 航空ガイダンス。 赤実線: ETS、青破線: BI。エラーバーは 95% 信頼区間を表す。横軸は風速の閾値で、単位は GSM, MSM ガイダンスが [m/s]、航空ガイダンスが [ノット] である。

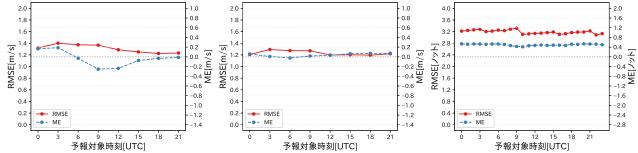


図 4.5.5 最大風速ガイダンスの予報対象時刻別スコア。(左) GSM ガイダンス、(中) MSM ガイダンス、(右) 航空ガイダンス。赤実線:RMSE、青破線:ME。

近いものの、15 m/s (30 ノット)以上では予測頻度が 過少になっている(図 4.5.1 中、右)。 ETS では、各ガ イダンスとも風速 13 m/s までは 0.4 と比較的高い値と なっており、精度が高いことがわかる。しかし、GSM ガイダンスは 13 m/s から 30 m/s にかけての ETS の 低下は小さいが、MSM (航空)ガイダンスでは13 m/s (25 Jット) から 30 m/s(50 Jット) にかけて低下し ている。この BI と ETS の違いを詳しく見るため、空 振り率と捕捉率を図 4.5.2 に示す。 GSM ガイダンスで は風速 15 m/s 以上でも空振り率と捕捉率がほぼ一定な のに対し、MSM (航空)ガイダンスでは風速 15 m/s 以上で捕捉率の低下が顕著である。この MSM (航空) ガイダンスの捕捉率の低下がBIとETSの違いをもた らしている原因である。このような特性は 2013 年の 風ガイダンスの改良(小泉 2013)から見られる。今後、 MSM (航空)ガイダンスの頻度バイアス補正について 閾値の見直し等を含めた調査を行っていく必要がある。

図 4.5.3 の RMSE と ME を見ると、いずれのガイダンスでも風速が強い日中ほど RMSE が大きくなっている。また ME では、予報対象時刻によらず各ガイダンスとも概ね正バイアスであるものの、僅かではあるが日中午後の ME が低下するという日変化を示す。特にGSM ガイダンスでは、09 UTC で負バイアスとなっている(図 4.5.3 左)。予報対象時刻毎に層別化を行っているにもかかわらず、RMSE と ME にこのような特性が見られる原因についての調査を今後行っていく必要がある。

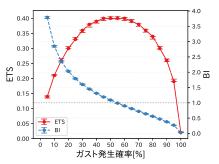
# (2) 最大風速ガイダンス

図 4.5.4 と図 4.5.5 に GSM, MSM および航空の最大風速ガイダンスの検証結果を示す。図 4.5.4 の BI と ETS を見ると、各ガイダンスとも 20 m/s (40 ノット)以上では予測頻度が過少になっているが、GSM より MSM (航空)ガイダンスの BI が低いのは定時風ガイダンスと同様である。ETS では、風速 15 m/s まで GSM と MSM (航空)ガイダンスに大きな差は見られないが、20 m/s 以上では GSM ガイダンスの方が ETS が大きく、精度が高いという定時風ガイダンスと同様の傾向を示している。最大風速と定時風ガイダンスは同様の作成方法を用いていることから、抱えている問題点も定時風ガイダンスと同様であると考えられる。最大風速ガイダンスについても、MSM (航空)ガイダンスの頻度バイアス補正について閾値の見直し等を含めた調査を今後行っていく必要がある。

図 4.5.5 の RMSE と ME を見ると、定時風ガイダンスと同様に、いずれのガイダンスでも風速が強い日中で RMSE が大きくなっている。また ME では、予報対象時刻によらず GSM, MSM ガイダンスで概ね  $0.0~\mathrm{m/s}$  に近いものの、日中午後の ME が低下するという日変化を示す傾向があることも定時風ガイダンスと同様である(図 4.5.5 左、中  $\mathrm{\cite{le}}$ 

## (3) ガスト発生確率ガイダンス

図 4.5.6 にガスト発生確率の閾値別  $BI \ge ETS$ 、および、信頼度曲線を示す。 $BI \ge ETS$  を見ると、確率を 55% としたときに BI が最も 1 に近く、ETS は最大となっている。このことは、平均的には 55% を閾値とす



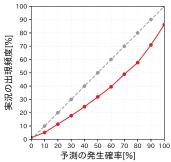
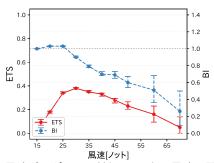


図 4.5.6 ガスト発生確率ガイダンスの確率値別スコアと信頼度曲線。(左)確率値別 BI 及び ETS。赤実線: ETS、青破線: BI。 (右)信頼度曲線。赤実線:信頼度、灰色破線:理想直線。エラーバーは 95%信頼区間を示す。



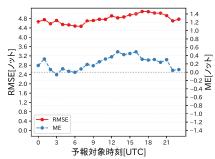
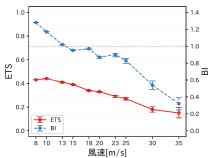


図 4.5.7 ガスト風速ガイダンスの検証図。(左)風速の閾値別 BI 及び ETS。赤実線:ETS、青破線:BI。(右)予報対象時刻 毎の RMSE 及び ME。赤実線:RMSE、青破線:ME。



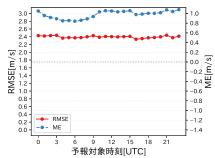


図 4.5.8 最大瞬間風速ガイダンスの検証図。(左)風速の閾値別 BI 及び ETS。赤実線:ETS、青破線:BI。(右)予報対象時刻毎の RMSE 及び ME。赤実線:RMSE、青破線:ME。

ることで「ガストあり」を最も精度良く予測できることを意味している。ただし、空港や風向によってこの 閾値は異なるため、空港毎および風向毎に最適な閾値 を調査する必要がある。信頼度曲線を見ると、10%以上の閾値で 45 度の理想直線よりも右側にあり、実況の 出現率と比べてガストの予測確率が高いことがわかる。この傾向は運用開始当初から見られたが、2017年 2 月の MSM の変更の際に予測式の再作成を行ったものの 過大傾向はさらに拡大しており、今後改めて予測式の作成が必要となっている。

# (4) ガスト風速ガイダンス

図 4.5.7 にガスト風速ガイダンスの風速の閾値別の BI と ETS、および、予報対象時刻別の ME と RMSE を示す。BI と ETS を見ると、30 ノット以上の閾値で BI が 1 より小さく、予測頻度が過少となっている。ETS は 30 ノット付近で最大で、それ以上の風速では徐々に低下し、精度が次第に低下することがわかる。ME と RMSE を見ると、ME は夜間に正バイアスを示し、

RMSE も夜間から早朝にかけて誤差が拡大する傾向があることがわかる。

# (5) 最大瞬間風速ガイダンス

図 4.5.8 に最大瞬間風速ガイダンスの風速の閾値別の BI と ETS、および、予報対象時刻別の ME と RMSE を示す。ガスト風速と同様に、風速が強まるにつれて BI が下がり ETS も下がる傾向にあることは同じであるが、ETS は 10~m/s 付近で最大となっている。ME は全ての予報対象時刻で正バイアスを示すが、日中の ME がその他の時刻に比べて低い傾向があることがわかる。RMSE は全ての予報対象時刻で 2.5~m/s 前後で、日変化などの傾向は見られない。

## 4.5.4 事例検証

ここでは風ガイダンスの予測特性や利用上留意すべ き特性が現れている特徴的な事例を3つ示す。

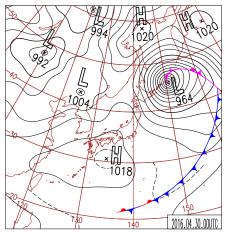


図 4.5.9 2016 年 4 月 30 日 00UTC を対象とする日本域地 ト天気図.

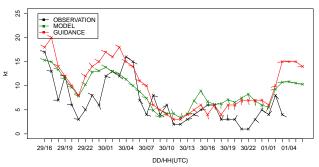


図 4.5.10 仙台空港における 2016 年 4 月 29 日 15UTC 初期値の MSM 航空定時風ガイダンスの予測結果。赤:ガイダンス、緑:MSM 地上風、黒:観測の時系列図。縦軸は風速 [ノット]、図中のバーは風向を、横軸は予測対象日時 [UTC] を表す。

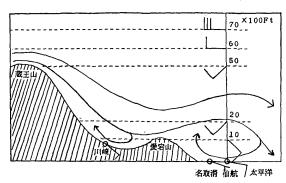


図 4.5.11 西風卓越時の仙台空港周辺の風の流れの模式図 (葛巻 (1991) から引用)。

# 一般風に対して稀な風が吹く場合の予測事例

図 4.5.9 に 2016 年 4 月 30 日 00UTC の日本域地上天 気図を、図 4.5.10 に仙台空港で観測された正時の風の 時系列と、2016 年 4 月 29 日 15UTC 初期値の MSM および航空定時風ガイダンスの予測を示す。30 日 00UTC には千島近海に発達した低気圧があり、北日本では冬型の気圧配置となっていた。30 日 00UTC から 08UTC の仙台空港の風を見ると、MSM では北西の風を予測しているが、観測では南東の風となっている。このときの航空定時風ガイダンスの予測は、MSM と同様に北西風を予測し、MSM の風速を強めて改悪している。

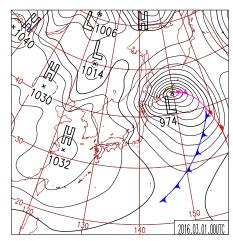


図 4.5.12 2016 年 3 月 1 日 00UTC を対象とする日本域地 上天気図。

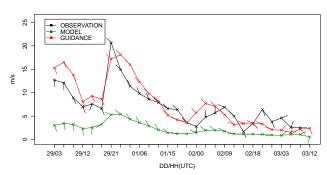


図 4.5.13 札幌における 2016 年 2 月 29 日 00UTC 初期値の GSM 最大風速ガイダンスの予測結果。赤:ガイダンス、緑:GSM 地上風、黒:前 3 時間内に観測された最大風の時系列図。縦軸は風速 [m/s]、図中のバーは風向、横軸は予測対象日時 [UTC] を表す。

仙台空港周辺では、冬型の気圧配置や寒冷前線の通過後に等圧線が南北に走行する際、風上側の山の影響を受けて、山越えの西風が風下側で東風に逆転する場合があることが知られている(葛巻(1991)、図4.5.11)。本事例もこれと同様の理由でモデルやガイダンスと逆向きの風が観測されたと考えられる。定時風ガイダンスではモデルの風向で層別化しているため、一般風に対して稀な風が吹く場合には、ガイダンスの予測が大きく外れる可能性があることに注意が必要である。

## 頻度バイアス補正の最大閾値を超える風の予測

図 4.5.12 に 2016 年 3 月 1 日 00UTC の日本域地上天 気図を、図 4.5.13 に札幌における前 3 時間内に観測された最大風の時系列と、2016 年 2 月 29 日 00UTC 初期値の GSM および GSM 最大風速ガイダンスの予測を示す。1 日 00UTC には千島近海に発達した低気圧があり、日本付近は冬型の気圧配置となっている。1 日明け方には石狩湾小低気圧が札幌を通過し(図略)、風が北西風に変わったタイミングの 0451JST に、札幌で 3 月の歴代 5 位となる 21.2 m/s の日最大風速を観測した。29 日 00UTC 初期値の GSM では石狩湾小低気圧の通過を予測できていたものの、小低気圧中心付近の局所的な強風は予測できていなかった(図略)。一方ガイダ

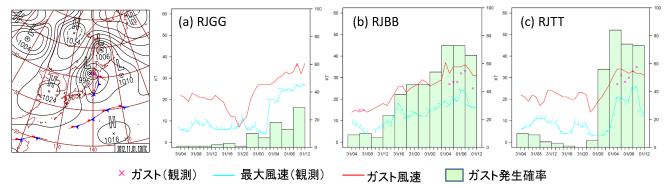


図 4.5.14 2012 年 11 月 1 日 12UTC を対象とする日本域地上天気図と (a) 中部国際空港 (RJGG) (b) 関西国際空港 (RJBB) (c) 東京国際空港 (RJTT) における 2012 年 10 月 31 日 03UTC 初期値のガスト発生確率ガイダンス (右軸) ガスト風速ガイダンス及び観測された最大風速とガスト (左軸) の時系列図。ガストの観測は風速のみ。最大風速の観測のバーは風向を表す。

ンスでは、0600JST で 17.2 m/s の最大風速を予測して おり、モデルの予測を大きく改善していた。ただし実 況と比べるとやや弱い予測となっている。この理由と して、モデルが局所的な強風を予測できなかったこと に加えて、最大風速ガイダンスの頻度バイアス補正に おいて、観測側の閾値の最大を 15.0 m/s に設定してい ることも挙げられる。頻度バイアス補正は実況の頻度 に合うように予測値を補正する手法(第2.9節)であ るが、観測側の閾値の最大を上回るような実況に対し ては予測を適切に補正できない場合がある。実際、統 計検証(図4.5.4左)でも示したように、風速15 m/s を超える最大風速の予測については BI は 1 よりもや や小さくなっている。特に、閾値の最大を超える実況 が稀な地点では、カルマンフィルタの係数や頻度バイ アス補正の閾値が適切に学習できないことに注意が必 要である。

# ガスト発生確率ガイダンスの予測事例

図4.5.14に2012年11月1日12UTCの日本域地上天気図と、中部国際空港(RJGG)、関西国際空港(RJBB)、東京国際空港(RJTT)における、最大風速とガストの実況及び、2012年10月31日03UTC初期値のガスト発生確率とガスト風速ガイダンスの予測の時系列図を示す。1日00UTCの日本域地上天気図(図略)では日本海北部に低気圧があり、寒冷前線が中国地方沿岸にのびていた。この前線は本州を通過し、12UTCには三陸沖から関東の沿岸に達した。この寒冷前線の位相が通過した影響で、RJGGでは0911UTCに、RJBBでは0430~1030UTCに、RJTTでは0423~0902UTCにガストを観測した。

ガスト発生確率ガイダンスの予測を見ると、RJGG、RJBB, RJTT とも、確率のピーク時にガストが観測されたことは共通している。一方、確率の最大値を比較すると、RJGG では 29%であるのに対し、RJBB では 73%、RJTT では 84% と高い値になっていた。これは RJGG はガストが発生する頻度が相対的に少ないこと

が影響している。ガスト発生確率は地点と風向で層別化しているため、ガストが発生しやすい空港や風向では高い確率が予測されやすく、ガストが発生しにくい空港や風向では低くなりやすい傾向がある。ガスト発生確率を利用する場合には、空港と風向の特性を考慮して、目安となる確率を設定しておく必要がある。

# 4.5.5 利用上の留意点と今後の展望

本節の最後に、統計検証及び事例検証を踏まえた利用上の留意点を3点と風ガイダンスの今後の展望を述べる。

1 つ目の留意点として、層別化の影響を挙げる。定時風、最大風、ガスト発生確率ガイダンスでは、予測対象地点に加えてモデルの地上風向でも層別化を行っているため、台風や前線の位置ずれなどにより、モデルの一般風の風向が外れた場合や、一般風に対して通常と異なる風が吹く場合にはガイダンスの誤差が大きくなる。

2 つ目の留意点として、頻度バイアス補正の閾値の 影響を挙げる。頻度バイアス補正を用いている定時風、 最大風速、ガスト風速、最大瞬間風速ガイダンスでは、 頻度バイアス補正の閾値の最大を超えるような強風は、 予測頻度が実況頻度と比べて過少になる傾向がある。 統計検証の結果、この傾向は MSM・航空の定時風・最 大風速ガイダンスで顕著であった。

3 つ目の留意点として、ガスト風速ガイダンスの利用方法を挙げる。ガスト風速ガイダンスは、ガストがあったとした場合の風速を予測するガイダンスであり、ガストが発生するか否かの情報を含んでいない。そのため、まずガスト発生確率ガイダンスなどでガストの有無を判断し、ガストの発生が見込まれる場合にガスト風速ガイダンスでガスト風速を予測する、という利用を想定している。その際、4.5.3 で述べたように、ガストの有無を判断するガスト発生確率の閾値は空港によって異なるので、適切な閾値を調査して利用願いたい。また、ガスト風速ガイダンスの説明変数には安定度や上昇流を用いているものの、不安定降水に伴う突

風のような、数値予報モデルで予測が困難なスケールの現象に伴うガストは予測することはできない。雷雨等が予測される場合は過去の知見などからガストの予測をしていただきたい。統計検証の結果、ガスト発生確率は実況と比べて高い確率を予測する傾向が見られた。現状では確率値そのものを利用するというよりは、確率の時間変化や空港毎の特性を考慮して利用していただきたい。

最後に、風ガイダンスの今後の展望を述べる。定時風・最大風速ガイダンスについては、頻度バイアス補正の閾値は全地点で一律の値を用いているため、強風が吹きやすい地点では頻度バイアス補正が十分に機能していない可能性がある。上述した MSM (航空)の定時風・最大風速ガイダンスで強風時に予測頻度が過少となる点も含め、頻度バイアス補正の閾値を地点の特性に合わせて設定するなどの調査を行いたい。

ガスト発生確率ガイダンスは、2017年2月の MSM の変更時に係数の再作成を行ったにも関わらず、MSM 変更以降に予測精度が低下している状態であることは 既に述べた。また、統計検証の結果、信頼度曲線が理 想直線と比べて高確率側に大きくずれていることも分 かっている。このため、説明変数の追加や層別化の見直 し等を含めた根本的な仕様の検討を行っていく必要が あると考えている。新しい数値予報モデルを利用した 風ガイダンスの開発も進めていく。 LFM では地形が詳 細に表現されるため、これまで予測が困難だった局地的 な風も予測できる可能性があることから、LFM を入力 としたガイダンスの開発を進めている。また、2019年 度からはメソアンサンブル予報システム (MEPS)の 本運用が予定されており、予測の信頼度や最大・最小 などの情報が提供が可能になることから、MEPS を利 用した風ガイダンスの開発も進めていく予定である。

## 参考文献

Bechtold, P. and J. Bidlot, 2009: Parametrization of convective gusts. *ECMWF Newsletter*, *ECMWF*, **119**, 15–18.

木村陽一, 1998: 風ガイダンスの統計的特徴と風速補 正. 平成 10 年度量的予報研修テキスト, 気象庁予報 部, 79-84.

気象庁予報部予報課, 1980: 数値予報資料の利用と問題 点の検討. 気象庁研究時報, **32**, 251-290.

小泉友延, 2011: ガストガイダンスの開発. 平成 23 年 度数値予報研修テキスト, 気象庁予報部, 30-36.

小泉友延, 2013: 風ガイダンスの改良. 平成 25 年度数値予報研修テキスト, 気象庁予報部, 49-57.

國次雅司, 1997: 風ガイダンスの開発. 平成 9 年度量的 予報研修テキスト, 気象庁予報部, 39-44.

葛巻祐二,1991: 北西風場における仙台空港の局地的な 東風について. 気象庁研究時報,43,51-60.

松本逸平, 2003: RSM 及び MSM 最大風速ガイダンス.

平成 15 年数度値予報研修テキスト, 気象庁予報部, 43-46.

## 4.6 天気ガイダンス1

## 4.6.1 はじめに

天気の予測を支援する資料として、天気ガイダンス とお天気マップがある。

天気ガイダンスには、GSM と MSM を入力とした 2 種類のガイダンスがあり、それぞれ  $20~\rm{km}$ ,  $5~\rm{km}$  格子毎に前  $3~\rm{bhl}$ 間の卓越天気を「晴れ」・「曇り」・「雨」・「雨または雪」・「雪」の  $5~\rm{chl}$ 0のカテゴリに判別している。天気ガイダンスは後述する日照率ガイダンスに加え、平均降水量ガイダンス(第  $4.2~\rm{ml}$ 2 解水種別ガイダンス(第  $4.3~\rm{ml}$ 3 の結果を利用し、設定した閾値によって分岐するフローチャートで天気を判別する。天気ガイダンスは天気分布予報の作成を支援するため  $1996~\rm{ml}$ 4 に運用を開始した (柳野 1995)。天気ガイダンスの雨雪判別は運用開始当初は気温ガイダンスと地上相対湿度を用いていた (鎌倉 2007)が、 $2010~\rm{ml}$ 6 年の予報作業支援システム更新以降は上空の気温も考慮した降水種別ガイダンスを利用している (澤田 2009)。

日照率ガイダンスには、天気ガイダンスと同様に  $GSM \cdot MSM$  ガイダンスがあり、それぞれ 20 km, 5 km 格子毎に前3時間日照率2を予測する。日照率ガイダ ンスは、1982年6月から気象官署での日照の確率を線 形重回帰で予測する日照確率ガイダンスとして運用を 開始し、1985年6月からは曇天率(1-日照率)を予 測する曇天率ガイダンスに変更された (気象庁予報部 1986)。1996年3月からは天気ガイダンスの晴れ曇り 判別のために日照率をニューラルネットワーク(第2.6 節)で予測するようになり(柳野 1995)、その後、2001 年3月から目的変数にアメダスの観測値を追加した(佐 藤 2000)。この日照率の予測は天気ガイダンスの内部 で計算し配信していなかったが、2009年より天気ガイ ダンスの参考資料として配信を開始した(國次 2009)。 その際に、過去の経緯と夜間の予測値も配信するため、 ガイダンスとしては曇天率を配信している。

お天気マップは、GSM, MSM, LFM の予測値を直接的に利用して、フローチャートで天気を「快晴」・「晴れ」・「薄曇り」・「曇り」・「雨」・「みぞれ」・「雪」の7つのカテゴリに判別するアルゴリズムである。格子間隔は対応する数値予報モデルと同じである。お天気マップは、数値予報の予測結果から晴れや曇りといった天気を見るために1992年に開発され(瀬上1992)、その後、雨雪判別の手法の変更(萬納寺1994)や、モデル変更に伴う閾値の見直し(安藤2007; 高桑2015)を行ってきた。2018年現在では、アデス(気象情報伝送処理システム)の統合ビューワでお天気マップが利用されている。

本節では、天気ガイダンス、日照率ガイダンスとお 天気マップについて、まず第 4.6.2 項で仕様と作成方法 を、続いて第 4.6.3 項と第 4.6.4 項で統計検証と事例検 証を、最後に第 4.6.5 項で利用上の留意点と今後の展 望を述べる。

## 4.6.2 仕様

本項では、日照率ガイダンス、天気ガイダンス、お 天気マップの仕様を順に説明する。

### 日照率ガイダンス

日照率ガイダンスの仕様を表 4.6.1 に示す。日照率ガイダンスは、前 3 時間日照率(格子形式)を予測する。まず、日照計があるアメダス地点でニューラルネットワークを作成し、その重み係数を用いて格子毎の日照率を計算する。学習は日中のみ行うが、夜間の晴れ曇り判別を行うために、日中と同じ重み係数を利用して夜間の予測も行う。日照率ガイダンスの作成方法を以下で述べる。

## (1) 地点毎に日照率の重み係数を学習

格子形式の前 3 時間日照率を予測するために、まず観測地点毎にニューラルネットワークの重み係数を求める。使用しているネットワークは、入力層に 11 ユニット、1 層の中間層に 5 ユニット、出力層に 1 ユニットを配置した 3 層順伝播型である。中間層と出力層の活性化関数はそれぞれロジスティック関数と線形関数を用いている。説明変数には GSM, MSM ガイダンスで共通の 11 個の変数(表 4.6.1)を用いる。これらの変数は、対象地点の周囲 4 格子のモデル予測値から内挿することで算出している。なお、各説明変数は上限・下限値(相対湿度は  $0\sim100\%$ 、850 hPa と 500 hPa の気温差は  $10\sim45$  K、対数変換後の降水量は  $0\sim1.6$ )を用いて  $0\sim1$  に規格化する。

ニューラルネットワークの重み係数は地点と季節(暖候期(4 月~9 月)、寒候期(10 月~3 月))で層別化している。ただし、季節の変わり目での予測の急変を防ぐため、それぞれの季節が始まる 30 日前から重み係数の学習を開始する。他のガイダンスと異なり、初期時刻、予報時間の層別化は行っておらず、すべての時刻で同じ係数を用いる。GSM は 12UTC、MSM は 15UTC 初期値で、12, 15JST に対応する予報時間の重み係数を学習している。

# (2) 格子毎の日照率予測

次に、(1) で求めたニューラルネットワークの重み係数を用いて格子毎に前 3 時間日照率を予測する。格子形式の日照率の計算に用いる重み係数は、次のように決める。

- 1. 格子内に気象官署があれば、気象官署(1地点)の ニューラルネットワークの重み係数を用いる。
- 2. 格子内にアメダス地点があれば、格子の中心から近い順に最大3地点の重み係数を選ぶ。
- 3. 200 km 以内に気象官署・アメダス地点があれば、 格子の中心から近い順に最大3地点の重み係数を

<sup>1</sup> 小林 大輝

<sup>2</sup> 前3時間の日照時間の割合

表 461	口昭家	( 暴天 寒 )	ガイダン	スの什様一覧	占

作成対象	20 km 格子 (GSM) 、5 km 格子 (MSM)
作成方法	ニューラルネットワーク (第 2.6 節 )
作成対象とするモデル	GSM, MSM
予報対象時間単位	前 3 時間
予報期間と間隔	GSM: FT=6~84 の 3 時間間隔、MSM: FT=3~39 の 3 時間間隔
逐次学習の有無	あり
説明変数	各層 (1000, 925, 850, 700, 500, 400, 300 hPa) の前 3 時間平均相対湿度 [%]、前 3 時間平均気温差 (850 hPa-500 hPa) [K]、前 3 時間降水量 (FT-3 h, FT±0 h, FT+3 h) [mm] ただし、降水量は $\log_{10} (1 + 降水量)$ で変換する。
目的変数	12, 15JST の前3時間日照率(観測された前3時間の日照時間を3時間で割った値)。
層別化処理の対象	対象地点、暖候期( $4{\sim}9$ 月)と寒候期( $10{\sim}3$ 月) 暖・寒候期の係数切替の $30$ 日前から事前学習を行う。
備考	アメダス地点でニューラルネットワークを作成し、それを元に格子形式の日照率を計算する。 予測結果は曇天率( $1-$ 日照率)として配信する。

表 4.6.2 天気ガイダンスの仕様一覧

作成対象	20 km 格子 (GSM) 、5 km 格子 (MSM)		
作成方法	日照率、降水量(第 4.2 節 ) 降水種別(第 4.3 節)から天気を判別。 天気を晴れ、曇り、雨、雨または雪、雪に判別。		
作成対象とするモデル	GSM, MSM		
予報対象時間単位	前 3 時間		
予報期間と間隔	GSM: FT=6~84 の 3 時間間隔、MSM: FT=3~39 の 3 時間間隔		

#### 選ぶ。

4. 上記のいずれにも該当しない場合、地方予報中枢 官署がある 11 地点 3 の重み係数を用いる。

こうして選ばれた各重み係数と格子毎に計算した説明 変数から、次の式を用いて前3時間日照率を計算する。

前 3 時間日照率 = 
$$\sum_{n=1}^{N} \frac{g_n}{d_n^2} / \sum_{n=1}^{N} \frac{1}{d_n^2}$$
 (4.6.1)

ここで、 $g_n$  は選ばれた地点 n の重み係数と格子毎に計算した説明変数から求めた前 3 時間日照率を、 $d_n$  は格子の中心から地点 n までの距離を表し、N は格子毎に用いる地点数(重み係数の数)を表す。ただし、地方予報中枢官署 11 地点の場合は  $d_n=1$  としている。

## 天気ガイダンス

天気ガイダンスの仕様を表 4.6.2 に示す。天気ガイダンスは格子毎の天気判別に、降水種別ガイダンスの前3時間降水種別 (PTYP3) と、降水ガイダンスの3時間平均降水量 (MRR3)、日照率ガイダンスの前3時間曇天率を用いる。天気判別の方法は GSM ガイダンスと MSM ガイダンスで異なる。これは、GSM ガイダンスに利用するデータの格子間隔に違いがあるためである。 MSM ガイダンスの場合は降水種別・平均降水量・日照率ガイダンスはすべて 5km 格子である。これに対して、GSM ガイダンスの場合は降水種別ガイダン

スは 5 km 格子で、平均降水量・日照率ガイダンスは 20 km 格子である。

以下では、 ${
m MSM}$  ガイダンスと  ${
m GSM}$  ガイダンスに分けて天気判別の方法を説明する。図 4.6.1 に  ${
m MSM}$  ガイダンスのフローチャート図を示す。

## MSM 天気ガイダンス

- 1. 降水のカテゴリを以下のように分類する。
  - (a) PTYP3 が「雨」で MRR3 が 1.0 mm/3h 以 上の場合、天気は「雨」。
  - (b) PTYP3が「雨か雪」でMRR3が1.0 mm/3h 以上、または、PTYP3が「雪か雨」でMRR3 が0.5 mm/3h以上であれば、天気は「雨ま たは雪」。
  - (c) PTYP3 が「雪」で MRR3 が 0.5 mm/3h 以 上の場合、天気は「雪」。
- 「雨」「雨または雪」「雪」のいずれでもない場合、 降水なし(図 4.6.1 の )に進む
- 前3時間曇天率が0.5より大きければ天気は「曇り」、0.5以下なら「晴れ」。

# GSM 天気ガイダンス

まず、PTYP3(5km 格子)とMRR3(20km 格子)を用いて降水に関する判別を行う。20km 格子内の5km 格子で、PTYP3が「雨」でMRR3が1.0 mm/3h以上、PTYP3が「雨か雪」でMRR3が1.0 mm/3h以上、PTYP3が「雪か雨」でMRR3が0.5 mm/3h以上、PTYP3が「雪」で

<sup>&</sup>lt;sup>3</sup> 札幌、仙台、東京、新潟、名古屋、大阪、広島、高松、福岡、鹿児島、那覇

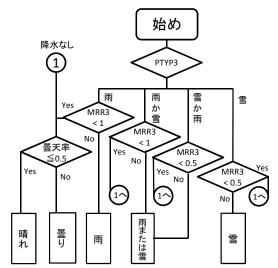


図 4.6.1 MSM 天気ガイダンスのフローチャート図。PTYP3 は降水種別ガイダンスの前3時間降水種別、MRR3は降水ガイダンスの3時間平均降水量、曇天率は日照率ガイダンスの前3時間曇天率を示す。

MRR3 が 0.5 mm/3h 以上の場合の 4 通りに分類 し、それぞれに相当する格子数を  $w_1, w_{21}, w_{22}, w_3$  とする。

- 2. 1 で分類された 20 km 格子内の 5 km 格子数が半 分(8 格子)に満たない場合と、半分以上ある場 合に分類する。
  - (a) 半分に満たない場合は降水なしとし、前3時間曇天率が0.5より大きければ天気は「曇り」、0.5以下であれば天気は「晴れ」。
  - (b) 半分以上ある場合は降水ありとし、次の式の結果 w から分類する。

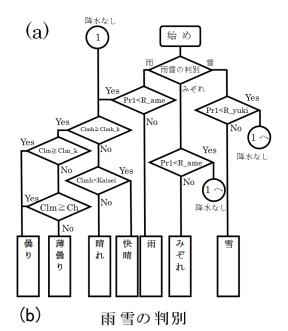
$$S = (w_1 \times 1) + (w_{21} \times 2/3)$$
$$+ (w_{22} \times 1/3) + (w_3 \times 0)$$
$$w = S/(w_1 + w_{21} + w_{22} + w_3)$$

- i. wが 0.85 より大きい場合、天気は「雨」。
- ii. w が 0.85 以下で 0.15 より大きい場合、 天気は「雨または雪」。
- iii. w が 0.15 以下の場合、天気は「雪」。

## お天気マップ

お天気マップの天気判別方法を図 4.6.2 に、閾値の一覧を表 4.6.3 に示す。なお、表 4.6.3 で示した各閾値は気象官署及び特別地域気象観測所の地上実況気象通報の天気を考慮し、調整している。天気判別は次のように行う。

- 1. 図 4.6.2 (b) により、「雨」「みぞれ」「雪」の判別を行う。
- 2.1の判別結果に応じて、以下のように天気を分類 する
  - (a) 「雪」に判別された場合、モデルが予測する



RH(%) 850hPaのT(℃)の基準:-8.0

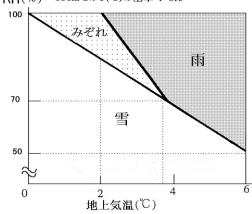


図 4.6.2 (a) お天気マップ天気判別のフローチャート図。 $\Pr1:$ 前 1 時間降水量、 $\Pr1:$  所 1 時間降水量、 $\Pr1:$  所 1 時間降水量、 $\Pr1:$  中層雲量、 $\Pr1:$  所 1 時間降水量、 $\Pr1:$  中層雲量、 $\Pr1:$  所 1 時間降水量、 $\Pr1:$  中層雲量。 $\Pr1:$  所 1 的 1 所 1 的 1 所 1 的 1 所 1 的 1

表 4.6.3 天気判別に用いる閾値と対応するモデル予測値の 一覧。各閾値は気象官署及び特別地域気象観測所の地上実 況気象通報の天気を考慮している。

13172 1K-157 151 C 2 1 1 C 0				
閾値	GSM	MSM	LFM	予測値
R_yuki	0.05	0.05	0.03	Pr1
R_ame	0.4	0.1	0.1	Pr1
Clmh_k	0.5	0.4	0.4	Clmh
Clm_k	0.5	0.4	0.4	Clm
Kaisei	0.1	0.1	0.1	Clmh

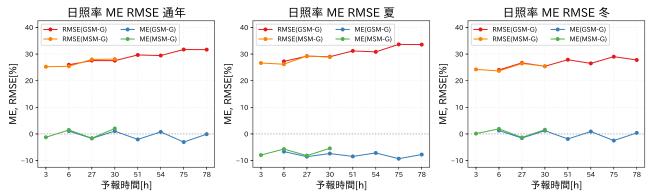


図 4.6.3 00UTC 初期値における日照率ガイダンスの期間、予報時間別の RMSE と ME。左から順に通年、夏期間、冬期間の検証結果で、GSM-G (MSM-G) は GSM (MSM) ガイダンスを表す。ただし、対象とする予測値は日照率。

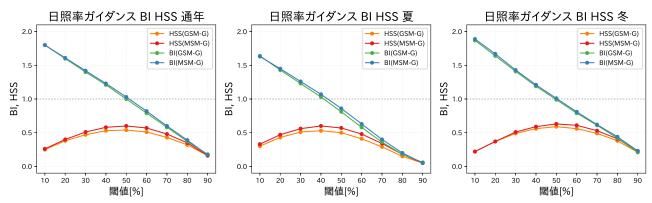


図 4.6.4 00UTC 初期値における日照率ガイダンスの期間、閾値別の BI と HSS。左から順に通年、夏期間、冬期間の検証結果で、GSM-G (MSM-G) は GSM (MSM) ガイダンスを表す。各スコアは晴れに対するスコアで、観測された日照率が 50%以上のときに晴れが、50%未満のときに曇りが観測されたとし、ガイダンスの日照率については横軸の閾値以上の時に晴れが、閾値未満の時に曇りが予測されたとして計算したスコアである。BI は 1 より大きいときに晴れの予測が多いことを示す。

前 1 時間降水量  $(\Pr 1)$  が  $R_{
m yuki}$  以上であれば天気は雪。

- (b) 「みぞれ」に判別された場合、 $\Pr1$  が R\_ame 以上であればみぞれ。
- (c) 「雨」に判別された場合、Pr1 が R\_ame 以上であれば雨。
- 3. 「雪」「みぞれ」「雨」のいずれでもない場合、降水なし(図 4.6.2(a)の )に進む。
- 4. 「快晴」「晴れ」「薄曇り」「曇り」は、ランダム オーバーラップ (Morcrette and C.Jakob 2000) を 仮定して、以下の式で定義する全雲量 (Clmh) 及 び中下層雲量 (Clm) から判別する。

$$Clmh = 1 - (1 - Cl) (1 - Cm) (1 - Ch)$$
  
 $Clm = 1 - (1 - Cl) (1 - Cm)$ 

ここで、Cl, Cm, Ch はそれぞれ、モデルが予測する下層雲、中層雲、上層雲の雲量を表す。

- 5. Clmh が Clmh k 以上の場合、
  - (a) Clm が Clm\_k 以上であれば「曇り」。
  - (b) Clm が Clm\_k より小さい場合、
    - i. Clm が Ch 以上なら「曇り」。
    - ii. Clm より Ch の方が大きければ「薄曇 り」。

- 6. Clmh が Clmh\_k より小さい場合、
  - (a) Clmh が Kaisei 以上であれば「晴れ」。
  - (b) Clmh が Kaisei より小さければ「快晴」。

# 4.6.3 統計検証

ここでは、通年(2016年1月~12月)、夏期間(2016年6月~8月)、冬期間(2016年12月~2017年2月)についてアメダス地点における日照率ガイダンスの検証結果と、気象官署及び特別地域気象観測所における天気ガイダンス・お天気マップと地上実況気象通報の天気を比較した検証結果を示す。天気の検証では、天気ガイダンスが前3時間の卓越天気を予測していること、1 mm/3h以上の雨(0.5 mm/3h以上の雪)を対象としている点で地上実況気象通報の天気と異なることに留意する必要がある。なお、数値予報モデルについては2018年3月現在運用しているものと同じ仕様の予測値を用いた。

### 日照率ガイダンス

まず、期間別に GSM・MSM 日照率ガイダンスのアメダス地点における RMSE と ME ( 巻末付録 A.2.1 ) を図 4.6.3 に示す。 GSM・MSM ガイダンスとも、季節によらず、RMSE は  $25\sim35\%$ 程度となっており、同じ予報時間での GSM と MSM の差は小さい。 GSM・

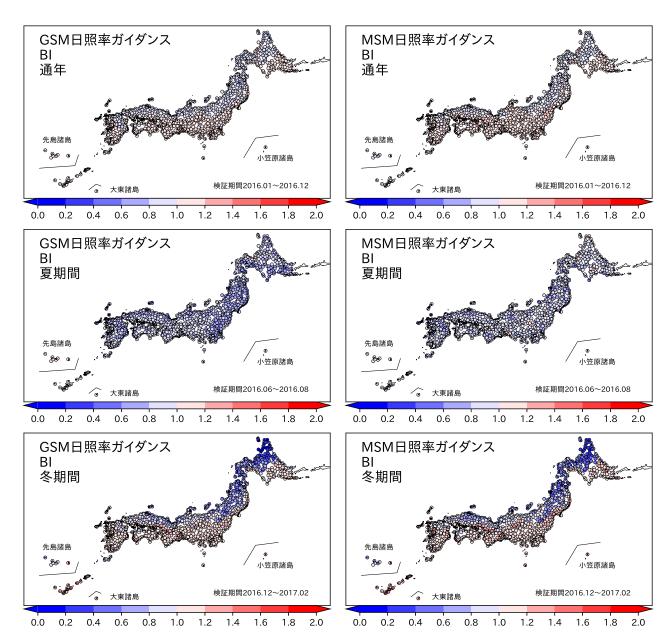


図 4.6.5 00UTC 初期値における日照率ガイダンスの地点、期間別の BI。左列が GSM ガイダンス、右列が MSM ガイダンスで、上から順に通年、夏期間、冬期間の検証結果。なお、BI は晴れに対するスコアで、観測・予測について日照率の晴れ・曇りの閾値を 50%としたスコアである。

MSM ガイダンスとも夏は負バイアスとなっており冬 や通年の結果と比べて RMSE が大きい。

次に、天気ガイダンスへの入力として日照率ガイダンスがどのような予測特性を持っているか説明する。図 4.6.4 に、観測された日照率が 50%以上のときに晴れが、50%未満のときに曇りが観測されたとし、日照率ガイダンスの各閾値別に晴れを予測した場合のバイアススコア (BI) と Heidke Skill Score (HSS) を示す(BI: 巻末付録 A.3.8、HSS: 巻末付録 A.3.12)。ここで示す BI は晴れを対象としており、1 より大きいと晴れの予測が実況に比べて多いことを示す。GSM・MSM 日照率ガイダンスは、通年と冬では閾値を 50%としたときに最も精度が良い結果となった。一方、夏は 50%の

閾値では BI が 1 より低く、日照率を過小に予測してしまうため、閾値を 40% としたときに最も精度が良い結果となった。ただ、閾値を 50% とした場合でも HSS に大きな違いはなく、年間を通して 50% を閾値としても概ね精度良く晴れ・曇りを判別できることが分かる。

日照率ガイダンスで 50%を閾値として晴れを予測した場合の BI および HSS の全国分布をそれぞれ図 4.6.5 と図 4.6.6 に示す。通年で見ると、BI は太平洋側で 1 より大きく日本海側で 1 より小さい。夏では全国的に BI が 1 より小さい。冬の BI は通年と同様、太平洋側で 1 より大きく日本海側(特に北海道と東北地方)で 1 より小さい。これは太平洋側では晴れが多く、日本海側では曇りが多いことにより、それぞれ頻度が多い

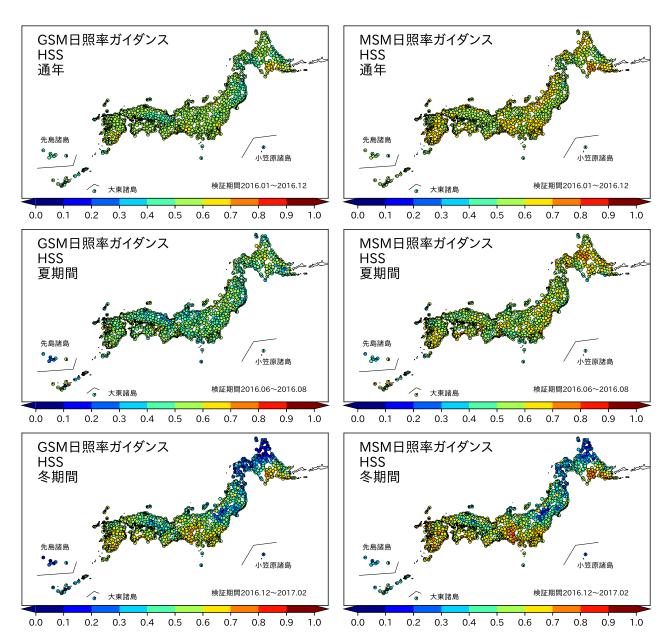


図 4.6.6 00UTC 初期値における日照率ガイダンスの地点、期間別の HSS。左列が GSM ガイダンス、右列が MSM ガイダンスで、上から順に通年、夏期間、冬期間の検証結果。なお、HSS は、観測・予測について日照率の晴れ・曇りの閾値を 50%としたスコアである。

天気の予測に偏ったためと考えられる。太平洋側では下層雲が広がる際に曇りの予測が不十分な例も見られる(後述)。また、HSS は日本海側で小さく太平洋側で大きい。日本海側でHSS が小さい理由は冬の日本海側で晴れの頻度が少なくかつ天気の変動が大きく、晴れの予測が難しいためと考えられ、太平洋側でHSS が大きい理由は日本海側に比べ晴れの日が多くかつ天気の変動が小さく、天気が予測しやすいためと考えられる。全国的にどの期間も GSM ガイダンスより MSM ガイダンスの方が HSS が大きく精度が良い。晴れ曇り判別に日照率ガイダンスを利用する場合は、より精度の良い MSM 日照率ガイダンスを利用していただきたい。

# 天気ガイダンスとお天気マップ

天気ガイダンスとお天気マップの精度を比較するために、まず図 4.6.7 に各天気カテゴリに対する BI と ETS (巻末付録 A.3.11)を示す。通年を見ると、「晴れ」の予測についてはお天気マップは BI が 1 より小さい一方で、天気ガイダンスの BI は 1 に近く ETS も大きい。「曇り」の予測についてはお天気マップは BI が 1 より大きい一方で、天気ガイダンスの BI は 1 に近い。ETS はお天気マップと天気ガイダンスで大きな差はない。「雨」の予測については、お天気マップと天気ガイダンスで BI に大きな差はなく共に 1 より小さいが、天気ガイダンスの方が ETS は大きい。「みぞれ」については、お天気マップと天気ガイダンスとも

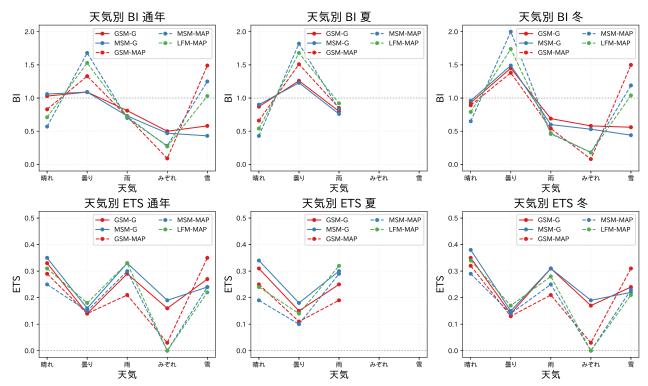


図 4.6.7 00UTC 初期値の天気ガイダンスとお天気マップの、天気と期間別の BI と ETS。上段が BI、下段が ETS で左から順に通年、夏期間、冬期間の検証結果。「モデル名-G」は天気ガイダンスの検証結果、「モデル名-MAP」はお天気マップの検証結果を示す。お天気マップの「快晴」「晴れ」「薄曇り」を「晴れ」、天気ガイダンスの「雨または雪」を「みぞれ」として計算している。

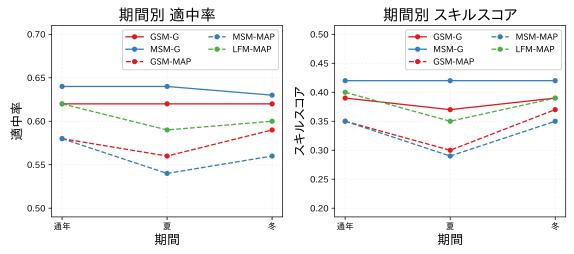


図 4.6.8 00UTC 初期値の天気ガイダンスとお天気マップの期間別適中率とスキルスコア ( 付録 4.6.A )。図の見方は図 4.6.7 と同じ。

にBIは1より小さいが天気ガイダンスの方が1に近く、ETS も大きい。「雪」については、天気ガイダンスはBIが1より小さく、お天気マップは1より大きい。ETS は、GSM お天気マップが最も大きくなっている。夏期間と冬期間についても、同様の傾向があり、「晴れ」「曇り」「雨」「みぞれ」についてはどの期間も概ねお天気マップより天気ガイダンスの方が精度が良く、GSM ガイダンスよりも MSM ガイダンスの方が精度が高い。特に、「晴れ」「みぞれ」についてはどの期間でもお天気マップより天気ガイダンスの方が精度が

良い。しかし、「雪」については天気ガイダンスは BI が低い傾向がある。これは天気ガイダンスは降水量が 0.5 mm/3h 以上の雪をターゲットとしているためである。天気ガイダンスは冬期間について他の期間と比べ「曇り」の BI が1より大きく予測頻度が多くなっており、ETS が小さくなっている。これは冬の日本海側で日照率ガイダンスが十分に「晴れ」を捕捉できず、天気ガイダンスが「曇り」を多く予測するようになっているためと考えられる。

次に、期間別の天気ガイダンスとお天気マップの適

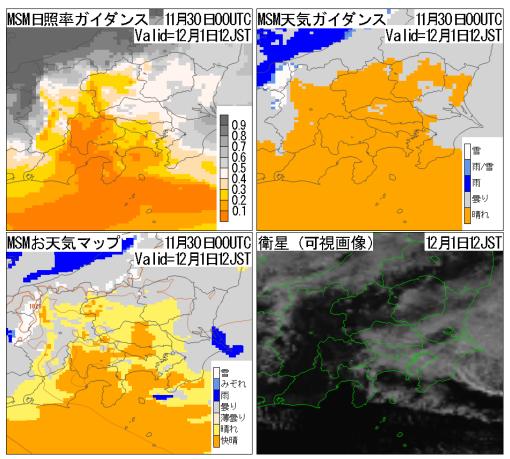


図 4.6.9 上段は左から 2017 年 11 月 30 日 00 UTC 初期値の MSM 日照率ガイダンス (前 3 時間曇天率 ) MSM 天気ガイダンス、下段は左から MSM お天気マップの予測 (対象は 12 月 1 日 12 JST ) と 2017 年 12 月 1 日 12 JST における衛星可視画像。天気ガイダンスの「雨/雪」は「雨または雪」を示す。

中率とスキルスコア(付録 4.6.A)を図 4.6.8 に示す。 どの期間も共通して、天気ガイダンス、特に MSM 天気ガイダンスの精度が良い結果となった。以上より、「晴れ」「曇り」「雨」「みぞれ」については全体的に精度が良い MSM 天気ガイダンスの利用を推奨するが、0.5 mm/3h 未満の弱い「雪」については天気ガイダンスの予測頻度が少ないので、お天気マップもあわせて利用していただきたい。

### 4.6.4 事例検証

ここでは、天気ガイダンス(日照率ガイダンス)とお天気マップについて、晴れ・曇り判別や降雪に関する予測例を紹介する。

## (1) 天気ガイダンスの晴れが多い事例

初めに、天気ガイダンスで晴れを多く予測した事例を示す。2017 年 12 月 1 日は、冬型の気圧配置の中で関東の南東海上に収束線が形成され、関東南部を中心に下層雲が広がった  $^4$ 。図 4.6.9 に、2017 年 11 月 30 日 00 UTC 初期値の FT=27 の MSM 日照率ガイダンス(前 3 時間曇天率)、MSM 天気ガイダンスと MSM お

天気マップ、12 月 1 日 12JST の衛星可視画像を示す。 衛星可視画像で関東南部周辺に見られる雲は、衛星赤 外画像(図略)などから下層雲であったと考えられる。 衛星可視画像によると晴れの領域が長野県から群馬県 にかけて広がっているが、お天気マップは曇りの領域 が多い。これに対して日照率ガイダンスと天気ガイダ ンスは晴れの領域を広げ、実況に近い予測となってい た。一方、埼玉・東京・神奈川付近の下層雲について は、天気ガイダンスとお天気マップで予測できておら ず、ともに晴れ(または快晴)を予測している。ここ で日照率ガイダンスの予測結果を見ると、埼玉県では 前3時間曇天率(日照率ガイダンス)が0.4以上で比較 的大きく予測している。これは実況で下層雲が広がっ ている領域と対応しており、お天気マップの予測より 実況に近い。日照率ガイダンスは上空の相対湿度を説 明変数に用いているが、参照している気圧面の高度方 向の間隔が広いため、薄い下層雲による曇天域を十分 予測できない場合がある。この事例のような場合には 日照率ガイダンスは曇天域を過小評価してしまい、天 気ガイダンスが「晴れ」を広く予測してしまうことが

この事例では、日照率ガイダンスは曇天率を比較的

<sup>4</sup> 本事例は、鈴木・安藤 (1992) で紹介されている関東地方で起こる特徴的なパターンに同じである。

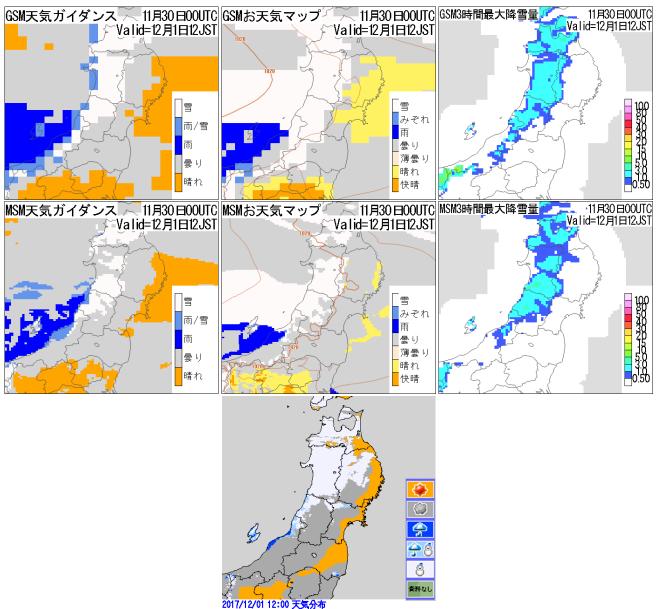


図 4.6.10 2017 年 11 月 30 日 00UTC 初期値の 12 月 1 日 12JST の予測。上段は左から順に GSM 天気ガイダンス、GSM お 天気マップ、GSM3 時間最大降雪量ガイダンス。中段は左から順に MSM 天気ガイダンス、MSM お天気マップ、MSM3 時 間最大降雪量ガイダンス。一番下の図は 2017 年 12 月 1 日 12JST の推計気象分布 (天気) である。天気ガイダンスの「雨/雪」は「雨または雪」を表し、推計気象分布 (天気) は、橙が「晴れ」、灰色が「曇り」、青が「雨」、水色が「雨または雪」、白が「雪」を表す。

大きく予測していた。この事例のように晴れ曇りを判別する際には、天気ガイダンスに加えて日照率ガイダンスも利用することで、ガイダンスの予測を修正できる場合がある。

# (2) 雪の予測事例

次に、天気ガイダンスとお天気マップの雪の予測事例を示す。2017年 12 月 1 日 12 JST の東北地方の雪について 11 月 30 日 00 UTC 初期値の予測と実況(推計気象分布)を図 4.6.10 に示す。MSM 天気ガイダンスでは、秋田県と岩手県の雪域を実況より狭く予測している。これは前述したように天気ガイダンスは 0.5~mm/3h 以

上の雪をターゲットとしているためである。MSM お天気マップは弱い雪もターゲットとしているため、天気ガイダンスよりも雪域を広く予測しており弱い雪を含んだ実況に近い。また、MSM3 時間最大降雪量ガイダンスは、弱い降雪も含めて量的に予測し、かつ統計手法を使っているため、秋田県と岩手県の雪域についてMSM お天気マップと MSM 天気ガイダンスより実況に近い。GSM 天気ガイダンスについても同様で岩手県の雪域を狭く予測しているが、GSM お天気マップは広く予測しており実況に近い。ただし、GSM お天気マップは雪域を福島県まで広げやや過大評価している。GSM3 時間最大降雪量ガイダンスを見ると、秋田県と

岩手県の雪域は GSM お天気マップと同様に予測できており、また、山形県から新潟県にかけての雪域の予測も適切に近い。このように、弱い雪を予測する際は、最大降雪量ガイダンスやお天気マップも参考にしていただきたい。

# 4.6.5 利用上の留意点と今後の課題

最後に、天気ガイダンスとお天気マップについて利用上の留意点と今後の課題を述べる。

### 利用上の留意点

天気ガイダンスとお天気マップを利用する場合、以 下の点に留意していただきたい。

- お天気マップでは数値予報モデルの出力値を直接 的に利用するのに対し、天気ガイダンスでは数値 予報モデルのバイアスが補正された日照率、降水 量および降水種別ガイダンスを利用して天気を判 別している、という違いがある。統計検証の結果 では天気ガイダンスはお天気マップと比べて予測 精度が高いことから、基本的には天気ガイダンス の利用を推奨する。
- 第 4.6.4 項 (1) のとおり、下層の薄い雲に対しては 日照率ガイダンスの予測が晴れ曇りの閾値となる 50%に満たない場合があり、天気ガイダンスで晴 れが予測される場合がある。このような状況が予 測される場合は、日照率ガイダンスで 50%に近い 領域の予測を参考にすることで天気ガイダンスの 予測を修正できることがある。
- お天気マップは降水の有無の閾値を観測にあわせて決めているが、天気ガイダンスは天気予報の降水の付加基準(雨は1 mm/3h、雪は0.5 mm/3h)を閾値にしており、第4.6.4 項(2)のように弱い雪を含む雪域を狭く予測する場合がある。弱い雪を予測する際にはお天気マップ、3時間最大降雪量ガイダンスを参考にしていただきたい。

# 今後の課題

日照率ガイダンスのニューラルネットワークの基本的な作りは1996年から変わっておらず、また、第4.6.4項で述べたように説明変数に利用している相対湿度は高さ方向に粗く取得しているなどの問題点があることから、これらについて改良を行う必要がある。また、第4.6.3項で述べたように、天気ガイダンスの入力としている日照率ガイダンスは夏では閾値を50%とするよりも40%とした方が精度が良いため、閾値について検討する必要がある。天気ガイダンスについては「雪」の判別精度が低いという問題もあるので、閾値調整についてランダムフォレスト(第2.10節)の導入により精度が向上する可能性がある。天気ガイダンスは天気カテゴリを予測するガイダンスであるから、推計気象分布(天気)などを目的変数としてニューラルネットワークや多項ロジスティック回帰も利用できる可能性

がある。また、メソアンサンブル予報システムの予測 結果を入力とする日照率ガイダンスの開発を行い、ア ンサンブル平均を利用することで日照率ガイダンスの 精度を向上させることも今後の課題となっている。

## 参考文献

- 安藤昭芳, 2007: お天気マップ. 平成 19 年度数値予報 研修テキスト, 気象庁予報部, 94-97.
- 鎌倉智之, 2007: 天気ガイダンス. 平成 19 年度数値予報研修テキスト, 気象庁予報部, 73-75.
- 気象庁予報部, 1986: 短期間予報用ガイダンスの解説. 予報技術資料第 35 号, 気象庁予報部.
- 國次雅司, 2009: 概説. 平成 21 年度数値予報研修テキスト, 気象庁予報部, p. 20.
- 萬納寺信崇, 1994: お天気マップ. 平成 6 年度数値予報 研修テキスト, 気象庁予報部, 90-91.
- Matsuo, T., Y. Sasyo, and Y. Sato, 1981: Relationship between types of precipitation on the ground and surface meteorological elements. *J. Meteor. Soc. Japan*, **59**, 462–476.
- Morcrette, J.J. and C.Jakob, 2000: The Response of the ECMWF Model to Changes in the Cloud Overlap Assumption. *Mon. Wea. Rev.*, **128**, 1707–1732.
- 佐藤豊, 2000: RSM 天気ガイダンスの改善と GSM 天 気ガイダンスの開発. 平成 12 年度量的予報研修テキスト, 気象庁予報部, 63-70.
- 澤田康子, 2009: その他のガイダンス (4) 天気ガイダンス、(5) 日照率ガイダンス. 平成 21 年度数値予報研修テキスト, 気象庁予報部, 44-45.
- 瀬上哲秀, 1992: お天気マップ. 平成 4 年度数値予報研修テキスト, 気象庁予報部, 69-82.
- 鈴木和史, 安藤義彦, 1992: 関東地方に局地的悪天をも たらすメソ スケール雲システムのメソ天気系モデ ル. 気象庁研究時報, 44, 63-79.
- 高桑健一, 2015: お天気マップの改良. 平成 27 年度数値予報研修テキスト, 気象庁予報部, 79-81.
- 柳野健, 1995: ニューラルネットによるガイダンス. 平成7年度量的予報研修テキスト, 気象庁予報部, 54-69.

# 付録 4.6.A 天気ガイダンス及びお天気マップの適 中率とスキルスコア

第 4.6.3 項で天気ガイダンスとお天気マップの検証に用いた適中率とスキルスコアについて説明する。以下では、各天気カテゴリ「晴れ」「曇り」「雨」「みぞれ」「雪」に対して、順に  $1\sim 5$  の数字を割り振り、i 番目の天気カテゴリについて、ある期間内の事例数  $\mathbb N$  に対して制測された回数を  $\mathbb M_i$ 、予測された回数を  $\mathbb F_i$ 、そのうち予測が適中した数を  $\mathbb FO_i$  とする。

# 適中率

基本的な考え方は巻末付録 A.3.2 の適中率と同じである。天気判別の適中率は次の式で定義される。

適中率 
$$\equiv \frac{\sum_{i=1}^{5} \mathrm{FO}_i}{\mathrm{N}} \tag{4.6.2}$$

適中率は最大値の1に近いほど予測の精度が高いことを示す。

# スキルスコア

基本的な考え方は巻末付録 A.3.12 の HSS と同じである  $^5$ 。天気判別のスキルスコアは次のように定義する。

スキルスコア = 
$$\frac{\sum_{i=1}^{5} (\text{FO}_i - \text{S}_i)}{\text{N} - \sum_{i=1}^{5} \text{S}_i}$$
 (4.6.3)

ただし、

$$S_i = P_i \cdot F_i, \quad P_i = \frac{M_i}{N}$$
 (4.6.4)

ここで、 $P_i$  は天気カテゴリi の気候学的出現率で、 $S_i$  は天気i を  $F_i$  回ランダムに予測した場合、予測と実況の天気が一致する事例数である。SS は最大値の1 に近いほど精度が高いことを示し、0 のときは気候値予測と精度が同程度、負のときは気候値予測より精度が低いことを示す。

<sup>5</sup> このスキルスコアを HSS と呼ぶこともある

### 4.7 発雷確率ガイダンス<sup>1</sup>

## 4.7.1 はじめに

発雷確率ガイダンスは、全国の 20 km 格子毎の前3時 間に発雷する確率 (PoT) を予測する格子形式のガイダ ンスであり、雷注意報の発表判断や飛行場予報などに用 いられている。2018年現在、モデル予報値に GSM を利 用した GSM 発雷確率ガイダンス (GSM-PoT)、MSM を利用した MSM 発雷確率ガイダンス (MSM-PoT) が 運用されている。また、空港地点向けに MSM-PoT の 空港周囲の4格子を空間内挿した値を航空ガイダンス として作成している。雷に関するガイダンスの歴史は 長く<sup>2</sup>、1981 年には統計手法に線形重回帰 (第 2.4 節) を用いた雷雨確率ガイダンスが運用開始されている (気 象庁予報部 1986)。その後、1996年3月からは統計手 法はニューラルネットワーク (第2.6節)に、名称も発雷 確率ガイダンスに変更された (田沢 1996)。2007年か らは GSM-PoT と MSM-PoT の 2 つのガイダンスが 運用開始となり、MSM-PoT には統計手法としてロジ スティック回帰 (第 2.5 節) が導入された (高田 2007)。 2009 年には GSM-PoT にも MSM-PoT と同様の手法 が導入され (高田 2009)、両ガイダンスの予測手法は 統一された。近年は、2015年に予測手法の改良(白山 2015) が実施されているほか、モデル更新と同時に予 測式の再作成を行うことにより、精度の維持・向上を 図っている。このように、発雷確率ガイダンスの統計 手法はこれまでに何度か変更されてきたが、どの手法 に対しても一括学習を用いているという特徴がある。 雷は発生頻度の低い現象であるため、第 2.2.3 項で述 べられている通り、一括学習を行うことで予測精度を 維持している。

本節ではまず、第 4.7.2 項でガイダンスの仕様と作成 方法を述べ、次に第 4.7.3 項と第 4.7.4 項で統計検証と 事例検証から予測特性について説明し、最後に第 4.7.5 項で利用上の留意点と今後の展望について述べる。

## 4.7.2 仕様

発雷確率ガイダンスの仕様を表 4.7.1 に示す。GSM-PoT は 20 km 格子毎、MSM-PoT は 5 km 格子毎 3 に、PoT をロジスティック回帰を用いて予測する係数 固定型のガイダンスである。目的変数を各格子の周囲 60 km 四方における発雷の有無とし、6 つの説明変数を用いた回帰式(予測式と呼ぶ)により予測を行う。発 雷確率は 20 km 格子毎に計算を行うが、よりメリハリのついた予測値とするため、また、雷雲から離れた場所への落雷事例に対応するために、60 km 四方の地域確率(地域のどこかで現象が起きる確率)を目的変数

1 土田 尚侑

に設定している  $^4$ 。予報時間は GSM-PoT が FT=6 から 84 まで、MSM-PoT が FT=3 から 39 まで、予報時間間隔はいずれも 3 時間である。予測式は、地域、 $-10\,^\circ$ C 高度、予報時間により層別化を行い、 $-10\,^\circ$ C 高度が  $^5$  km 以上の場合のみ予報対象時刻でも層別化を行う。また、予測値は LAF 法 (第  $^2$   $^2$   $^5$ 0  $^5$ 0 GSM-PoT とMSM-PoT の主な仕様は統一されており、両ガイダンスに対する作成方法の違い等は意識することなく利用できる。

# (1) 作成方法

発雷確率ガイダンスの作成は以下の通り、 予測式の作成、 PoT 計算の順に行う。

### 予測式の作成

過去数年間の発雷実況から作成した目的変数と数値予報予測値から計算した説明変数を用いてロジスティック回帰を行い、予測式を作成する。発雷確率ガイダンスは係数固定型ガイダンスであるため、この手順は説明変数の作成に用いる数値予報モデルが更新されて予測特性が変わる場合や、ガイダンス改良などの際に実施する。最近では2015年の GSM, MSM ガイダンスの改良 (白山 2015)、2016年の GSM 物理過程改良 (後藤 2016)、2017年の MSM モデル更新 (後藤 2017b) および GSM 物理過程改良 (後藤 2017a) でそれぞれ予測式の再作成を行っている。

## PoT 計算

数値予報予測値から説明変数を計算し、格子毎にPoTを計算する<sup>6</sup>。ただし、各格子のPoTは周囲9格子の係数を用いた平均値とする。これは、複数の格子をまとめた地域毎に予測式を層別化しているために、地域の境界でPoTの分布が不連続にならないようにするためである。ここで計算したPoTに対し、LAF法により過去初期値を用いた重み付け平均を行ったものを最終的なPoTとする。

# (2) 目的变数

目的変数である発雷の有無には、全国の空港に設置された  $LIDEN^7$  による雷観測(対地放電及び雲放電)を用いる。ただし、以下に示す品質管理処理を行う。

● LIDEN の標定位置から 20 km 以内、標定時刻の± 10 分以内のレーダーエコー強度 (10 分毎、1 km 格子の全国合成レーダーデータ)に対して季節や

<sup>2</sup> 来歴の詳細は第1.3 節を参照。

 $<sup>^3</sup>$  予測値は  $20~{
m km}$  格子で作成し、 $20~{
m km}$  格子内の  $5~{
m km}$  格子に同じ値を割り振る。

<sup>4</sup> 地域確率は、対象とする地域の面積に依存する。開発時に面積を変えて実験した結果、60 km 四方程度が最も利用しやすいと判断した (高田 2007)。

<sup>&</sup>lt;sup>5</sup> LAF 法を適用する前の予測値も配信している。

<sup>&</sup>lt;sup>6</sup> 海岸線から離れた一部の海上格子では予測式を作成していないため、予測対象格子から近い地域の予測式を選択して利用している。このとき利用する予測式は、地域層別化を導入する際に、発雷の頻度分布を参考に主観的に決定した。

<sup>&</sup>lt;sup>7</sup> 雷監視システム:LIghtning DEtection Network system

表 4.7.1 発雷確率ガイダンスの仕様一覧

作成対象	GSM: 20 km 格子、MSM: 5 km 格子 <sup>3</sup>
作成方法	ロジスティック回帰
予報対象時間単位	3 時間
予報期間と間隔	GSM: FT=6 から FT=84 まで 3 時間間隔、MSM: FT=3 から FT=39 まで 3 時間間隔
逐次学習の有無	なし
説明変数	表 4.7.2 を参照
目的変数	対象とする 20 km 格子を含む周囲 9 格子(60 km 四方)における前 3 時間の発雷の有無。 LIDEN、飛行場観測、地上気象観測から作成する。詳細は第 4.7.2 項を参照。
層別化処理の対象	地域、 $-10~^{\circ}\mathrm{C}$ 高度、予報時間、予報対象時刻。詳細は第 $4.7.2$ 項を参照。
備考	予測は $\mathrm{LAF}$ (過去初期値との重み付き平均 ) ありと $\mathrm{LAF}$ なしを作成する。詳細は第 $4.7.2$ 項を参照。

地域に応じた閾値を設定し、閾値を下回っていた 場合には LIDEN の標定を誤りとする。

- LIDEN の標定の有無にかかわらず、METAR<sup>8</sup>、 SPECI<sup>9</sup> の飛行場観測で雷を観測していれば空港 を含む格子を発雷あり、観測がなければ発雷なし とする。
- 地上気象観測で雷を観測していれば発雷ありとする。対象格子は記事の発雷の方向、距離を用いて 決定する。

発雷確率は 20 km 毎に作成するが、その際周囲の雷観 測も含めることで、周囲も含めた 9 格子 (60 km 四方) 内での発雷の有無を目的変数としている。

# (3) 説明変数

説明変数は表 4.7.2 で示す 12 個の中から 6 個を選択して使用する。このうち、FRR3, CAPE, SSI の 3 つは発雷との関係が深いことがわかっていることから、地域ごとのバラつきを抑えて利用しやすいガイダンスとするために必ず使用する必須変数とし、残りの 3 変数を候補変数の中から選択する。候補変数からは、必須変数と候補変数の全組み合わせの回帰式を総当りで試し、 $AIC^{10}$  が最小となる組み合わせの 3 個を選択する。目的変数の対象範囲の予測値を考慮するため、説明変数は各格子の周囲 60 km の平均値とする。また、FRR3 は各格子の周囲 60 km の平均値とする。ま

SSI は、一般には  $850~\mathrm{hPa}$  と  $500~\mathrm{hPa}$  の間で計算した値が用いられるが、発雷確率ガイダンスでは冬季 ( $-10~\mathrm{^{\circ}C}$  高度が  $3~\mathrm{km}$  未満の場合)に限り  $925~\mathrm{hPa}$  と  $700~\mathrm{hPa}$  の間で計算した値 (SSI9) を用いる。これは SSI よりも SSI9 の方が夏季に比べて背の低い冬季の対流雲との対応がよいと考えられ、白山 (2015)による 調査の結果、冬季の利用に有効と判断されたためである。また、 $\mathrm{DXQV}$  は冬季の利用を想定した変数のため、

-10 °C 高度が 5 km 以上の予測式では使用しない。

### (4) 予測式の層別化

予測式は次に示す通り、地域、 $-10\,^{\circ}\mathrm{C}$  高度、予報時間、予報対象時刻( $-10\,^{\circ}\mathrm{C}$  高度が  $5~\mathrm{km}$  以上のときのみ)による層別化を行っており、発雷やモデルの特性に対応した予測を可能にしている。

### 地域層別化

格子毎に予測式を作成すると、精度の高い予測式を作成するために必要なサンプル数を確保できないため、予測対象格子を発雷特性の似た格子同士をまとめた35地域に分割し、地域毎に予測式を作成している。地域分けを図4.7.1に示す。これにより、サンプル数を増やして予測精度の高い回帰式を得やすくすると共に、地域ごとの発雷特性にも対応することができる。この地域分けは、2009年に地域層別化を導入した際に、1年分のPoT予測値をウォード法1 (Ward 1963)を用いてクラスター分けしたものである。

## -10°C 高度による層別化(季節層別化)

-10 °C 高度が 3 km 未満、3-5 km、5 km 以上の 3 段階に対応する予測式をそれぞれ作成している。これにより、季節ごとの特性の異なる雷に対応することができる。-10 °C 高度が 3 km 未満では主に冬季雷を、5 km 以上では主に夏の雷を、3-5 km では主に春と秋の界雷をターゲットにしている。なお、予測対象格子の-10 °C 高度が 1 km 未満の場合は発雷確率を 0% とする。この設定は、-10 °C 高度が低い場合には発雷しないという観測事実に基づいている 12。また、予測式が切り替わる境界での予測値の不連続を軽減するため、-10 °C 高度が

<sup>8</sup> 航空気象定時観測気象報

<sup>9</sup> 航空気象特別観測気象報

<sup>&</sup>lt;sup>10</sup> AIC(Akaike's Information Criterion):赤池情報量基準。詳 細は 第 2.3.8 項を参照。

<sup>11</sup> 階層型クラスター分析の1手法。あるクラスタ同士を結合したときに、結合後の重心までの距離の総和と結合前の各クラスタ内の重心までの距離の総和との差が最小となるクラスタ同士を結合する。

<sup>12</sup> 一例として Michimoto (1993) では、-10°C 高度が 1.4 km 以下の場合に発雷しないとしている。発雷確率ガイダンスでは、一発雷の見逃しを減らすために余裕をもたせた 基準を設定している。

<b>悪</b> 179	<b>発電確率ガイダンスに用いる説田変数</b>	ī

	変数名	説明変数の内容
	FRR3	数値予報モデルの $3$ 時間降水量予測 $[\mathrm{mm/3h}]$ 。 $\log_{10}\left(1+\mathrm{FRR3} ight)$ と変換する。
必須変数	CAPE	対流有効位置エネルギー $[\mathrm{J~kg^{-1}}]$ 。 モデル地上面及び $925~\mathrm{hPa}$ 面から持ち上げた $\mathrm{CAPE}$ のうち大きい方を選択。 $\mathrm{CAPE} \times 1/1000$ と変換する。
	SSI	ショワルター安定指数。 $-10~^\circ\mathrm{C}$ 高度が $3~\mathrm{km}$ 未満の場合には $925~\mathrm{hPa}$ と $700~\mathrm{hPa}$ の間で計算した SSI9 を用いる。
	ZM10	$-10~^\circ\mathrm{C}$ 高度 $[\mathrm{km}]$ (大気の温度が $-10~^\circ\mathrm{C}$ となる高度)。
	LLU	地上から $700~\mathrm{hPa}$ の間の平均風の東西成分 $[\mathrm{m/s}]$ 。 $-10~\mathrm{^{\circ}C}$ 高度が $3~\mathrm{km}$ 以上のときは最大 $15~\mathrm{m/s}$ に制限する。
	LLV	地上から $700~\mathrm{hPa}$ の間の平均風の南北成分 $[\mathrm{m/s}]$ 。 $-10~\mathrm{^{\circ}C}$ 高度が $3~\mathrm{km}$ 以上のときは最大 $15~\mathrm{m/s}$ に制限する。
候補変数	TPWR	気柱相対湿度。鉛直方向に飽和していると仮定した飽和可降水量に対する可降水量の比率。
	VOR5	500 hPa 渦度 [10 <sup>-6</sup> s <sup>-1</sup> ]。最大 99×10 <sup>-6</sup> s <sup>-1</sup> に制限する。
	LAPS	地上から $850~\mathrm{hPa}$ の間の気温減率 $[^{\circ}\mathrm{C~km}^{-1}]$ 。
	DXQV	冬型の降水指数。風向別降水率 $\times 850~\mathrm{hPa}$ の風速 $[\mathrm{m/s}] \times$ 海面と下層温位の飽和比湿差 $[\mathrm{g~kg^{-1}}]$ 。詳細は降水ガイダンス (付録 $4.2.\mathrm{A}$ ) を参照。
	VSHR	850 hPa と 500 hPa の間の風ベクトル差の大きさ [m/s]

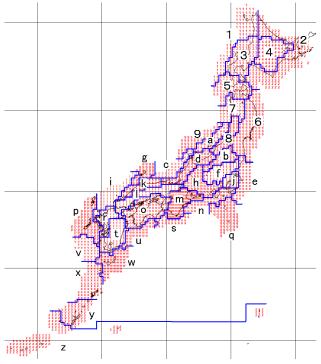


図 4.7.1 予測式作成の区域分け図。図に示す 1-9, a-z の 35 地域毎に予測式の作成を行う。

 $3~\rm km$  未満、 $3-5~\rm km$ 、 $5~\rm km$  以上を対象とする予測式の作成にはそれぞれ  $3.5~\rm km$  未満、 $2.5-5.5~\rm km$ 、 $4.5~\rm km$  以上の発雷事例を用いている。また、予測値を作成する際には、 $-10~\rm ^{\circ}C$  高度が  $4-5~\rm km$  の場合に  $3-5~\rm km$  と  $5~\rm km$  以上の予測式を用いた結果をそれぞれ計算して線形内挿している。

# 予報時間による層別化

GSM-PoT では FT=6-12, 15-24, 27-36, 39-48, 51-60, 63-72, 75-84 の 7 通り、MSM-PoT では

FT=3-9, 12-15, 18-21, 24-27, 30-33, 36-39 の 6 通りで予測式を分けている。これは、予報時間とともに予測精度が下がるという数値予報の特性に応じて、先の予測ほど高確率が出にくい予測式を作成するためである。これにより、先の予測に対しても PoT の信頼度を維持することができる。予報対象時刻による層別化

-10 °C 高度が 5 km 以上の場合は予報対象時刻が午後 3 時から午後 9 時まで (06-12UTC) と午後 9 時から翌日午後 0 時 (12-翌日 03UTC) の 2 つの予測式を作成している。これは、前者について午後の熱雷を区別するためである。午後 9 時 (12UTC)を対象とする予測値は、2 つの予測式を用いて計算した結果の平均値とする。

# (5) LAF 法による重み付け平均

GSM-PoT では過去 2 初期値、MSM-PoT では過去 8 初期値を使って重み付き平均を行っている。表 4.7.3 および表 4.7.4 に各初期値に対する予報時間ごとの重み を示した。GSM-PoT では各初期値の FT=84 までを 利用した重み付け、MSM-PoT では 1, 3, 5, 7 初期値前 は FT=15 まで、2, 4, 6, 8 初期値前は FT=33 までを 利用した重み付けを行う。最新初期値に対する重みを 最も大きくしているが、GSM-PoT の 06(18)UTC 初 期値では1初期値前の00(12)UTC初期値も同じ重み で利用する。また、MSM-PoT では、過去初期値の予 報後半を利用すると精度が悪くなったことから、LAF に利用しない予報時間が多くなっている。この設定は MSM の予報時間延長 (越智・石井 2013) 前から変更 されていないため、より効果的な設定に見直すことは 今後の課題である。第2.2.6項で述べられている通り、 LAF 法は一般的に予測精度の向上が期待できるが、最 新初期値でモデルの予測が大きく変化した場合には、

表  $4.7.3~\mathrm{GSM-PoT}$  の LAF 法で用いる重み (括弧内は 06 及び  $18\mathrm{UTC}$  初期値で用いる重み )。 FT は最新初期値を基準とする。

初期値 \FT	6	9	12	 72	75	78	81	84
最新初期値	2	2	2	 2	2	2	2	2
1 初期値前	1(2)	1(2)	1(2)	 1(2)	1(2)	1(2)		
2 初期値前	1	1	1	 1				

表 4.7.4 MSM-PoT の LAF 法で用いる重み。FT は最新初期値を基準とする。

初期値 \FT	3	6	9	12	15	18	21	24	27	30	33	36	39
最新初期値	25	25	25	25	25	25	25	25	25	25	25	25	25
1 初期値前	22	22	22	22									
2 初期値前	19	19	19	19	19	19	19	19	19				
3 初期値前	16	16											
4 初期値前	13	13	13	13	13	13	13						
5 初期値前													
6 初期値前	7	7	7	7	7								
7 初期値前													
8 初期値前	1	1	1										

LAF 法を適用しない予測値の方が適切な場合もある。

## 4.7.3 統計検証

# (1) GSM-PoT と MSM-PoT の比較検証

GSM-PoT および MSM-PoT の統計的な検証結果を示す。検証は 2016 年 2 月 ~ 2016 年 11 月と 2016 年 12 月 ~ 2017 年 1 月の 2 つの期間に分けて実施した。これは、期間全体の発雷のうち日本海側で発生する冬季雷が占める割合が多くなる 12 月 ~ 1 月については検証結果の傾向が大きく異なるためである。検証対象とする格子は、予測式作成に使用する格子とした 13 。数値予報モデルには 2018 年 3 月現在で最新の GSM および MSM を用いた。検証に用いる予測値の作成にあたり、GSM の全期間、MSM の 2016 年 2 月 ~ 11 月がそれぞれ GSM-PoT,MSM-PoT の係数作成期間であること から、当該期間は 1 か月抜きクロスバリデーション(第 2.3.11 項 (2) 参照)による検証結果を示す。

検証指標にはエクイタブルスレットスコア (ETS) と ブライアスキルスコア (BSS) を用いた。検証対象の初期値は、GSM-PoT が 00, 06, 12, 18UTC, MSM-PoT が 03, 09, 15, 21UTC であり、ETS 及び月別の BSS では、GSM-PoT は FT= $12\sim33$  を対象とし、予報時間別の BSS は、それ ぞれ FT= $6\sim84$ , FT= $3\sim39$  の 3 時間毎を対象とした。また、予報時間別の信頼度を、予報時間の層別化単位を対象として、GSM-PoT は FT= $12\sim84$  を 12 時間毎、MSM-PoT は FT= $9\sim39$  を 6 時間毎に示した。

# 春~秋季(2月~11月)

はじめに、検証期間 2016 年 2 月 ~ 11 月について結 果を示す。図 4.7.2 に GSM-PoT と MSM-PoT の月別

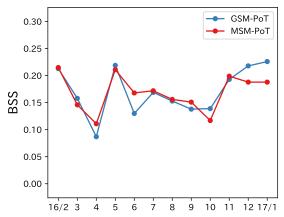


図 4.7.2 2016 年 2 月から 2017 年 1 月までの月別の BSS。 横軸は (年/) 月。青線が GSM-PoT、赤線が MSM-PoT を示す。

の BSS の比較図を示す。2 月  $\sim$  11 月は月ごとのバラつきはあるものの、概ね MSM-PoT の予測精度が高いことがわかる。

図 4.7.3(a) に閾値別の ETS を示す。この図から、GSM-PoT, MSM-PoT ともに ETS がピーク値をとる確率値は 15%であり、また、この付近では MSM-PoT の方が ETS の値が大きいことがわかる。これは、GSM-PoT, MSM-PoT ともに 15%を閾値として発雷の有無を判断する場合に最も精度がよく、また、MSM-PoT の方がその予測精度が高いことを意味する。この判断に利用する閾値は 白山 (2015) で示されている 20%より低い値となっているが、モデル変更により予測特性が変化したためと考えられる。一方で、高確率を閾値とした場合は MSM-PoT の値が GSM-PoT の値よりも低くなっている。また、本検証は全国を対象としているが、地方別にはピークをとる確率値が 15%でない地域もあり、例えば関東甲信から中部の GSM-PoT、九

<sup>13</sup> 海岸線から離れた格子は予測式作成に用いていないため検 証対象外とした。

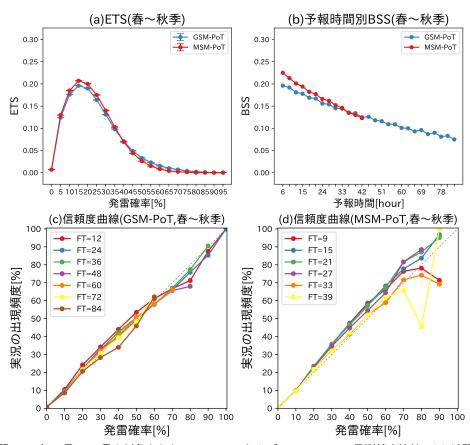


図 4.7.3 検証期間 2016 年 2 月 ~ 11 月を対象とした GSM-PoT および MSM-PoT の予測精度比較。(a) は閾値別の ETS、(b) は予報時間別の BSS で、青線で GSM-PoT、赤線で MSM-PoT を示す。(c) 及び (d) は予報時間別の信頼度曲線で、(c) は GSM-PoT を FT=12 から FT=84 まで 12 時間毎、(d) は MSM-PoT を FT=9 から FT=39 まで 6 時間毎に示す。予報時間別の BSS では、対象時刻を合わせるために GSM-PoT の予報時間を基準に表示している。

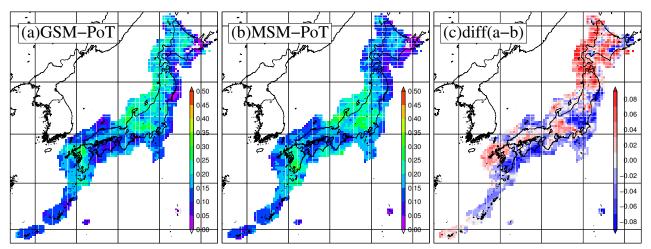


図 4.7.4 検証期間 2016 年 2 月 ~ 11 月を対象とした GSM-PoT および MSM-PoT の BSS 分布図とその差分図。(a) は GSM-PoT, (b) は MSM-PoT の BSS 分布図、(c) はその差分図を示し、(c) では赤色ほど GSM-PoT の方が予測精度が高く、青色ほど MSM-PoT の方が予測精度が高いことを示す。

州の GSM-PoT, MSM-PoT で 20%となっている。

図 4.7.3(b) に予報時間別の BSS を、(c), (d) にそれぞ れ GSM-PoT 及び MSM-PoT の予報時間別の信頼度 曲線を示す。予報時間別の BSS の図から、MSM-PoT の FT=30 程度までは MSM-PoT の方が GSM-PoT よ りも予測精度が高く、それ以降は両者に大きな差はな いことが確認できる。また、予報時間別の信頼度曲線 の図から、GSM-PoT の信頼度曲線は傾き 45° の理想 直線上に乗っており、予報時間によらず信頼度が高い ことがわかる。一方で MSM-PoT は、直線より上にな る傾向があり、実況の出現確率よりも PoT を低めに 予測する傾向が見られる。この傾向は特に高確率側で 顕著に見られ、高確率の信頼度が低いことを示してお り、このために MSM-PoT の ETS が高確率側の閾値 で低くなっていると考えられる。なお、FT=9 および FT=33 の確率値 90%や FT=39 の確率値 80%以上で は直線から大きく離れているが、これは検証対象の発 雷事例数が非常に少ないためである。

図 4.7.4(a), (b) 及び (c) にそれぞれ、20 km 格子毎に検証した GSM-PoT 及び MSM-PoT の BSS の分布図、その差分図を示す。それぞれの検証図を見ると、GSM-PoT は北海道や東北北部、中部、九州北部などで、MSM-PoT は関東や九州南部などで予測精度が高くなっている。MSM-PoT は北海道~東北北部で高確率の予測頻度が少ない傾向が特に強くあり、GSM-PoT と比較して BSS が低いことの一因となっている。

以上示した通り、春~秋は統計的には MSM-PoT の方が GSM-PoT よりも予測精度は高いが、高確率の信頼度が低いことに注意する必要がある。予測分布が GSM-PoT と近いが確率値が大きく異なるような場合には、信頼度の高い GSM-PoT も参考にして、発雷の有無を判断していただきたい。

## 冬季(12月~1月)

次に、検証期間 2016 年 12 月~2017 年 1 月について結果を示す。この期間は両ガイダンスともに年によらず安定して統計的な予測精度が高い特徴があり、図 4.7.2 の月別の BSS の比較図からわかる通り GSM-PoT の方が予測精度が高い。また、図 4.7.5 に示す通り、この期間は日本海側で発生する冬季雷が全体の 7 割を占めることが特徴である。

図 4.7.6(a) に閾値別の ETS を示す。ETS のピーク値は GSM-PoT が MSM-PoT を上回っており、GSM-PoT の方が予測精度が高いことがわかる。目安となる確率値は GSM-PoT, MSM-PoT ともに 15%であり、他の季節と同様に利用できる。また、ピーク値をとる確率値より高確率側でも MSM-PoT よりも GSM-PoT の ETS が高くなっている。

図 4.7.6(b) に予報時間別の BSS, (c) 及び (d) にそれ ぞれ GSM-PoT 及び MSM-PoT の予報時間別の信頼 度曲線を示す。予報時間別の BSS の図からは、全予報

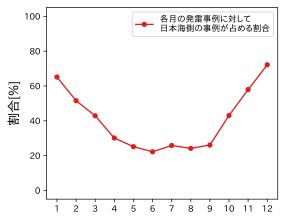


図 4.7.5 日本海側の発雷事例の全国の発雷事例に占める割合 [%]。横軸は月。2014 年から 2017 年までの 4 年間の平均値。

時間を通して GSM-PoT の予測精度が MSM-PoT を上回っていることが確認できる。また、予報時間別の信頼度曲線の図からは、両ガイダンスとも傾き  $45^\circ$  の理想直線からは離れており、また、予報時間によらず実況の出現確率よりも PoT が低い傾向であることがわかる。MSM-PoT では GSM-PoT に比べて高確率側ほどこの傾向が顕著であり、信頼度が低くなっている。これは ETS の図で GSM-PoT よりも高確率の予測精度が低いことと整合する。

図 4.7.7(a), (b) 及び (c) にそれぞれ、20 km 格子毎に検証した GSM-PoT 及び MSM-PoT の BSS の分布図、その差分図を示す。両ガイダンスとも同様の分布で、この期間に雷が多く発生する日本海側で予測精度が高いことがわかる。一方で、日本海側以外の地域では、BSS は低く、負値となっている地域も目立つ。事例検証でも取り上げるが、発雷確率ガイダンスは予測手法に統計手法を用いているため、冬季に発雷事例数の多い日本海側は予測精度が高いが、事例数の少ない内陸や太平洋側の発雷に対しては実況に即した PoT を予測しにくく、予測精度は低い。また、差分図からは全国的に MSM-PoT よりも GSM-PoT の予測精度が高いことがわかる。

以上示した通り、冬季の発雷はそのほとんどが日本海側で発生する。その予測精度は両ガイダンスともに高いが、MSM-PoT は他の季節と同様に高い確率の予測頻度が少ない傾向が強く、GSM-PoT の方が MSM-PoT よりも予測精度が高い。ただし、発雷確率ガイダンスは統計手法を用いたガイダンスのため、事例数の少ない内陸部や太平洋側など日本海側以外の発雷に対しては高い確率を出しにくく、予測精度は低い。

## (2) LAF 法の検証結果

図 4.7.8 に LAF を適用した場合 (すなわち PoT 予測値) としない場合について精度検証を行った結果を示す。検証期間や予報時間などの検証条件は (1) に同じ

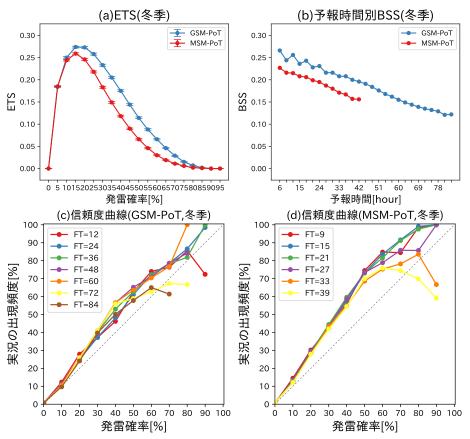


図 4.7.6 検証期間 2016 年 12 月 ~ 2017 年 1 月を対象とした GSM-PoT および MSM-PoT の予測精度比較。図の説明は図 4.7.3 と同じ。

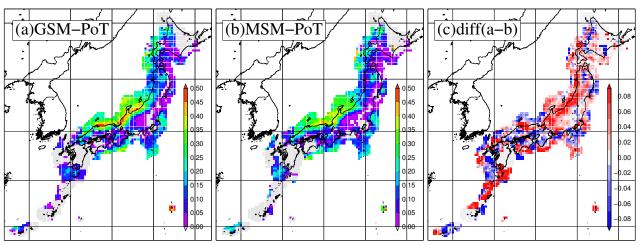


図 4.7.7 検証期間 2016 年 12 月~2017 年 1 月を対象とした GSM-PoT および MSM-PoT の BSS 分布図とその差分図。図の 説明は図 4.7.4 と同じ。

である。GSM-PoTではLAFの適用前後で予測精度に大きな違いはないが、主に高確率側の信頼度を改善していることがわかる。一方でMSM-PoTはLAFを適用することで、特に予報期間の前半で予測精度が低下していることがわかる。また、信頼度については、

LAF を適用することにより高確率側ほど予測頻度が少なくなっている。MSM-PoT の予測精度、信頼度の低下については、第4.7.5項で詳しく述べるが、係数の再作成を行うことで改善されると考えている。ただ、本来のLAF 法の目的である予測精度の向上は両ガイダン

スについても得られていない。PoT の LAF 法は 2008 年の 高田 (2008) による導入以降大きな変更はなされていないため、最新のモデル特性に応じてより効果的な設定に見直す必要があり、今後の課題である。

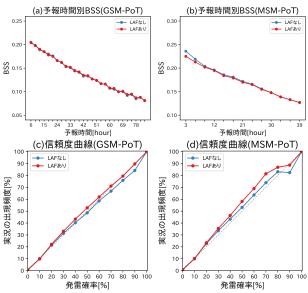


図 4.7.8 LAF を適用した場合としなかった場合の予測精度と信頼度の比較。検証期間は 2016 年 2 月 ~ 2017 年 1 月。
(a) および (b) は予報時間別の BSS で (a) に GSM-PoT、(b) に MSM-PoT の検証結果を示す。(c) および (d) は信頼度曲線で (c) に GSM-PoT、(d) に MSM-PoT の検証結果を示す。それぞれ青線に LAF 適用なし、赤線に LAF 適用ありの結果を示す。

#### 4.7.4 事例検証

ここではいくつかの事例を通して、 $\operatorname{PoT}$  の予測特性 や利用上の留意点を示す。なお、事例 (1) から (3) までの天気図はまとめて図 4.7.13 に示す。

#### (1) PoT の一般的な予測特性

まず、PoT の一般的な予測特性を比較的よく示している事例を 2 例紹介する。PoT はモデルで降水が予測されており  $^{14}$ 、大気の安定度が低い場合に高確率となる。ガイダンスが高確率を予測している場合や GSM-PoT と MSM-PoT の予測値の差が大きい場合など、その予測の妥当性を判断する必要がある場合には、モデルの降水及び大気安定度の妥当性を考慮することが有効となる。

図 4.7.9 は 2016 年 6 月 21 日 06UTC を対象とした 事例である。日本の南の前線上で発雷が観測されてお り、その一部が九州南部にかかっている。九州南部の発 雷は両ガイダンスとも予測しているが、GSM-PoT は 東海地方や紀伊半島にも高めの予測をしている。PoT の予測値は主に説明変数の FRR3, CAPE, SSI の値、 すなわちモデルの降水量及び大気安定度に大きく左右 されるが、この事例ではモデル間でこれらの予測が異なっていたことが、GSM-PoTとMSM-PoTとの差につながっている。九州南部では両モデルとも降水を予測しており、かつ、大気安定度も低いため、両ガイダンスともに高めの PoT が予測されている。一方、東海地方では両モデルとも降水を予測しているが、MSMでは大気安定度が高いため、MSM-PoT の予測値は低い。また、紀伊半島では MSM は大気安定度が比較的低い地域もあるが、降水が予測されていないため PoT は低い。

図 4.7.10 は 2016 年 6 月 30 日 00UTC を対象とした事例で、九州北部に前線があり、その暖域側の九州南部で発雷している。両ガイダンスとも九州に発雷を予測しているが、GSM-PoT は九州北部、MSM-PoT は九州南部にそれぞれ高い確率を予測している。予測に差が生じた理由は降水予測の差にある。GSM は九州北部を中心に、MSM は九州南部を中心にそれぞれ強い降水を予測している。大気安定度の定性的な分布はモデル間で大きな差はなく、降水の予測位置のズレがそのまま PoT の予測の差となっている。

## (2) 稀少事例に対する予測特性

次に、発雷確率ガイダンスが統計的な確率予測手法 を用いているゆえの予測特性を紹介する。図 4.7.11 は 2017年1月14日21UTCを対象としたGSM-PoTの 予測事例で、発雷事例数が少ない稀少事例に対する予 測特性を示している。日本付近は強い冬型の気圧配置 となり、山陰から近畿にかけての日本海側や東北地方 で発雷した。GSM は山陰から近畿にかけてと東北のど ちらにも降水を予測しており、また、同程度の大気安 定度を予測している。しかし、GSM-PoT は山陰から 近畿にかけてのみ高い確率を予測し、東北地方はほと んど予測していない。PoT の予測は統計手法を用いて いるため、観測数が少ない発雷事例に対しては高確率 を出しにくい。冬季の東北地方 (特に太平洋側)は発 雷が少ないことから、高い PoT は予測されなかったと 考えられる。そのほかにも稀少事例として春の熱雷な どがあげられる。

# (3) 春先の寒気南下時の予測事例

2 月~4 月に発生する日本海側の雷は、日本海の海面水温が低下するため、大規模な発雷に至らないことも多い。しかし、-10 °C 高度が 3 km 以下の予測式は 12 月~1 月の大規模な冬季雷に最適化されているため、PoT が広範囲に高い確率で予測されて空振りとなることがある。図 4.7.12 は 2017 年 4 月 12 日 21 UTC を対象とした GSM—PoT の予測事例である。図から、北陸地方は-10 °C 高度が 3 km 以下であることがわかる。このとき北陸から東北地方にかけて大気安定度が低くなることが予測され、広い範囲に高い PoT を予測した。しかし、実際にはこの予測時刻の後に新潟県で数回程度の発雷があったのみで多くの地域では空振

<sup>14</sup> モデル降水量は説明変数として利用する際に対数変換されるため、値の大小よりも降水の有無による影響が大きい。

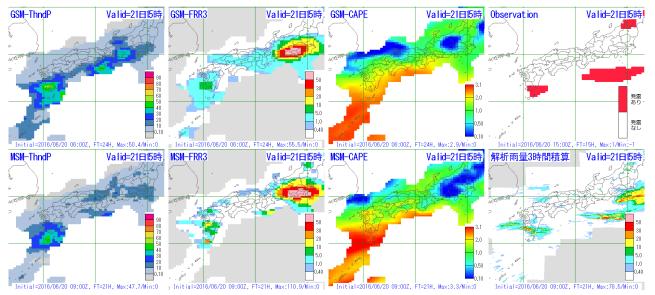


図 4.7.9 2016 年 6 月 21 日 06UTC を対象とした予測事例。上段左から GSM-PoT[%]、説明変数の GSM の FRR3[mm/3h],  $CAPE[10^3 J/kg]$ 、発雷実況、下段左から MSM-PoT[%]、説明変数の MSM の FRR3[mm/3h],  $CAPE[10^3 J/kg]$ 、前 3 時間解析雨量 [mm/3h]。発雷実況は赤い格子が発雷ありを示す (以降の図でも同様)。

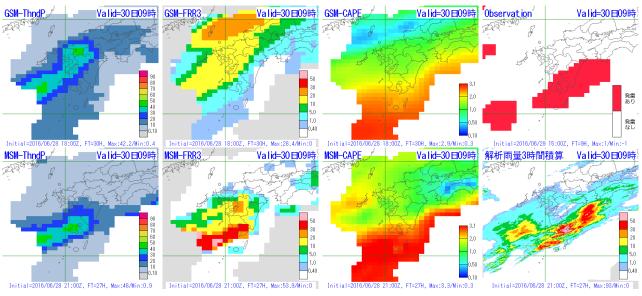


図 4.7.10 2016 年 6 月 30 日 00UTC を対象とした予測事例。図の配置は図 4.7.9 と同じ。

りとなった。海面水温を考慮した説明変数に DXQV があるが、北陸から東北地方の日本海側の予測式には採用されていない。今後利用方法の再検討や他に有効な説明変数を模索する必要がある。

## (4) 予報時間により予測確率が変化する例

PoT の予測式は予報時間によって層別化している。そのため、予報時間後半ほど数値予報モデルが持つ予測の不確実性が考慮され、高確率が予測されにくくなっている。図 4.7.14 は 2016 年 8 月 1 日 09UTC を対象とした GSM-PoT の予測事例で、左から FT=69, 45, 21 の予測値で、右側ほど予報対象時刻に近く、新しい初期時刻の予測値を示している。初期値の変化に対し

モデルの予測値の変化は小さいが、実況で雷が観測された地域の PoT は次第に高い確率値に変化していることがわかる。このように、予報時間が短くなり数値予報の精度が上がるにつれて、同じ大気安定度でも次第に高い PoT が予測されるようになり、実況との対応が良くなる。

# (5) MSM-PoT の予測過小事例

最後に、2018 年現在の PoT が抱える問題点として、MSM-PoT の予測過小事例を紹介する。図 4.7.15 は、2016 年 5 月 31 日 09UTC を対象とした予測事例で、上空に寒気を伴った低気圧が通過し北日本の広い範囲で発電した事例である。GSM-PoT と MSM-PoT の予測

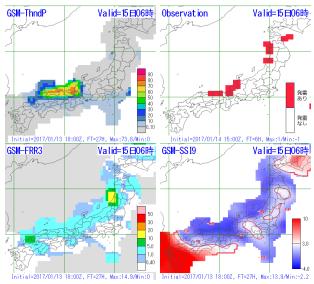


図 4.7.11 2017 年 1 月 14 日 21UTC を対象とした予測事例。上段左から GSM-PoT[%]、発雷実況、下段左から説明変数の GSM の FRR3[mm/3h] と SSI9。

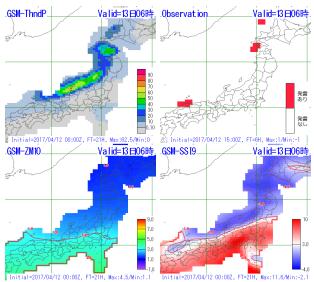


図 4.7.12 2017 年 4 月 12 日 21UTC を対象とした予測事例。上段左から GSM-PoT[%]、発雷実況、下段左から説明変数の GSM の -10 °C 高度 [km] と SSI9。

値を比較すると、どちらも北海道から東北にかけて発雷を予測しているが、MSM-PoT はその範囲が狭く、値も小さい。これは主に CAPE と SSI の差によるものである。それぞれの図を見ると、GSM と比べて MSMでは CAPE が小さく、SSI は大きいことから、大気安定度が高くなっている。MSM は 2017 年 2 月のモデル更新 (原 2017) により大気安定度をより適切に表現するようになった。今回の事例では更新前と比べてより大気安定度を高める方向に変化しているが、本来は PoTの予測式を更新することでこの特性変化に対応できるはずである。2018 年現在の MSM-PoT は、予測式を更新する際に最新モデルの予測値のみでは開発データが

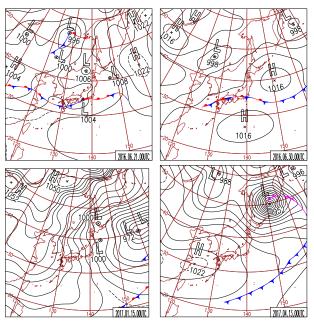


図 4.7.13 事例 (1) から (3) の日本域地上天気図。上段左から 2016 年 6 月 21 日 00UTC、2016 年 6 月 30 日 00UTC、下段左から 2017 年 1 月 15 日 00UTC、2017 年 4 月 13 日 00UTC。

足りなかったため、モデル更新前の MSM の予測値を含めて作成された予測係数を用いている。今後最新モデルの予測値が十分に蓄積された段階で係数の再作成を実施する予定であるが、それまでは MSM-PoT が低い確率となる傾向に留意いただきたい。また、第 4.7.3 項 (1) でも述べた通り、春から秋の北海道~東北北部では MSM-PoT が予測過小となる傾向が特に強いため、注意が必要である。

# 4.7.5 利用上の留意点と今後の展望

本項では、これまでに示した利用上の留意点をまとめて紹介し、最後に今後の展望を述べる。利用上の注 意点として以下の通り紹介した。

- 第 4.7.3 項では GSM-PoT 及び MSM-PoT の予測 精度について、春季~秋季は MSM-PoT の方が高 く、冬季は GSM-PoT の方が高いことを示した。 ただし、MSM-PoT は予測頻度が過小の傾向があ るため、低めの確率値となることに注意して利用 する必要がある。特に、春から秋の北海道~東北 北部ではこの傾向が強いことから注意する必要が ある。
- 第 4.7.4 項 (1) では、PoT の予測値が説明変数の FRR3, CAPE, SSI の値に大きく左右されることを示した。GSM-PoT と MSM-PoT の予測値が大きく異なり両ガイダンスの妥当性を判断する必要があるようなときには、各モデルの降水予測の有無及び大気安定度の予測を確認することが有効である。

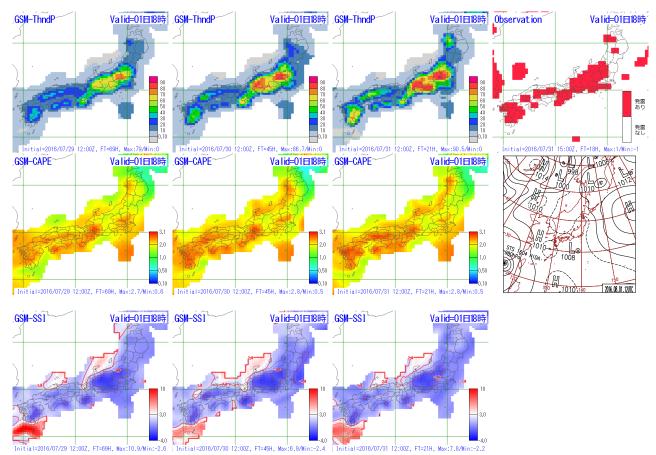


図 4.7.14 2016 年 8 月 1 日 09UTC を対象とした予測事例。上段に GSM-PoT[%] および発雷実況、中段に説明変数の GSM の  $CAPE[10^3 J/kg]$  および 2016 年 8 月 1 日 12UTC の日本域地上天気図、下段に説明変数の GSM の SSI を示す。左から順に FT=69、FT=45、FT=21 の予測値で、右側ほど新しい初期値の予測となっている。

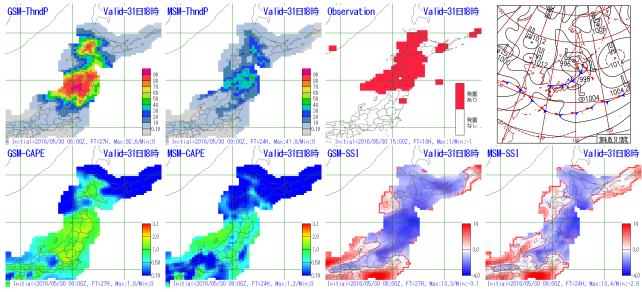


図 4.7.15 2016 年 5 月 31 日 09UTC を対象とした予測事例。上段左から GSM-PoT[%]、MSM-PoT[%]、発雷実況、2016 年 5 月 31 日 12UTC の日本域地上天気図、下段左から説明変数の GSM および MSM の  $CAPE[10^3 J/kg]$  と SSI。

- PoT は予測手法に統計手法を用いているため、第4.7.4項(2)で示したような稀少事例、第4.7.4項(3)で示したような季節外れの事例については予測が難しい。
- 第4.7.4項(4)に示した通り、PoTは予報時間が先の予測では、高確率を予測しにくくすることで信頼度を維持している。モデルの予測傾向が変わらない場合でも、予報時間が短くなるにつれ高確率に変わっていく場合があるので留意が必要である。

続いて今後の展望を述べる。まずは MSM-PoT の予 測過小傾向を改善する必要がある。これについては、 MSM 更新に伴う予測式の再作成時に開発データを増 やすために更新前のモデルの予測事例を利用したこと が原因であり、予測式の作成を更新後のモデルのみを 用いて行うことで改善が見込まれる。当時は更新後の モデルのみでは開発データが足りず、精度が出ないと いう事情があったが、2018年現在で更新後のモデルの 予測値も蓄積されてきていることから、なるべく早い 時期の対応を検討している。このように、係数固定型 である発雷確率ガイダンスは、モデル更新に伴って予 測特性に大きな影響を受けることがある。第 4.7.1 項 でも述べた通り、逐次学習型の手法を利用することは 難しいが、モデル更新の影響の軽減について検討して いく必要がある。また、第 4.7.3 項 (2) で述べた通り、 LAF の重み付けのより適切な設定についても検討する 必要がある。そして、これまでの精度向上の取り組みに より信頼度は大きく向上し、利用しやすくなってきて いるが、予測精度についてはまだ改善の余地があると 考えている。例えば、説明変数の CAPE は地上または 925 hPa を基準とした大気の不安定指標として利用し ているが、これよりも上層の大気の不安定状態と発雷 が関係していることがある。より上層からの持ち上げ についても考慮することで、上層の不安定による発雷 の見逃しを減らせる可能性がある。このほか、DXQV など既存の説明変数の計算手法の見直しや精緻化、新 たな説明変数の利用調査についても継続して取り組む 必要がある。

### 参考文献

- 後藤尚親, 2016: ガイダンスの特性の変化. 平成 28 年 度数値予報研修テキスト, 気象庁予報部, 36-41.
- 後藤尚親, 2017a: GSM 改良のガイダンスへの影響. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 31-34.
- 後藤尚親, 2017b: MSM ガイダンスの特性の変化. 平成 29年度数値予報研修テキスト, 気象庁予報部, 56-60.
- 原旅人, 倉橋永, 2017: メソ数値予報システムの特性の 変化. 平成 29 年度数値予報研修テキスト, 気象庁予 報部, 48-55.
- 気象庁予報部, 1986: 短期間予報用ガイダンスの解説. 予報技術資料第 35 号, 気象庁予報部.

- Michimoto, K., 1993: A Study of Radar Echoes and their Relation to Lightning Discharges of Thunderclouds in the Hokuriku District. Part II: Observation and Analysis of "Single-Flash" Thunderclouds in Midwinter. J. Meteor. Soc. Japan, 71, 195–204.
- 越智健太, 石井憲介, 2013: 領域拡張・予報時間 39 時間化されたメソモデルの特性. 平成 25 年度数値予報研修テキスト, 気象庁予報部, 1-17.
- 白山洋平, 2015: 発雷確率ガイダンスの改良. 平成 27 年度数値予報研修テキスト, 気象庁予報部, 64-72.
- 高田伸一, 2007: 航空気象予報ガイダンス. 平成 19 年度数値予報研修テキスト, 気象庁予報部, 87-93.
- 高田伸一, 2008: 発雷確率ガイダンス. 平成 20 年度数値予報研修テキスト, 気象庁予報部, 88-89.
- 高田伸一, 2009: 発雷確率ガイダンス. 平成 21 年度数値予報研修テキスト, 気象庁予報部, 39-43.
- 田沢秀隆, 1996: 予報支援資料. 平成7年度量的予報研修テキスト, 気象庁予報部, 43.
- Ward, J. H., Jr., 1963: Hierarchical Grouping to Optimize an Objective Function. *Journal of the Amer*ican Statistical Association, 58, 236–244.

### 4.8 湿度ガイダンス 1

### 4.8.1 はじめに

2018年現在、湿度ガイダンスとして、最小湿度と時系列湿度ガイダンスがある。最小湿度ガイダンスでは日最小湿度を、時系列湿度ガイダンスでは毎正時の湿度に加えて日平均湿度・実効湿度  $^2$  を予測している。これらの湿度ガイダンスは乾燥注意報の発表を支援するプロダクトとして利用されている。

最小湿度ガイダンスは 1990 年に運用を開始し(第 1.3 節)、1996 年に線形重回帰からニューラルネットワークへ統計手法を変更した (柳野 1995)。また、2001年3月には作成対象地点数を大幅に増やし、それまではニューラルネットワークの重み係数を固定としていたが、逐次学習により日々更新するように変更した (佐藤 2000)。2009年にはそれまでの GSM 最小湿度ガイダンスに加えて MSM 最小湿度ガイダンスの運用も開始した (國次 2009)。時系列湿度ガイダンスは、乾燥注意報の発表支援などを目的に 2016年8月から運用を開始した (高桑 2015)。

本節では、湿度ガイダンスの仕様、統計検証、事例検証や利用上の留意点と今後の展望について述べる。以降、 $\mathrm{GSM}\ (\mathrm{MSM})$  湿度ガイダンスを  $\mathrm{GSM}\ (\mathrm{MSM})$  ガイダンスと記述する。

### 4.8.2 仕様

湿度ガイダンスは、全国の気象官署 (特別地域気象観測所を含む)について、日最小湿度、時系列湿度、日平均湿度および実効湿度の予測を行っている。GSM (MSM) ガイダンスは  $00\sim18$ UTC ( $00\sim21$ UTC) の 6時間毎 (3 時間毎)に作成される。本項では、最小湿度ガイダンス、時系列湿度ガイダンスの順に仕様・作成方法を述べる。

### 最小湿度ガイダンス

最小湿度ガイダンスの仕様を表 4.8.1 に示す。最小湿度ガイダンスは統計手法にニューラルネットワーク (第 2.6 節)を用いた逐次学習型のガイダンスで、日最小湿度の観測データを用いて学習を行う。使用しているニューラルネットワークは、3 層の順伝播型で入力層 (23 ユニット)、中間層 (8 ユニット)、出力層 (1 ユニット)で構成され、中間層の活性化関数はロジスティック関数、出力層の活性化関数は線形関数を利用する。表 4.8.1 に示したように、モデルの地上の最小湿度・比湿・気温・最高気温の他に上空の湿度や風速なども用いており、GSM・MSM ガイダンスで同じ説明変数としている。これらの説明変数は対象地点の周囲4格子から線形内挿することで計算する。

予報対象日について、GSM (MSM) ガイダンスは最長で3日後(翌日)まで予測する。予報対象日やその日数が初期時刻によって異なるのは、数値予報モデルの予報時間の制約によるものである。なお、18UTC 初期値の MSM ガイダンスでは日最小湿度は予測しない。これは、1日を通した説明変数が得られないからである。

ニューラルネットワークの重み係数は、地点と季節で層別化している。季節は、夏期(4月~9月)と冬期(10月~3月)としている。ただし、季節の変わり目で予測結果が大きく変わることを防ぐために、重み係数の利用を開始する30日前から学習を行う。重み係数の学習は、GSM (MSM) ガイダンスは12UTC (09UTC) に行い、学習には前日に観測された日最小湿度を用いる。また、すべての予報対象日で同じ重み係数を用いる。

### 時系列湿度ガイダンス

時系列湿度ガイダンスの仕様を表 4.8.2 に示す。時系列湿度ガイダンスは毎正時の湿度(時系列湿度)と日平均湿度・実効湿度を予測する。なお、18,21UTC 初期値の MSM ガイダンスでは、日平均湿度と実効湿度の予測を行っていない。これは、数値予報モデルの予報時間の制約によるものである。

### (1) 時系列湿度

時系列湿度ガイダンスは、時系列湿度の予測にカルマンフィルタ(第 2.7 節)を用いている。目的変数は湿度の観測値と数値予報モデルの予測値との差とし、その差の予測値に数値予報モデルの予測値を足すことで湿度の予測結果とする。説明変数は、地上相対湿度、地上の東・西・南・北風成分、地上風速、中・下層量、前 1 時間降水量、海面更正気圧を用いている。地上の各風向の風速、地上風速は、複数の格子点を用いることで風向の違いにより各風速が過小に評価されることを防ぐため、作成対象地点に最も近い格子のデータを使用する。これら以外の説明変数はモデルにより計算方法が異なり、GSM ガイダンスの場合は作成対象地点の周囲 4 格子を用いて数値予報モデルの予測値を線形内挿し、MSM ガイダンスの場合は作成対象地点周辺の海陸分布 3 を考慮して線形内挿している。

時系列湿度ガイダンスは前述の最小湿度ガイダンスと独立して計算しており、両者の整合性をとる必要があるため、時系列湿度ガイダンスの結果が予報対象日における最小湿度ガイダンスの予測結果を下回った場合には時系列湿度の値を最小湿度の値で置き換える。

予測式は、地点、初期時刻、予報時間で層別化し、係数の学習は毎初期値行う。初期時刻と予報時間で層別化することによって、対象時刻と予報時間で変わるモデルの誤差特性に対応することができる。

<sup>1</sup> 小林 大輝

<sup>&</sup>lt;sup>2</sup> 実効湿度は木材の乾燥の程度を示す指標で、数日前からの 湿度が考慮されており、日最小湿度とともに乾燥注意報の発 表基準となっている。

 $<sup>^3</sup>$  MSM で利用されている地形データの、モデル格子における陸の割合 ( $0{\sim}1$ ) から、陸の割合が 0.5 以上の場合は 1 (陸格子) 小さい場合は 0 (海格子) としたもの。

表 4.8.1 最小湿度ガイダンスの仕様一覧。

作成対象	気象官署(	気象官署 (特別地域気象観測所含む)					
作成方法	ニューラル	ニューラルネットワーク(第 2.6 節)					
作成対象とするモデル	GSM, MS	GSM, MSM					
予報対象時間単位	24 時間 (15	5~15UTC)					
	使用する数値予報モデルと予報対象日は以下のとおり。 「翌日」等は、日本時間で初期時刻から見た日付を示す。						
	モデル	初期時刻	当日	翌日	2 日後	3 日後	
		00UTC					
	GSM	06UTC		0	0	0	
		12UTC		0	0	0	
		18UTC		0	0		
初期時刻と予報対象日	MSM	00UTC					
		03UTC		0			
		06UTC		0			
		09UTC		0			
		12UTC		0			
		15UTC	0				
		18UTC					
		21UTC					
逐次学習の有無	あり						
	予報対象日における、 03, 12, 21JST の地上気温、03, 12, 21JST の 850 hPa 風速、						
 説明変数	03, 12, 21JST の 3 層 (1000, 925, 800 hPa) 平均相対湿度、						
D/M/10X XX		JST の 1000					
		hPa の日最 :、1000, 925					湿、地上最小湿度、 対湿度
目的変数	観測された	観測された日最小湿度(1分値から算出)					
層別化処理の対象	作成対象地	作成対象地点、夏期(4月~9月)、冬期(10月~3月)。					

### (2) 日平均湿度・実効湿度

(1) で得られた時系列湿度の予測結果と最新の観測 値を用いて日平均湿度・実効湿度の予測を行っている。 湿度ガイダンスでは実効湿度を以下の式で定義する。

対象日の実効湿度 = 対象日前日の実効湿度  $\times$  0.7 + 対象日の日平均湿度  $\times$  0.3

(4.8.1)

まず、 初期時刻日の日平均湿度を、観測値および時系列湿度の予測値から求める。 初期時刻日の実効湿度を、観測値から得られた初期時刻日前日の実効湿度と の日平均湿度から求める。次に 翌日以降の実効湿度を時系列湿度の予測値から計算した日平均湿度とで求めた実効湿度から求める。ただし、GSM ガイダンスには FT=1,2の予測値がないため、当日の日平均湿度には観測値と FT=3 以降の時系列湿度の予測値をあわせて22時間分の平均値を使用する。また、MSM ガイダンス15UTC 初期時刻 (00JST) では日本時間で当日の予測をするため、前日の湿度の観測データを利用する。

### 4.8.3 統計検証

本項では、夏期間 ( 2016 年 6 月  $\sim$  8 月 ) 8 期間 ( 2016 年 12 月  $\sim$  2017 年 2 月 ) 及び通年 ( 2016 年 1 月  $\sim$  12 月 ) における湿度ガイダンスについてモデルとの精度比較、期間毎の GSM ガイダンスと MSM ガイダンスの比較を示し、最後に地点毎の検証結果を示す。なお、数値予報モデルには 2018 年 3 月現在運用しているものと同仕様の予測値を用いた。

### 最小湿度ガイダンス

まず、12UTC 初期値における全作成対象地点で平均した GSM 及び MSM ガイダンスの二乗平均平方根誤差 (RMSE) と平均誤差 (ME) を期間毎に図 4.8.1 に示す。ME は正バイアスの傾向があり、MSM ガイダンスはその傾向が強い。RMSE は、GSM 及び MSM ガイダンスともに通年の方が夏期間や冬期間と比べて大きくなっている。これは、季節の変わり目で梅雨前線や秋雨前線の影響により、最小湿度の変動が大きくなることと、数値予報モデルにおいて梅雨前線や秋雨前線の位置の予測が難しいことが原因と考えられる。また、MSM ガイダンスは夏の方が冬より精度が高いが、GSM ガイダンスは夏の方が精度が低くなっている。不安定性降水などの予測精度の違いが原因の一つとして

表 4.8.2 時系列湿度ガイダンスの仕様一覧

作成対象	気象官署(特別地域気象観測所含む)						
統計手法	カルマンフィルタ (第 2.7 節 )						
予報対象とするモデル	GSM, MS	M					
予報対象時間	毎正時						
予報期間と間隔		=3 から FT =1 から FT					
逐次学習の有無	あり						
説明変数	モデル予測値(地上相対湿度、地上の東・西・南・北風成分、地上風速、中・下層雲量、前1時間降水量、海面更正気圧) ただし、各風成分と地上風速は作成対象地点に最も近い格子点のモデル予測値を使用し、 それ以外は GSM では周囲の 4 格子から線形内挿、MSM ではモデルの海陸分布を考慮 した内挿値を用いる。						
目的変数	湿度観測値	[ – モデルの	湿度予	則値			
層別化処理の対象	作成対象地	点、初期時	刻、予輔	時間			
備考	日平均湿度	正時の湿度 と実効湿度 は、日本時 初期時刻 00UTC 06UTC 12UTC 18UTC	の初期時	対に対	する予報	対象日は以	<b>削している。</b> 人下のとおり。
	MSM	00UTC 03UTC 06UTC 09UTC 12UTC 15UTC	0	0 0 0 0			
		21UTC					

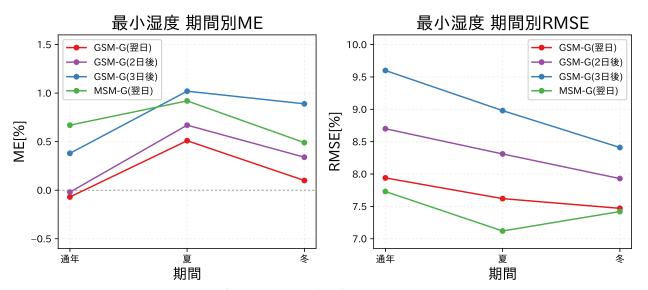


図 4.8.1 12UTC 初期値における GSM 及び MSM 最小湿度ガイダンスの期間別 ME ( 左 )・RMSE ( 右 )。GSM-G (MSM-G) は GSM (MSM) ガイダンスを示す。

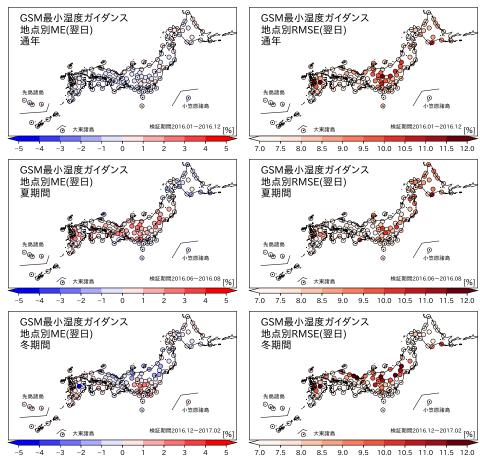


図 4.8.2 翌日を予報対象とする 12 UTC 初期値の GSM 最小湿度ガイダンスの期間別 ME ( 左列 )・RMSE ( 右列 ) のスコアマップ。上から通年、夏期間、冬期間の結果。

## 考えられる。

続いて、12UTC 初期値の GSM 及び MSM ガイダ ンスにおける翌日の最小湿度の予測について地点毎の RMSE と ME を図 4.8.2 と図 4.8.3 に示す。通年で見る と、ME について GSM ガイダンスは正バイアスと負 バイアスの地点があるが、概ね1%以内となっている。 MSM ガイダンスは正バイアスの地点が多くなってい る。RMSE については GSM, MSM ガイダンスともに 東日本内陸、北海道沿岸および九州で値が大きく、精 度が低い傾向がある。夏で見ると、ME について GSM と MSM ガイダンスの傾向は同じで内陸で正バイアス の傾向があり、最小湿度を高めに(乾燥の程度を弱く) 予測する傾向がある。北海道沿岸では逆に最小湿度を 下げすぎる傾向がある。RMSE については GSM ガイ ダンスより MSM ガイダンスの方が小さく、精度が高 い。冬で見ると ME について GSM, MSM ガイダンス ともに関東で正バイアスがあり、冬の乾燥の程度を弱 く予測する傾向がある。RMSE については夏と同様に GSM より MSM ガイダンスの方が値が小さく、精度が 高い。

## 時系列湿度ガイダンス

時系列湿度ガイダンスについて、GSM 及び MSM ガイダンスの 00UTC 初期値における全作成対象地点で平均した予報時間毎の RMSE と ME を図 4.8.4 に示す。GSM (MSM) ガイダンスの RMSE はどの期間もモデルより小さく、モデルを改善していることが分かる。また、どの期間も共通して、ガイダンスの ME は 0%付近となっており、モデルが持つ時刻ごとのバイアスの変動をほぼ除去できている。これは第 4.8.2 項 (1) で述べたように、初期時刻と予報時間で層別化しているためである。期間で比較すると夏期間の方が冬期間よりも RMSE が小さく精度が高い。RMSE についてGSM ガイダンスと MSM ガイダンスを比較すると、夏期間では予報期間の前半 (FT=15まで)で MSM ガイダンスの方が RMSE が小さく精度がやや高く、通年・冬期間では精度は同程度である。

次に、地点毎の予測特性を述べるため、00UTC 初期値における FT=24 の GSM 及び MSM ガイダンスについて作成対象地点毎の ME と RMSE を図 4.8.5 と図 4.8.6 に示す (共に朝 9 時のガイダンスの特性を示している)。 通年の ME をみると、どの地点も概ね ME が 0 に近くなっている。 夏期間の ME と RMSE をみると、

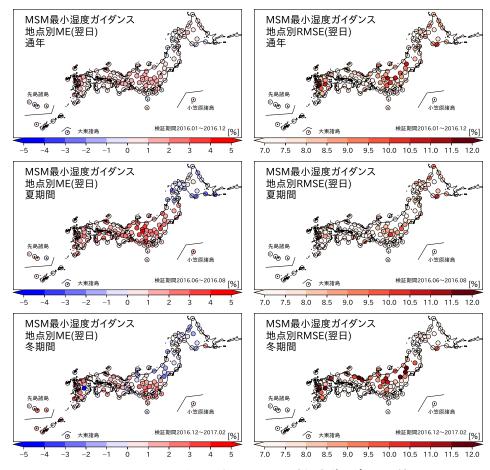


図 4.8.3 図 4.8.2 と同じ。ただし、MSM 最小湿度ガイダンスの結果。

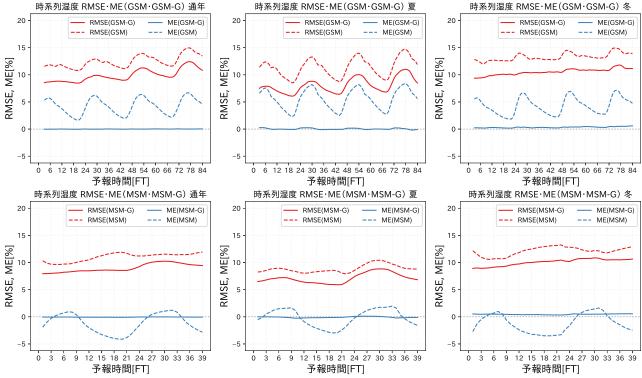


図 4.8.4 00UTC 初期値におけるモデルとガイダンスの時系列湿度の予報時間毎の期間別 ME ( 青線 )・RMSE ( 赤線 )。上段が GSM、下段が MSM で、GSM-G (MSM-G) は GSM (MSM) ガイダンスを示す。左から通年、夏期間、冬期間の結果。

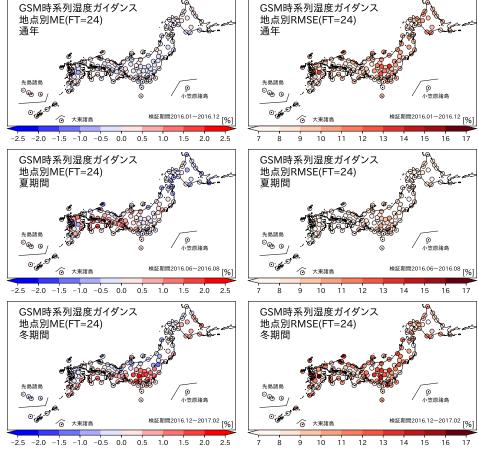


図 4.8.5 00UTC 初期値 FT=24 における GSM 時系列湿度ガイダンスの期間別 ME ( 左列 )・RMSE ( 右列 ) のスコアマップ。 上から通年、夏期間、冬期間の結果。

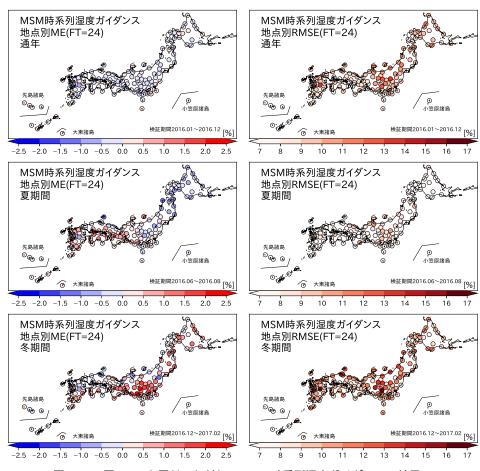


図 4.8.6 図 4.8.5 と同じ。ただし、MSM 時系列湿度ガイダンスの結果。

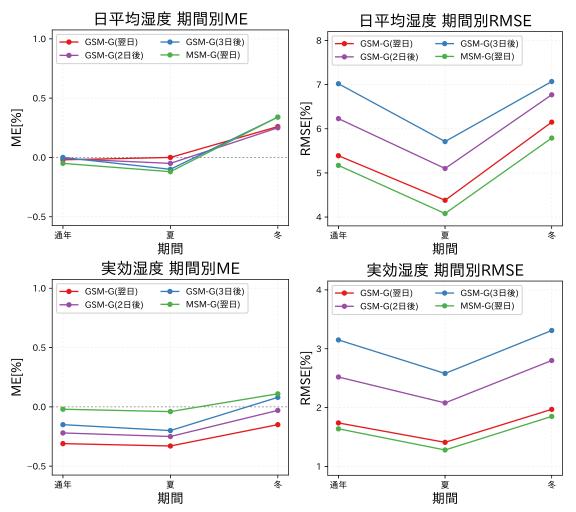


図 4.8.7 12UTC 初期値における日平均(上段)・実効湿度(下段)の期間別 ME(左)・RMSE(右)。GSM-G (MSM-G) は GSM (MSM) ガイダンスを示す。

ME は GSM・MSM ガイダンス共に西日本でバイアスが正となっている地点があること、RMSE は 10%以下と精度は通年に比べて高いことが分かる。GSM ガイダンスと MSM ガイダンスの比較からは、MSM ガイダンスの方が RMSE が小さく精度が高いことが分かる。冬の ME と RMSE を見ると、最小湿度と同様に関東平野で正バイアスがあり、湿度を高めに予測する傾向がある。また、通年・夏期間と比べて全国的に RMSE が大きく精度が低い傾向がある。これは、関東の正バイアスに加え、時雨等の断続的な降水が多いことや放射冷却の予測が難しいことが影響していると考えられる。

最後に、全地点で平均した予報対象日毎の日平均湿度と実効湿度の検証結果を図 4.8.7 に示す。日平均・実効湿度は、時系列湿度のガイダンス予測結果を用いているため、夏に比べて冬の精度が低いことについては共通している。時系列湿度より日平均湿度の方が RMSEが小さいが、これは日平均を計算する際に各時刻のランダム誤差が互いに打ち消しあうためと考えられる。時系列湿度と同様、日平均・実効湿度については MSM ガイダンスの方が GSM ガイダンスより概ね精度が高

い。ここで、GSM ガイダンスの実効湿度に平均湿度よりも負バイアスの傾向がある理由は、12UTC 初期値の FT=1, 2 の予測値がなく、日中に比べて湿度が高くなりやすい 22, 23JST の湿度が使われないことで日平均湿度が低くなってしまったためと考えられる。今後 GSM も FT=1, 2 の予測を行うよう改良する必要がある。

### 4.8.4 事例検証

湿度ガイダンスの予測特性が現れている事例について考察する。

### 乾燥注意報事例

東京で乾燥注意報が発表された事例を示す。2017年11月21日は、日本周辺は冬型の気圧配置であった(図略)。このため、東京では日中晴れが続き空気が乾燥することが予測されていた(図 4.8.8)。MSM の地上相対湿度は実況より低めに予測する一方で、MSM 時系列湿度ガイダンスは実況に近く、特に空気が最も乾燥した14時頃の予測はほぼ正確に予測していた。また、MSM 時系列湿度ガイダンスによる実効湿度では、20

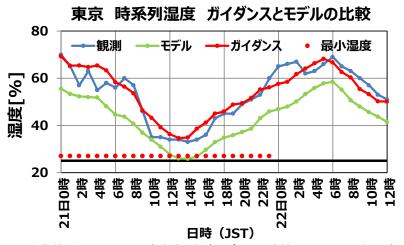


図 4.8.8 11 月 20 日 12UTC 初期値における MSM 時系列湿度ガイダンス(赤線)と MSM の毎正時における湿度予測(緑線) の結果(東京)。赤点線は同初期時刻における MSM 最小湿度ガイダンスの 21 日の予測結果。青線は毎正時における湿度の観測値で、黒線は東京の日最小湿度の乾燥注意報基準(25%)を示す。

日 12UTC 初期値で 58.8%で、同初期値における MSM 最小湿度ガイダンスの予測結果は 27%であった。21 日の実況では、実効湿度が 58.3%で最小湿度が 31%となり、最小湿度ガイダンスはやや低めの予測ではあったが、湿度が注意報発表基準(最小湿度 25%以下かつ実効湿度 50%以下)を下回るほどの乾燥ではないことを予測できていた。本事例は MSM の負バイアスを MSM ガイダンスが修正した事例となった。

### 4.8.5 利用上の留意点と今後の課題

最後に、第 4.8.3 項及び第 4.8.4 項を踏まえて、最小湿度ガイダンスと時系列湿度ガイダンスの利用上の留意点をまとめるとともに、湿度ガイダンスの今後の展望を述べる。

最小湿度ガイダンスについては、統計的に MSM ガイダンスの方が GSM ガイダンスよりも精度が高いので、MSM ガイダンスの利用を推奨する。ただし、GSM と MSM 予測の妥当性を判断しながら両者をあわせて利用していただきたい。また、GSM・MSM ガイダンス共に内陸で精度が低い傾向があること、夏の東日本以南と冬の関東で最小湿度を高く予測し、乾燥が弱めとなる傾向があるので留意願いたい。

時系列湿度ガイダンスは、夏期間で概ね GSM ガイダンスより MSM ガイダンスの方が精度が高い。GSM・MSM ガイダンス共に、夏に比べて冬に精度が低い傾向があり、これは放射冷却、時雨等の断続的な降水および関東の乾燥の予測が難しいことに起因していると思われる。日平均湿度・実効湿度の予測についても、一年を通して GSM ガイダンスより MSM ガイダンスの方が精度が高い。

続いて、湿度ガイダンスの今後の展望を述べる。2018年現在、時系列湿度ガイダンスと最小湿度ガイダンスは統計手法と説明変数が異なるため、予測特性も異なっている。両ガイダンスの整合性や利便性を考慮すれば、

統一的な方法でガイダンスを作成する必要があると考 えており検討中である。特に、最小湿度ガイダンスは 1996年から説明変数や基本的な手法は変わっておらず、 見直しが必要である。また、降水種別ガイダンス(第 4.3節)にはモデルの地上相対湿度を用いているために、 モデルのバイアスが雨雪判別に影響している。そこで、 地点形式の気温ガイダンスを格子形式気温ガイダンス に変換している(第4.4節)ように、湿度ガイダンス についてもモデルのバイアスを取り除いた格子形式湿 度ガイダンスの開発を検討している。降雪量ガイダン スでは防災気象情報の発表に用いるためにメソアンサ ンブル予報システムを入力としたガイダンス (MEPS ガイダンス)を開発する予定だが、降雪量ガイダンス は降水種別ガイダンスを入力とする(第4.3節)ため、 格子形式湿度ガイダンスについても MEPS ガイダンス を開発する必要がある。

### 参考文献

國次雅司, 2009: 概説. 平成 21 年度数値予報研修テキスト, 気象庁予報部, p. 20.

佐藤豊, 2000: RSM 最小湿度ガイダンスの改善と GSM 最小湿度ガイダンスの開発. 平成 12 年度量的予報研 修テキスト, 気象庁予報部, 71-78.

高桑健一, 2015: 時系列湿度ガイダンスの開発. 平成 27年度数値予報研修テキスト, 気象庁予報部, 73-78.

柳野健, 1995: ニューラルネットによるガイダンス. 平成7年度量的予報研修テキスト, 気象庁予報部, 54-69.

### 4.9 視程ガイダンス 1

### 4.9.1 はじめに

空港の視程は航空機の離着陸に大きく影響を与える。 パイロットが目視により自分の判断で飛行できる気象 の状態を有視界飛行状態 (VMC) と言う。離着陸の際の VMC の基準の一つとして視程 5000 m 以上が定められ ており、視程  $5000~\mathrm{m}$  未満では  $\mathrm{VMC}$  での離着陸はでき ない。また、計器着陸装置 (ILS) を利用した着陸では、 ILSの運用精度に応じてカテゴリー毎に基準が決まって おり、例えばカテゴリー a では滑走路視距離 (RVR)が 200 m 以上なければ着陸することはできない。この ように視程の基準が細かく設定されているため、運航 用飛行場予報 (TAF) において、視程 5000 m 未満では 100 m 毎に細かく予測しており、視程の予報変化基準 として、200, 400, 600, 800, 1600, 3200, 5000 m が国 際民間航空条約第3附属書 (ICAO Annex3) で定めら れている。また、視程は代替飛行場の選定基準にも用 いられている。

気象庁では空港の視程の予測を支援するため、1993 年から視程ガイダンスを運用しており、当時は線形重回 帰を用いて前3時間に視程が10km未満となる確率と 前3時間に10km未満となる場合の最小視程を予測対 象としていた。1996年には利用するモデルが領域モデ ル (RSM) に変わったが、仕様に変化はなく、2001 年に は予測対象を前3時間の最小視程と最小視程が5000 m 未満になる確率に、統計手法をカルマンフィルタに変 更し、頻度バイアス補正を導入した(大林 2001)。更 に 2002 年にはメソモデル (MSM) の運用開始に合わ せて、線形重回帰を用いた前1時間の最小視程を新た に予測対象に追加した(岩倉ほか 2002)。この前1時間 の最小視程は、一括学習型で MSM の非静力学化に対 応できないことや悪視程の捕捉率が低いという課題か ら、2004 年に予測手法を RSM 視程ガイダンスと同手 法のカルマンフィルタに変更した (高田 2004)。 MSM の予報時間が 33 時間に延長された 2007 年には、MSM ガイダンスの予報時間も延長し、それに伴い RSM ガ イダンスを廃止して MSM ガイダンスに統合した (高 田 2007)。その際に、予測対象を前1時間の最小視程と 平均視程、前3時間の最小視程が5000 m 及び1600 m 未満となる確率の4つとし、天気層別化(無降水、雨、 雪)と天気が無降水の場合に限り予報対象時刻別の層 別化を導入した。2010年には説明変数に雲水量を追加 するなどの改良を行った (松本・高田 2010)。

視程ガイダンスは空港での視程を予測するガイダンスであるが、空港以外の視程を予測するために、診断手法を用いて視程を面的に予測する MSM 視程分布予想の運用を 2011 年から開始した (井藤 2011)。 視程分布予想は濃霧注意報や海上濃霧警報などを発表する

際の参考資料として利用されている。2012 年には海上予警報業務の高度化へ資するために全般海上予報区 ( $0^{\circ}$ N $\sim$ 60 $^{\circ}$ N,  $100^{\circ}$ E $\sim$ 180 $^{\circ}$ E) の範囲内の海域の全てを予測対象領域とする GSM 視程分布予想も運用を開始した (井藤 2013)。 LFM 航空悪天 GPV でも視程分布予想は作成され、下層悪天予想図などに利用されている (金井ほか 2015)。

本節では視程を予測するガイダンスとして、空港の最小視程と平均視程を予測する視程ガイダンス、最小視程が 5000 m 未満となる確率 (PoV50) 及び 1600 m 未満となる確率 (PoV16) を予測する視程確率ガイダンス、面的な視程を予測する視程分布予想の 3 つについて述べる。第 4.9.2 項で各ガイダンスの仕様を、第 4.9.3 項と第 4.9.4 項で統計検証と事例検証を通じてガイダンスの予測特性を示し、最後に第 4.9.5 項で利用上の留意点と今後の展望について述べる。なお、飛行場予報などで扱う視程は全方向の平均的な視程である卓越視程であるため、本節では特に断らない限り卓越視程を視程と記述する。

### 4.9.2 仕様

視程ガイダンスと視程確率ガイダンスは空港における地点形式の予測を、視程分布予想は格子形式の予測を提供する。予測式の作成や更新には、目視観測である METAR  $^2$ , SPECI $^3$ , SCAN $^4$  を利用し、自動観測である METAR AUTO $^5$  による視程の観測値は利用しない $^6$ 。これは METAR AUTO の視程が RVR 観測装置または視程計の設置された場所における局所的な視程であり、目視観測や TAF で対象としている卓越視程と異なるためである。以下で各ガイダンスの仕様と作成手法について述べる。

### (1) 視程ガイダンス

視程ガイダンスは空港における前1時間の最小視程及び平均視程を予測対象とする。表4.9.1 に視程ガイダンスの仕様を示す。視程障害の発生は天気現象、特に降水現象に大きく依存するため、天気(無降水、雨、雪)毎に異なる予測式を作成しており、予測時は航空天気ガイダンス(付録4.9.A参照)で予測された天気によって予測式を使い分けている。

### 作成方法

視程ガイダンスの作成方法を説明する。

METAR などから前 1 時間における最小視程と最大視程を作成する。

<sup>2</sup> 航空気象定時観測気象報

<sup>3</sup> 航空気象特別観測気象報

<sup>4</sup> 航空気象観測所気象報

<sup>5</sup> 航空気象自動観測気象報

 $<sup>^6</sup>$  ただし、観測が完全自動化となった空港や観測時間においては他に観測がないため、自動  ${
m METAR/SPECI}$  報を使っている。

<sup>1</sup> 後藤 尚親

+ 401	30 TO 12 / 12	1 30 TO TO 25 25 2 25 2 5	<b>↑/14</b> ΕΕ
<del>75</del> 4 9 1	栂籽カイタン人	と視程確率ガイダンス	(/)什恁一智

作成対象	空港
作成方法	カルマンフィルタの予測に頻度バイアス補正を行う。
作成対象とするモデル	MSM
予報対象時間単位	1時間(視程) 3時間(視程確率)
予報期間と間隔	視程:FT=2 から FT=39 まで 1 時間間隔 視程確率:FT=6 から FT=39 まで 3 時間間隔
逐次学習の有無	あり
説明変数	地上相対湿度、雲水量、地上気温、地上風速、降水量
予測対象	視程:前 1 時間の最小視程と平均視程 視程確率:前 3 時間最小視程が 5000 m, 1600 m 未満 (1) か否 (0) か 観測は METAR, SPECI, SCAN を利用し、METAR AUTO は利用しない。
層別化処理の対象	作成対象地点(空港) 天気(無降水、雨、雪) 予報対象時刻(3 時間毎1日分、無降水のみ)
備考	頻度バイアス補正の閾値には以下を使用する。 視程:0.8, 1.6, 3.2, 5.0, 10.0 km 視程確率:0.0, 0.2, 0.35, 0.5, 0.65, 1.00

表 4.9.2 視程ガイダンスと視程確率ガイダンスの説明変数一覧。無降水、雨、雪の欄に各天気の予測式で利用する説明変数をで示し、式に示した変換を施す。

413.0 ( 2012.3	CNOC MENOLEGINE IN					
変数名	無降水	雨	雪	式	備考	
地上相対湿度				$\sqrt{1-RH}$	RH は相対湿度	
雲水量				$\sqrt{\sum_{i=2}^{9} \mathrm{QC}_i}$	$\mathrm{QC}$ は雲水量 $[\mathrm{g}\ \mathrm{kg}^{-1}]$ 、 $\mathrm{i}$ はモデル面の層番号	
降水量				$\sqrt{\text{RAIN}}$	RAIN は前 1 時間降水量 [mm]	
気温 × 風速				$T \times FF$	$ m T$ は気温 [ $^{\circ}$ C]、 $^{\circ}$ FF は風速 $^{\circ}$ EF は上限を $^{\circ}$ C $^{\circ}$ FF は上限を $^{\circ}$ C $^{\circ}$	

観測と予測の天気が一致している場合に、該当天 気の予測式を更新する。

MSMの地上面、モデル面から説明変数を作成する。 各天気の予測式から視程を算出する。

で算出した視程を頻度バイアス補正で補正し、各天気での視程とする。

航空天気ガイダンスで予測された天気に該当する 視程を各天気での視程から選択し、予測値とする。

視程ガイダンスでは、前 1 時間における最小視程と最大視程のそれぞれを目的変数とし、これらの予測値を平均することで平均視程とする( )。これは予測対象である平均視程を METAR などから求めるのは時間分解能の制約で難しい  $^7$  ためである。

航空天気ガイダンスで予測した天気と実況の天気が 一致している場合にカルマンフィルタによる係数更新、 頻度バイアス補正の閾値更新を行う()。

表 4.9.2 に示す説明変数を作成する( )。もやと霧の予測精度向上を目的として相対湿度と雲水量を、降水による視程悪化への効果として降水量を、吹雪による視程悪化への効果として気温 × 風速を用いる。気温 × 風速は気温が負の場合のみ値を持ち、前 1 時間における最小値を利用する。

で更新した予測式と で作成した説明変数を用いて、各天気における最小視程と最大視程を算出する()。算出した視程に頻度バイアス補正を行い、バイアス補正後の視程を各天気での視程とする()。各天気での視程は100 m単位となるように四捨五入を行い、最大値を9999 m、最小値を100 mとする。航空天気ガイダンスで予測された天気の視程を最終的な予測値とする()。ただし、みぞれの場合は雨と雪の予測式から求められた視程の平均値とする。また、最小視程と最大視程の平均値を予測対象である平均視程とする。層別化は、地点毎、天気毎、予報対象時刻毎(無降水のみ)で行っている。空港では夜間に観測が行われていない空港もあるので、その時間帯は最も近い時刻の予測式で代用する。また観測がほとんど入電しない空港では、近隣空港の予測式で代用する8。

### (2) 視程確率ガイダンス

視程確率ガイダンスは空港における前 3 時間の最小 視程が  $5000~\mathrm{m}$  及び  $1600~\mathrm{m}$  未満となる確率を予測対象 とする。表 4.9.1 に示す仕様の通り、視程確率ガイダンスの基本的なアルゴリズムは層別化を含め視程ガイダンスと共通だが、以下の点で視程ガイダンスと異なる。

<sup>7</sup> 定時の目視観測は最大でも30分毎であるため。

<sup>8 2018</sup> 年現在、小値賀空港と上五島空港の予測式を福江空港で、佐渡空港の予測式を新潟空港で代用している。

目的変数は前 3 時間最小視程が 5000 m もしくは 1600 m 未満 (1) か否 (0) か。

予測と観測の天気が一致しない場合も予測式を更 新する。

頻度バイアス補正の閾値が異なる。

天気や擾乱などの予測と実況の時間的なずれを軽減するために、前後の確率値と平均を行い、該当時刻の予測値とする。

視程確率ガイダンスでは天気が外れる確率も考慮した 尤もらしい確率となるように、天気が一致していない 場合においても予測式を更新する。また予報時間毎に 天気が変わる、もしくは予測対象である3時間内で天 気が変わると、天気層別の影響で予測される確率値が 大きく変動する。これを緩和するために前後の予報時間と平均を行う。この平均処理によって、予測値の変動を抑制するとともに、擾乱予測の位置ずれ等も緩和 され、予測精度が向上することを確認している。

### (3) 視程分布予想

視程分布予想は地上及び海上の視程を格子形式で予測するガイダンスである。表 4.9.3 に視程分布予想の仕様を示す。GSM 視程分布予想は予測対象である前 3 時間の最小視程を直接予測するが、MSM 視程分布予想では、1 時間毎に視程を予測し、前 3 時間内の 3 つの予測値の最小を前 3 時間の最小視程としている。気象要素(浮遊塵、霧、雨、雪)の視程への影響を考慮するため、後述する光消散係数を気象要素ごとに作成し、視程の予測値へ反映している。視程分布予想は GSM, MSM, LFM で作成されているが、LFM 視程分布予想は MSM 視程分布予想と作成方法が共通するところが多いため、ここでは GSM 視程分布予想と MSM 視程分布予想についての解説し、LFM 視程分布予想は MSM 視程分布予想との差について解説する。

### MSM 視程分布予想

MSM 視程分布予想では、MSM のモデル面予測値の雲水量、相対湿度、降水量、降雪量、風速から光消散係数  $\sigma$  を求め、それを気象光学距離 (MOR: Meteorological Optical Range) の定義に従って視程に変換している。世界気象機関により、MOR は色温度 2700K の白熱灯の並行ビームが大気や大気中に浮遊する粒子によって散乱吸収され、その照度 E が元の照度の 5%となるまでの距離と定義されている。

Lambert の法則によると、照度 E の距離 r に対する 変化率は  $\sigma$  を用いて以下のように書ける。

$$\frac{d\mathbf{E}}{dr} = -\sigma \mathbf{E} \tag{4.9.1}$$

ここで  $\sigma$  を距離が 0 から r の間でほぼ変わらないと考えて一定とすると、

$$\frac{\mathbf{E}(r)}{\mathbf{E}(0)} = \exp(-\sigma r) \tag{4.9.2}$$

MOR の定義より、視程 V[km] は照度 E が 5%になるまでの距離なので、

$$V = \frac{-\ln(0.05)}{\sigma} \approx \frac{3}{\sigma} \tag{4.9.3}$$

MSM 視程分布予想では、この視程と光消散係数の関係式を利用する。

 $\sigma$  は大気中に浮遊している浮遊塵  $\sigma_p$ 、雲粒  $\sigma_c$ 、雨粒  $\sigma_r$ 、雪  $\sigma_s$  による光消散係数の和で与えられる。

$$\sigma = \sigma_p + \sigma_c + \sigma_r + \sigma_s \tag{4.9.4}$$

浮遊塵による光消散係数  $\sigma_p$  は、岩倉・岡田 (1999) の 視程と相対湿度の関係式を参考に、METAR などから 求めた視程と  ${
m MSM}$  による湿度の予測との比較から作成した以下の近似式を用いる。

$$\sigma_p = 0.12 \times (1 - \text{RH})^{-0.5}$$
 (4.9.5)

湿度 RH は MSM 地上面の予測値を用い、前 1 時間の RH と平均して説明変数とする。ただし (4.9.5) 式は高湿度では成り立たなくなるため、RH は上限を 0.85 としている。また、前 1 時間の降水量が 0.1 mm 以上もしくは前 1 時間の雪の降水量が 0.03 mm 以上の場合、降水による浮遊塵の洗浄効果を考慮して  $\sigma_p=0$  とする。

雲粒による光消散係数  $\sigma_c$  は、Gultepe et al. (2006) などを参考に  $\mathrm{MSM}$  モデル面の第 2 層から第 9 層の雲水量の和  $\mathrm{QC}[\mathrm{g~kg^{-1}}]$  を利用した以下の近似式を用いる。

$$\sigma_c = 12.0 \times QC^{0.9}$$
 (4.9.6)

係数である 12.0 は  ${
m MSM}$  の  ${
m QC}$  と  ${
m METAR}$  等の視程を比較し、悪視程の予測頻度が適切になるように決定している。

雨粒による光消散係数  $\sigma_r$  は、鈴木 (1963) など古くから多くの調査が行われており、前 1 時間降水量 R[mm]を使った以下の近似式を用いる。

$$\sigma_r = 0.45 \times R^{0.55} \tag{4.9.7}$$

係数である 0.45 は MSM の降水量と METAR 等の視程との関係から決定した。

雪による光消散係数  $\sigma_s$  は、藤吉ほか (1983) などを参考に、雪の前 1 時間降水量  $\mathrm{SNOW}[\mathrm{mm}]$  とモデル面第 2 層の風速  $\mathrm{FF}[\mathrm{m/s}]$  を使った以下の近似式を用いる。

$$\sigma_s = 7.5 \times \text{SNOW}^{0.7} + 0.07 \times \text{FF}$$
 (4.9.8)

係数については前述の他の要素と同様に観測値とモデルの予測値から決定した。風速の項は吹雪による視程低下の効果を取り込むためであり、SNOWが 0.1~mm以上予測され、かつ降水(雨)が予測されていない場合に付加される。

ここまで示したとおり、光消散係数を算出する近似式の決定には MSM の予測値が利用されている。そのため、視程の予測頻度、精度は MSM の降水予測およ

表 4.9.3 視程分布予想の仕様一覧

作成対象	GSM: 20 km 格子 ( 等緯度経度座標 ) MSM: 5 km 格子 ( 等緯度経度座標 )
	LFM:10 km 格子 <sup>9</sup> (ポーラーステレオ座標)
作成方法	消散係数による診断手法 (第 4.9.2 項 (3) 参照 )
作成対象とするモデル	GSM, MSM, LFM
予報対象時間単位	GSM, MSM:3 時間、LFM:1 時間
予報期間と間隔	GSM: FT=3 から FT=84 まで 3 時間間隔 MSM: FT=3 から FT=39 まで 3 時間間隔 LFM: FT=1 から FT=9 まで 1 時間間隔
逐次学習の有無	なし
説明変数	湿度、雲水量、降水量、雪の降水量、風速
予測対象	GSM, MSM: 各格子の前3時間の最小視程、LFM: 各格子の正時の視程
層別化処理の対象	なし
備考	$\mathrm{GSM}:$ 降水の雨雪判別は天気ガイダンス(降水種別)に準じている。 $\mathrm{MSM}:$ モデル(ランベルト)とガイダンス(等緯度経度)の座標系の違いにより、南北端で一部欠損値が入る。 共通:視程の他に雲の消散係数 $\sigma_c$ 、雨の消散係数 $\sigma_r$ 、雪の消散係数 $\sigma_s$ から計算した視程も作成している。

び地上付近の雲水量の精度に大きく依存する。降水量や雲水量が予測されていない場合は浮遊塵のみが光消散係数に反映されるが、RHの上限が85%であるため、(4.9.5) 式より浮遊塵の効果のみでは9000 mを下回ることはない。なお、浮遊塵による視程悪化は、その濃度が考慮されておらず、相対湿度に対応した平均的な濃度における視程が示される。

最後に MSM 視程分布予想に導入している雲水量を増加させる処理について説明する。降水量が 0.2 mm 以下、雲水量が予測されているなどの簡易的な条件で霧雨と推定される際は、推定された領域に対して雲水量を増やす処理を行う。これは、2017 年 2 月に実施された MSM の改良 (原 2017) によって、モデル内でそれまで予測されていた雲水に代わって微小な粒径の降水が予測されるようになり、主に海上で雲水量の予測が減少し悪視程の予測頻度が下がることが判明したため導入した処理である。この処理によって、説明変数の雲水量のオーダーをモデルの変更前のオーダーに近づけ、悪視程の予測精度の悪化を防いでいる。この処理についての詳細は後藤 (2017) を参照していただきたい。

### GSM 視程分布予想

GSM 視程分布予想の基本的なアルゴリズムは MSM 視程分布予想と同じだが、予測対象領域が異なる、モデル面の代わりに指定気圧面 (P面)を利用する、雨と雪による視程低下を考慮しないことなどが異なる。雨と雪による視程悪化を考慮しないのは、GSM 視程分布予想の開発目的である海上予警報の対象が、霧ともやによる悪視程を対象とするためである。そのため、GSM 視程分布予想での光消散係数は以下の式で表さ

れる。

$$\sigma = \sigma_p + \sigma_c \tag{4.9.9}$$

なお、海上の霧予想が主なターゲットではあるが、予測式の作成にあたっては  $SHIP^{10}$  は用いず、MSM 視程分布予想と同様に METAR を用いた。GSM 視程分布予想では海上の視程を予測対象とするため、本来は SHIP を用いて予測式で作成するべきであるが、SHIP を用いて予測式では悪視程の予測頻度が低く、予測精度が悪かったため METAR を用いた。これは SHIP が主に 6 時間毎の通報であるため METAR に比べて学習に利用できるデータが少なくなることや、海上では視程目標物がないため視程観測の刻みが粗く  $^{11}$ 、観測の精度も METAR に比べて悪いためである。

 $\operatorname{GSM}$  視程分布予想で用いる各気象要素における光 消散係数の近似式を示す。まず浮遊塵による光消散係 数  $\sigma_n$  は、以下の近似式を用いる。

$$\sigma_p = 0.162 \times (1 - \text{RH})^{-0.5}$$
 (4.9.10)

RH の上限は  ${
m MSM}$  視程分布予想と同様に 85%であり、降水時に  $\sigma_p$  を 0 とする浮遊塵の洗浄効果も導入されている。

雲水による光消散係数  $\sigma_c$  は、以下の近似式を用いる。

オホーツク海:
$$\sigma_c = 22.7 \times \mathrm{CWC}_3^{0.96} \ (4.9.11)$$
  
その他の領域: $\sigma_c = 29.3 \times \mathrm{CWC}_2^{0.96} \ (4.9.12)$ 

 $\mathrm{CWC}_2$  は地表気圧より上層の  $\mathrm{P}$  面 2 層の雲水量の和、 $\mathrm{CWC}_3$  は地表気圧より上層の  $\mathrm{P}$  面 3 層  $^{12}$  の雲水量の

 $<sup>^9</sup>$  LFM 視程分布予想は、LFM 航空悪天 GPV の一要素として配信しているため、LFM 航空悪天 GPV の格子間隔である  $10~{
m km}$  となっている。

<sup>10</sup> 海上実況気象通報式

 $<sup>^{11}</sup>$  5000 m 以下の観測を例とすると、METAR は  $100~\mathrm{m}$  刻みであるのに対し、SHIP は  $50,\,200,\,500,\,1000,\,2000,\,4000~\mathrm{m}$  である。

<sup>&</sup>lt;sup>12</sup> 例えば地表気圧が 1013 hPa の場合、1000, 975, 950 hPa の 3 層となる。

和を表す。ただし、 $850~\mathrm{hPa}$ 以上の層は利用しない。また、雲量が0.5以上の場合は $\mathrm{CWC}$ を雲量で除算して格子平均の雲水量から格子内最大値への変換を行っている。 $\mathrm{GSM}$  視程分布予想では(4.9.11) 式に示すように、オホーツク海  $(45^\circ\mathrm{N}\sim60^\circ\mathrm{N}, 140^\circ\mathrm{E}\sim155^\circ\mathrm{E}$  の領域)とその他の領域で $\sigma_c$  の近似式を切り替えている。近似式を切り替える理由は、 $\mathrm{CWC}_2$  を用いて予測値を算出し、 $\mathrm{SHIP}$  を用いて検証を行ったところ、オホーツク海における霧の捕捉率が低くなることが確認されたためである。なお、オホーツク海とその他の領域の境界付近で予測値が急変するのを防ぐため、境界の $\pm1^\circ$ では消散係数 $\sigma_c$  を周囲  $5\times5$  格子で平均化した値を利用する。

なお、光消散係数  $\sigma$  には反映されないが、雨粒による光消散係数  $\sigma_r$  と雪による光消散係数  $\sigma_s$  も以下の近似式を用いて算出し、要素別の視程に変換している。これらは MSM 視程分布予想と同様に陸上における視程の予測を検討する際の参考資料として利用出来る。

$$\sigma_r = 0.403 \times R^{0.5}$$
 (4.9.13)

$$\sigma_s = 2.14 \times \text{SNOW}^{0.7} + 0.167 \times \text{FF} \quad (4.9.14)$$

R は前 3 時間内の 1 時間降水量の最大値を用いる。雪の降水量 SNOW は、MSM と異なり GSM では直接予測されないため、降水種別ガイダンス(第 4.3 節)を簡略化した手法  $^{13}$  を用いて雨雪判別を行い、雪と判定された格子の降水量 R を雪の降水量 SNOW として置き換えて計算している。

# LFM 視程分布予想

LFM 視程分布予想は、基本的には MSM 視程分布予想と同様のアルゴリズムで算出している。 MSM 視程分布予想との主な違いは以下の 3 点である。

前3時間の最小視程ではなく、1時間毎の正時の 視程を対象としている。

雲水量  $\mathrm{QC}$  はモデル面第 1 層から第 4 層の和としている。

霧雨と推定される領域でも雲水量を増加させる処理を行わない $^{14}$ 。

以下に LFM 視程分布予想で用いられる光消散係数の 近似式を示す。変数名の記述は MSM 視程分布予想の 説明と同様である。

$$\sigma = \sigma_p + \sigma_c + \sigma_r + \sigma_s \tag{4.9.15}$$

$$\sigma_p = 0.23 \times (1 - \text{RH})^{-0.5}$$
 (4.9.16)

$$\sigma_c = 6.0 \times QC^{0.9}$$
 (4.9.17)

$$\sigma_r = 0.45 \times R^{0.55}$$
 (4.9.18)

$$\sigma_s = 3.6 \times \text{SNOW}^{\frac{2}{3}} + 0.07 \times \text{FF}$$
 (4.9.19)

### 4.9.3 統計検証

視程ガイダンス、視程確率ガイダンス、GSM, MSM 視程分布予想の統計検証の結果を示す。検証には2016 年1月から12月の03,09,15,21UTCを初期時刻とす る FT=6 から FT=27 の予測値を用いた。GSM 視程 分布予想については同期間の 00, 06, 12, 18UTC を初 期時刻とする同 FT の予測値を用いた。視程において は値が小さい、つまり悪視程の精度が重要になるため、 閾値未満の視程が予測または観測されたときに「予測 あり」または「観測あり」としたカテゴリー検証を示 す。空港を対象として検証を行う場合は、地域航空気 象解説報でカテゴリー予想 15 が発表される空港を検証 対象としている。また全事例と記述した場合は全ての 事例を、各天気(無降水、雨、雪)の検証とした場合 は観測と予測の天気が一致した事例の検証である。便 宜上、以下では視程ガイダンスを VIS、視程確率ガイ ダンスを PoV、視程分布予想を PVIS と記載すること がある。なお、数値予報モデルには2018年3月現在運 用している GSM, MSM と同仕様の予測値を用いた。

### (1) 視程ガイダンス

図 4.9.1 は全事例について視程ガイダンスを閾値別に検証したエクイタブルスレットスコア (ETS) とバイアススコア (BI) である。 閾値  $5000~\mathrm{m}$  では BI は  $1~\mathrm{c}$  近く ETS は 0.3 程度だが、悪視程になるほど BI 及び ETS は低くなる傾向にある。これは悪視程ほど事例数が少ないことが影響していると考えられる。

図 4.9.2 は観測と予測の天気が一致した事例を対象 とした天気別検証である。 閾値 5000 m 及び 3200 m で は無降水が最も ETS が高く、1600 m 及び 800 m では 雨、雪の ETS が無降水を上回る。天気別に見ると無降 水は閾値 5000 m で BI が 1.2、ETS が 0.2 程度だが悪 視程になるほど BI は大幅に低くなり、ETS も低下す る。悪視程における ETS, BI が低いのは浮遊塵の効果 を相対湿度のみで計算しているため予測精度が十分で はないことや、雲水量の予測に関しても精度が十分で はない(特に放射霧の際に顕著)ことが影響している。 一方で、雨・雪の ETS, BI は悪視程でも低下していな い。これは雨・雪の悪視程の事例は無降水に比べて頻度 が高いことや、雲水量に比べて降水の予測精度が高い ことが影響している。なお、本検証は天気が一致した 場合が対象であり、雨や雪の予測が的中したほとんど の事例において視程が5000 m を下回るため、気候値予 報からの改善である ETS は 3200 m 以下より 5000 m の方が低くなっている。

図 4.9.3 に無降水事例における予報対象時刻別の検証を示す。閾値別検証と同様に、より悪視程である  $1600~\mathrm{m}$  の方が  $\mathrm{ETS}$ ,  $\mathrm{BI}$  ともに低い。閾値  $5000~\mathrm{m}$  では概ね  $\mathrm{BI}$ 

 $<sup>^{13}</sup>$  地上の相対湿度、気温と  $850~\mathrm{hPa}$  の気温から判定する。  $^{14}$  MSM と LFM はモデルの対流パラメタリゼーションの扱いに差があり、LFM では雲水量を増加させる必要がないため、

<sup>15</sup> 航空機の運航条件に対応した各種カテゴリー(VMC・IMC 及び周回進入条件及び雷の有無)を空港運用時間内の3時間 毎で予測して表形式で提供するもの(予報部予報課航空予報 室 2006)。

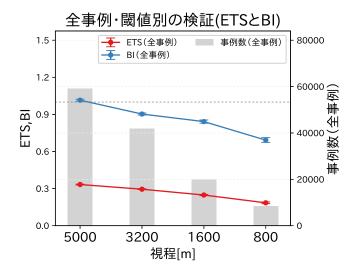


図 4.9.1 視程ガイダンスの全事例における閾値別の検証図。 左軸折れ線グラフで ETS と BI を、右軸棒グラフで事例数 を示す。ETS と BI のエラーバーは 95%信頼区間である。

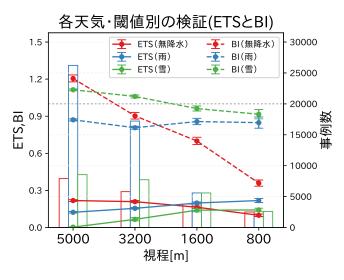
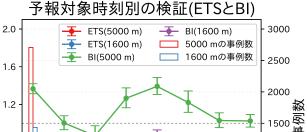


図 4.9.2 観測と予測の天気が一致した事例を対象とした視 程ガイダンスの天気別の閾値別検証図。左軸折れ線グラフ でETS(実線)とBI(破線)を、右軸棒グラフで事例数 を示す。ETS と BI のエラーバーは 95%信頼区間であり、 色分けは検証に用いた天気(赤:無降水、青:雨、緑:雪) である。

は1に近いが、1600 m では全ての時刻で BI が1を下 回る。また閾値 1600 m では 00UTC の事例数が多いも のの、ETS は他の時刻より低い。これは放射霧の予測 精度が不十分であることを示し、無降水事例における 悪視程の出現頻度の低さやモデルによる下層の雲水量 の予測が不十分であることが要因として考えられる。

図 4.9.4 に無降水事例における月別の検証を示す。こ れまでの検証と同様に、より悪視程である 1600 m の 方が ETS, BI ともに低い。事例数の多い暖候期に ETS が高く、事例数の少ない寒候期に ETS が低くなる傾向 がある。これは放射霧の予測精度が低いことが影響し



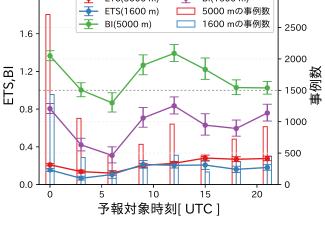


図 4.9.3 5000 m と 1600 m を閾値とした視程ガイダンス の無降水事例における予報対象時刻別の検証図。左軸折れ 線グラフで ETS と BI を、右軸棒グラフで事例数を示す。 ETS と BI のエラーバーは 95%信頼区間である。

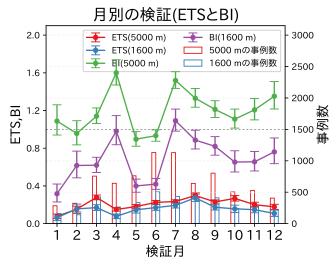


図 4.9.4 図 4.9.3 と同じ。ただし月別の検証結果を示す。

### ていると考えられる。

図 4.9.5 に全事例における 1600 m を閾値とした作成 対象地点別の検証スコアを示す。北日本や日本海側で BIが1に近くETSも高い地点が多いが、その他の地 域ではBIが1から大きく離れ、ETS も低い地点が多 い。北日本などで予測精度が高いのは雪の予測精度が 高いことが影響していると考えられる。南西諸島では、 無降水時には悪視程が発生する頻度が低く、降水事例 ではモデルで位置ずれの影響が大きい対流性の降水が 多いことが、予測精度が低い要因であると考えられる。 視程確率ガイダンスのスコア (ETS-PoV, BI-PoV) に ついては後述する。

# 視程ガイダンスと視程確率ガイダンスの地点別検証(閾値1600 m)

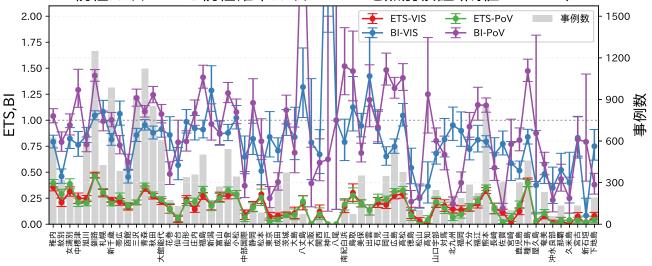


図 4.9.5  $1600~\mathrm{m}$  を閾値とした視程ガイダンスと視程確率ガイダンス (  $\mathrm{PoV16}$ 、実況ありとする閾値は 25% ) の地点別検証図。 左軸折れ線グラフで  $\mathrm{ETS}$  と  $\mathrm{BI}$  を、右軸棒グラフで事例数を示す。 $\mathrm{ETS}$  と  $\mathrm{BI}$  のエラーバーは 95%信頼区間である。

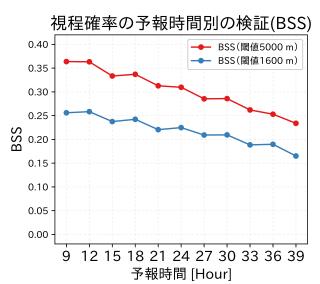


図 4.9.6 視程確率ガイダンスの予報時間別の検証図。PoV50 と PoV16 の BSS を示す。

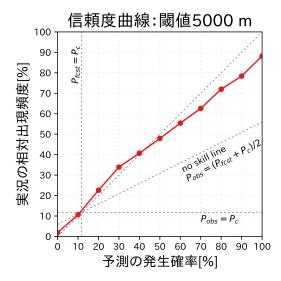


図 4.9.7 視程確率ガイダンスの信頼度曲線。検証対象は  ${
m PoV}50$ 。

### (2) 視程確率ガイダンス

図 4.9.6 に視程確率ガイダンスの予報時間別のブライアスキルスコア (BSS) を示す。PoV50 と PoV16 ともに予報時間内の BSS は正で確率予測としてスキルがあることが確認できるが、予報時間が進むにつれ BSS は徐々に下がる傾向があることがわかる。

図 4.9.7 と図 4.9.8 は視程確率ガイダンスの信頼度曲線である。図中点線で描画されている y=x の理想直線に概ね近いが、PoV50 は 50%以上の閾値で、PoV16 では 40%以上の閾値で理想直線からやや離れる。どちらも予測の発生確率が実況の相対出現頻度に比べて高いことから、視程確率ガイダンスは高い確率値において、実況に比べてやや高い予測値となる傾向があるこ

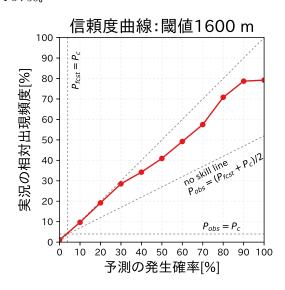


図 4.9.8 図 4.9.7 に同じ。ただし検証対象は PoV16。

### 視程と視程確率の比較(5000 m) 0.40 4.0 ETS-PoV BI-PoV **ETS-VIS** BI-VIS 0.35 0.30 3.0 0.25 2.5 $\overline{\mathbf{a}}$ 2.0 🗖 0.20 1.5 0.15 0.10 1.0 0.05 0.5 0.00 0.0 10 30 40 50 60 70 80 90 100 予測確率[%]

図 4.9.9 5000 m を閾値とした視程ガイダンスと視程確率ガイダンス (PoV50) の比較図。左軸で ETS を、右軸で BI を示す。エラーバーは 95%信頼区間である。

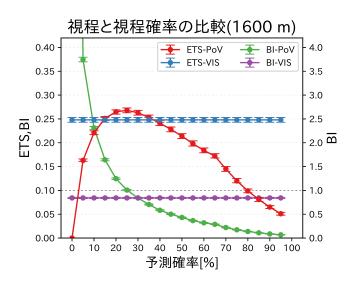


図 4.9.10 図 4.9.9 と同じ。ただし 1600 m を閾値としている。

### とを示している。

次に視程ガイダンスとの比較として、図 4.9.9 と図 4.9.10 に 5000 m 及び 1600 m を検証対象とした視程ガイダンスと視程確率ガイダンスの ETS と BI を示す。検証閾値 5000 m 及び 1600 m ともに視程確率ガイダンスの ETS (ETS-PoV) は 25%前後で最大となり、ETS が最大となる確率値付近では BI (BI-PoV) は 1 に近く、視程ガイダンスの ETS (ETS-VIS) を上回る精度となっている。予報作業において、視程が 5000 m もしくは 1600 m 未満となるか否かを判断する際に、視程確率ガイダンスの ETS が最大となる確率値を目安とすることで、予報の精度が最も高くなることが期待される。視程ガイダンスだけではなく、視程確率ガイダンスの予測も参考にしていただきたい。なお、図 4.9.5 に検証閾値を 25%とした PoV16 と VIS の地点別検証の比較

# MSM視程分布と視程の比較(全事例)

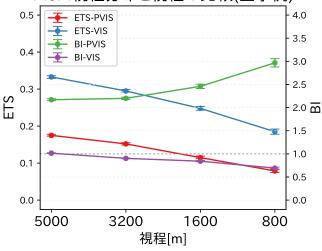


図 4.9.11 MSM 視程分布予想と視程ガイダンスの空港における閾値別検証。左軸で ETS を、右軸で BI を示す。エラーバーは 95%信頼区間である。

# GSMとMSMの視程分布比較(海域別)

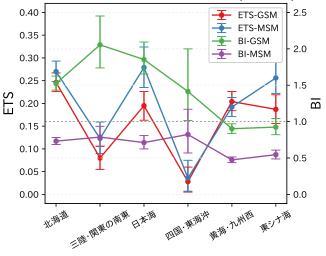


図 4.9.12 1000 m を閾値とした GSM 視程分布予想と MSM 視程分布予想の海域別検証。左軸で ETS を、右軸で BI を示す。エラーバーは 95%信頼区間である。

を示すが、差は小さいものの、多くの地点で PoV-ETS が VIS-ETS を上回っていることが分かる。

### (3) 視程分布予想

図 4.9.11 に MSM 視程分布予想と視程ガイダンスの空港における全事例の閾値別検証を示す。 MSM 視程分布予想はどの閾値でも BI は 1 より大きく、ETS は悪視程になるほど低くなる。視程ガイダンスと比較すると、MSM 視程分布予想は全ての閾値において視程ガイダンスの ETS を下回っている。更に BI で比較しても、視程ガイダンスは 1 に近く、カルマンフィルタによって係数を最適化しているほか、頻度バイアス補

# 2016年3月8日の東京国際空港の悪視程事例

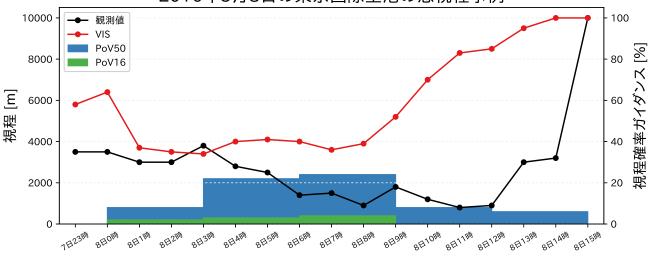


図 4.9.13 日本時間 2016 年 3 月 7 日夜遅くから 8 日にかけての東京国際空港の視程の観測値とガイダンス。ガイダンスは 2016 年 3 月 7 日の 12UTC 初期値で、左軸折れ線グラフで前 1 時間最小視程 (VIS) の観測値と視程ガイダンスを、右軸棒グラフで視程確率ガイダンス (PoV50 と PoV16) を示す。

正で予測頻度を補正している効果が確認出来る。このため、空港での視程の予測には視程ガイダンスを用いる方が通常は精度が高くなるが、視程ガイダンスは空港近傍のモデル予測値しか利用していないため、現象の予測の位置ずれが大きい場合には対応できない。これに対し、MSM 視程分布予想ではより広範囲における悪視程発生の可能性を確認出来るため、現象の予測に不確実性がある場合などは MSM 視程分布予想も参考にしていただきたい。

図 4.9.12 に GSM 視程分布予想と MSM 視程分布予想について、SHIP を用いた検証閾値 1000 m における無降水事例の海域別検証の比較を示す。図より、ほぼ全ての海域で MSM 視程分布予想の方が ETS が高いことがわかる。BI は MSM 視程分布予想は過少傾向、GSM 視程分布予想は過大の傾向が見られる。海域別に見ると、GSM, MSM 視程分布予想ともに黄海・九州西と東シナ海は予測頻度が他の領域に比べて低い。これは予測式を METAR で調整しているためと考えられる。予測式を調整する際は、観測点の緯度や経度、事例数の偏りなどを区別せずに利用するため、海霧の影響で霧の観測数が多い北日本の太平洋側の空港に最適化される傾向が大きい。そのため北の海域では精度が高く、南の海域で精度が低い結果になったと考えられる。

### 4.9.4 事例検証

ここでは霧の影響で50便以上が欠航となった東京国際空港での予測を紹介する。2016年3月8日は0時過ぎから関東地方の広範囲で濃霧となった。前日7日の日中には日本の東に中心をもつ高気圧の影響で、南からの暖湿気が流入し、17時ごろまで降水が継続していた。その後天気は回復したが高湿な気塊が関東平野に

残留し、夜間の放射冷却によって関東地方の各地で霧が観測された。東京国際空港においても、 内陸部で冷やされた高湿な気塊が弱い北風によって流入し、視程が悪化、午前7時には視程900mの霧を観測した。その後気温上昇に伴い一旦視程は回復したが、 午前9時頃から風向が東南東に変わり、今度は海上からの霧の流入により、午前11時前には再度視程が800mまで悪化した。

本事例について、図 4.9.13 に東京国際空港における 視程の観測値とガイダンスの時系列を、図 4.9.14 と図 4.9.15 に視程分布予想の予測と観測値、衛星画像を示す。 の放射霧について、視程ガイダンスは視程悪化を予測しているものの、最も視程が悪化する時刻で3500 m 以上であり、実況ほどの悪視程は予測できていなかった。PoV16 についても予測確率は数%と低く、高確率を予測出来なかった。これに対し、MSM 視程分布予想では東京国際空港周辺で100 m と実況よりも悪目だが、悪視程を予測出来ていた。GSM 視程分布予想では都心や空港付近では悪視程を予測できなかった。

の視程悪化は、図 4.9.15 の可視画像で赤線で囲んだ領域のように、東京湾に残留した海霧が東南東の風で侵入してきたことによって引き起こされたものである。視程ガイダンス、視程確率ガイダンスでは既に回復傾向で見逃しとなっている。 MSM 視程分布予想では東京湾に悪視程の領域が残ることは予測出来ているが、それでも 1800 m 程度であり、実況ほどの悪視程とはなっていない。 GSM 視程分布予想も悪視程を予測できていなかった。

今回の事例ので示したとおり、MSM 視程分布予想では予測がやや実況より悪目ながら悪視程が予測できているのに対し、視程ガイダンスでは悪視程を十分

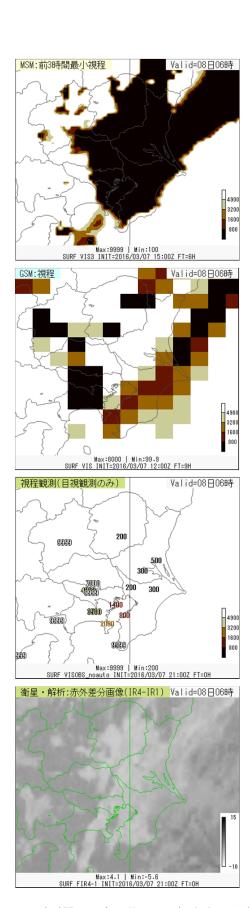


図 4.9.14 日本時間 2016 年 3 月 8 日 6 時の視程の予測値 [m] と観測値 [m] および衛星画像。上から MSM 視程分布予想 (2016 年 3 月 7 日 15UTC 初期値の FT=6)、GSM 視程 分布予想 (2016 年 3 月 7 日 12UTC 初期値の FT=9)、視程の観測値 [m]、赤外差分画像 [K]。視程の観測値の上限は 9999 m としている。

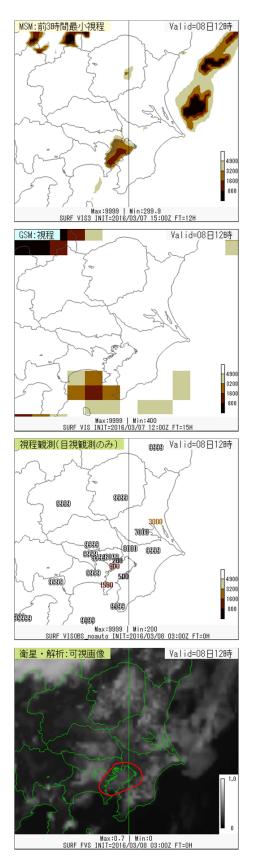


図 4.9.15 図 4.9.14 と同じ。ただし予報対象時刻は日本時間 2016 年 3 月 8 日 12 時で、赤外差分画像の代わりに可視画像を示す。可視画像図中の赤枠は同時刻に東京湾に発生していた海霧を示す。

に予測できていない。視程分布予想は統計処理を行わずモデルの雲水量を基に視程を予測しているが、視程ガイダンスはモデルの雲水量の外れも学習するため、図 4.9.11 で示したように、視程分布予想に比べ悪視程の予測は少なくなる。統計検証では視程ガイダンスの方が視程分布予想より精度は高いが、本事例のように視程ガイダンスの予測が不十分となる場合があるので、視程分布予想も参考にして予報作業を行っていただきたい。

### 4.9.5 利用上の留意点と今後の展望

本節では視程ガイダンスと視程確率ガイダンス、視程分布予想について開発や改良、仕様、検証結果を示した。空港で時系列の視程を予測する場合、視程ガイダンスの精度が最も高い。視程が5000 m もしくは1600 m 未満となるか否かを予測する場合には、視程確率ガイダンスも有効である。また事例検証で示したとおり、視程分布予想はメリハリの効いた面的予測を行うため、空振りは多いが、空港周辺の視程の予測も確認することができるため、予測の位置ずれなどが見込まれる際には有効となる。また、事例検証で示したとおり、統計処理を行っている視程ガイダンスでは予測が不十分になる場合でも、視程分布予想では悪視程を予測できる場合がある。

海上予警報に利用する場合は精度が高い MSM 視程分布予想を主に利用し、MSM 視程分布予想の予報領域外については GSM 視程分布予想を利用するのが良い。ただし、GSM 視程分布予想は実況に比べて広めに予測する傾向があることに注意が必要である。

最後に視程ガイダンスの今後の展望を述べる。視程ガイダンスは、予測式を天気で層別化しているため、この天気を決める航空天気ガイダンスの精度の影響を大きく受ける。航空天気ガイダンスのフローチャートの閾値について、最近のモデル改良の影響を反映した調整や、場合によっては手法の再検討が必要である。

視程確率ガイダンスは現在はカルマンフィルタで確率予測をしているが、精度向上を図るために、確率予測に適したロジスティック回帰に統計手法を変更することを検討する必要があるだろう。

視程分布予想は運用を開始して以降、モデル改良(米原 2017; 原 2017) に対応した予測式の調整が行っているが、説明変数の見直しや参照とする予測式を変更するなど、抜本的な調整は行っていない。例えば、視程の悪化にはエアロゾルの影響もあるが、現在の予測式では考慮されていない。エアロゾルの移流モデルなどの予測結果を用いて調査を行い、予測式にエアロゾルの影響を反映することを目指したい。また、MSM 視程分布予想で導入されている霧雨領域で雲水量を増やす処理は、主に海上での悪視程の予測頻度を増やすために導入したが、この処理の影響で陸上の悪視程は予測頻度が過大となることがわかっている(図略)。その

ため、今後より詳細な調査を行い、この処理を見直すことを検討している。

# 付録 4.9.A 航空天気ガイダンス

航空天気ガイダンスは、空港における卓越天気を予測するガイダンスである。1993年に運用を開始し、当時は前3時間の卓越天気が予測対象であり、天気ガイダンスから空港周辺の天気を抜き出すことで予測値を作成していた。その後2002年に MSM の運用開始に合わせ、お天気マップと同様のフローチャートを用いて、1時間単位の天気を予測するガイダンスに変更し、2007年に雨雪判別に航空気温ガイダンス $^{16}$ を利用するように変更した。なお2007年より、予報現場からの要望で海面更正気圧 [hPa] も要素の一つとして配信している。

航空天気ガイダンスのフローチャートを図 4.9.16 に示す。航空気象予報に合わせるため、以下の点がお天気マップと異なる。

予測するカテゴリが晴れ、曇り、雨、みぞれ、雪 である。

雨、みぞれ、雪については表 4.9.4 に示す閾値で 強度の予測を行う。

雨雪判別の気温には航空気温ガイダンスを利用する。

弱雪で  $850~\mathrm{hPa}$  と  $700~\mathrm{hPa}$  の相対湿度が共に 70% 未満の場合は降水なしとする  $^{17}$ 。

### 参考文献

藤吉康志, 若浜五郎, 遠藤辰雄, 入川真理, 小西啓之, 竹内政夫, 1983: 札幌における一冬間の降雪強度と視程の同時観測. 低温科学物理編, 42, 147-156.

後藤尚親, 2017: MSM ガイダンスの特性の変化. 平成 29年度数値予報研修テキスト, 気象庁予報部, 56-60.

Gultepe, I, M.D.Müller, and Z.Boybeyi, 2006: A New Visibility Parameterization for Warm-Fog Applications in Numerical Weather Prediction Models. J. Appl. Meteor. Climat., 45, 1469–1480.

原旅人, 2017: メソ数値予報システムの改良の概要. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 42-47.

井藤智史, 2011: 視程分布予想の開発. 平成 23 年度数値予報研修テキスト, 気象庁予報部, 25-29.

井藤智史, 2013: GSM 視程分布予想の開発. 平成 25 年度数値予報研修テキスト, 気象庁予報部, 58-62.

岩倉晋, 岡田菊夫, 1999: 東京国際空港における卓越視程の相対湿度依存性. 気象研究所研究報告, **50**, 81-90.

岩倉晋, 松本逸平, 大林正典, 2002: 航空ガイダンス. 平成 14 年度数値予報研修テキスト, 気象庁予報部, 46-51.

金井義文,満満男,工藤淳,2015:下層悪天予想図及び

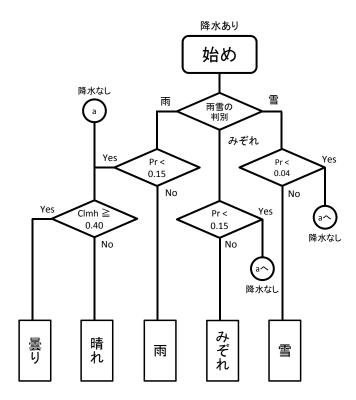


図 4.9.16 航空天気ガイダンスのフローチャート。各種文字の意味は以下のとおりである。Pr:前1時間降水量 [mm/h]、Cl:下層雲量、Cm:中層雲量、Ch:上層雲量、Clmh=1-(1-Cl)(1-Cm)(1-Ch):全雲量。雨雪判別のアルゴリズムについては第4.6節を参照されたい。

表 4.9.4 航空天気ガイダンスにおける降水強度を判定する 前 1 時間降水量 [mm] の閾値。

降水種別	弱	並	強
雨	0.15	1.5	8.0
雪	0.04	0.40	3.0
みぞれ	0.15	1.5	-

新しい狭域悪天予想図. 航空気象ノート第77号, 気象庁航空気象管理官.

松本智文, 高田伸一, 2010: 雲底確率ガイダンスの開発、 視程ガイダンスの改良. 平成 22 年度数値予報研修テ キスト, 気象庁予報部, 88-94.

大林正典, 2001: 航空気象予報支援資料. 平成 13 年度 数値予報研修テキスト, 気象庁予報部, 47-49.

鈴木弥幸, 1963: 視程と降雨(雪)強度の関係. 気象庁 研究時報, **15**, 483-487.

高田伸一, 2004: TAF-S 視程ガイダンス. 平成 16 年度 数値予報研修テキスト, 気象庁予報部, 53-55.

高田伸一, 2007: 航空気象予報ガイダンス. 平成 19 年度数値予報研修テキスト, 気象庁予報部, 87-89.

予報部予報課航空予報室, 2006: 新しい航空気象予報業 務について. 測候時報, **73**, 23-49.

米原仁, 2017: 全球数値予報システムの改良の概要. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 1-7.

<sup>&</sup>lt;sup>16</sup> 第 4.4 節を参照。MSM を入力として METAR, SPECI, METAR AUTO を観測値として空港向けの気温ガイダンス を作成する。

<sup>17</sup> 松本空港など高所にある空港で山を越えて流れてくる弱い 雪の予測は上空が乾いている時は空振りとなるため。

### 4.10 雲ガイダンス<sup>1</sup>

### 4.10.1 はじめに

空港の上空に低い雲があると、着陸時にパイロットが滑走路を視認出来る高度が低くなり、着陸が困難になるなど、雲が航空機の運航に与える影響は大きい。特に雲量が 5/8 以上となる最低雲層の雲底の高さ(シーリング、CIG)は航空機の運航に影響を与える指標の一つであり、例えば運航用飛行場予報 (TAF) の予報変化基準  $^2$  として  $100,\ 200,\ 600,\ 1000\ {\rm ft}^3$  が用いられている。また、CIG は代替飛行場の選定基準にも用いられている。

雲ガイダンスは空港上空の雲を予測するために 1993 年に運用を開始した。当時は線形重回帰を用いて前3 時間に CIG が 15000 ft 未満となる確率とその場合の CIG を予測対象としていた。1996年には利用するモデ ルが領域モデル (RSM) に変わり、CIG の予測対象高 度を 15000 ft から 7000 ft に変更、1998 年には診断手 法による上中下層雲の雲量と高さを追加した。2001年 には診断手法とカルマンフィルタを用いて空港上空38 観測を模して3層の雲を抽出する手法に変更した(大 林 2001)。この際に予測対象を前 3 時間の最低 CIG 時 における下から3層の雲量とその雲底高度に変更し、 CIG が閾値未満となる確率の予測を廃止した。2002年 にはメソモデル (MSM) を入力として予測対象時間を 前1時間としたガイダンスを追加し、頻度バイアス補 正を導入した (岩倉ほか 2002)。MSM の水平格子間隔 が 5 km となった 2006 年には前 1 時間を予測対象とす るガイダンスの予測手法をニューラルネットワークに 変更し、説明変数に地上と850 hPa の間の気温差を追 加した (高田ほか 2005)。2007年に MSM の予報時間 が 33 時間となる際に MSM ガイダンスの予報時間も延 長し、それに伴いカルマンフィルタを用いた RSM ガ イダンスを廃止して MSM ガイダンスに統合した (高田 2007)。2008年には説明変数に降水量を追加、予報対 象時刻や季節による層別化、頻度バイアス補正の調整 を行った(松澤ほか2008)。さらに目視観測のない時間 帯に自動観測である METAR AUTO<sup>6</sup> を利用するよう にした。2010年には前3時間にCIGが1000ft未満と なる確率 (PoC10) 及び 600 ft 未満となる確率 (PoC06) をロジスティック回帰を用いて予測する雲底確率ガイ

1 後藤 尚親

ダンスの運用を開始した(松本・高田 2010)。

本節では第 4.10.2 項でガイダンスの仕様を、第 4.10.3 項と第 4.10.4 項で統計検証と事例検証を通じてガイダンスの予測特性を示し、最後に第 4.10.5 項で利用上の留意点と今後の展望について述べる。

### 4.10.2 仕様

### (1) 雲ガイダンス

雲ガイダンスは空港上空の下から3層の雲量及びそ の雲底高度を予測するガイダンスである。雲ガイダン スの仕様を表 4.10.1 に示す。雲ガイダンスでは 38 層の 雲量を目的変数としており、目的変数の作成には目視 観測である METAR, SPECI<sup>7</sup>, SCAN<sup>8</sup> が優先して利用 されるが、目視観測のない時間帯には METAR AUTO も利用する。雲量は相対湿度が高くなると非線形に増 加することが知られている (大野・伊佐 1984) ため、統 計手法にニューラルネットワークを利用している。ネッ トワークは入力層のユニット数5個、中間層のユニット 数が6個、出力層のユニット数が1個の3層順伝播型で ある。中間層・出力層共に活性化関数はロジスティック 関数としている。他のニューラルネットワークのガイ ダンスの出力層は線形関数が多いが、雲量が $0\sim1$ の範 囲である事から雲ガイダンスでは出力層もロジスティッ ク関数を用いている。

### 作成方法

雲ガイダンスは以下の手順で作成する。

観測値から38層の雲量を作成する。

38 層各層において、ニューラルネットワークの重みと頻度バイアス補正の閾値を更新する。

MSM モデル面と気圧面から説明変数を作成する。 ニューラルネットワークで 38 層各層の雲量を予測 し、頻度バイアス補正で予測値を補正する。

38層から予測対象である3層を抽出し、雲ガイダンスの予測値とする。

雲ガイダンスでは予測対象である 3 層の値を直接予測する代わりに、まず空港上空の 38 層の雲量を求め、それから 3 層を抽出して予測値としている。そのために、まず目的変数として 38 層の雲量を作成する( )。雲の通報には雲の厚さの情報はないため、雲の厚さは一律に 2000 ft (霧の場合は 500 ft) としている。また通報には通常 3 層しか含まれないため、通報された最も高い雲層より上の層は欠測として扱う。観測時間は前 1 時間の最低 CIG の通報を利用するが、5/8 以上の雲がない場合は、第 1 層が最も低い通報を利用する。また鉛直視程が通報された場合にはその高度を第 1 層として扱い、それ以下の高度の雲量を 0 とする。

で求めた雲量を目的変数とし、38 層それぞれの予 測式を更新する( )。予測式は暖候期用と寒候期用で

<sup>&</sup>lt;sup>2</sup> 予報時間内に気象要素の重要な変化が予測される場合に、変化前後の状態、変化指示符、変化の開始時刻と終了時刻を TAF の変化群に示すための基準。

 $<sup>^{3}</sup>$  71 - 1 ft  $\simeq 0.3048$  m<sub>o</sub>

 $<sup>^4</sup>$   $0\sim1500$  ft は 100 ft 間隔、 $1500\sim5000$  ft は 500 ft 間隔、 $5000\sim10000$  ft は 1000 ft 間隔、 $10000\sim20000$  ft は 2000 ft 間隔とした 38 の層のこと。

<sup>5</sup> 航空気象定時観測気象報

<sup>6</sup> 航空気象自動観測気象報

<sup>7</sup> 航空気象特別観測気象報

<sup>8</sup> 航空気象観測所気象報

表 4.10.1 雲ガイダンスの仕様一覧

作成対象	空港
作成方法	ニューラルネットワーク (3層:活性化関数は中間層・出力層ともにロジスティック関数を使用)による予測に頻度バイアス補正を行う。
作成対象とするモデル	MSM
予報対象時間単位	1 時間
予報期間と間隔	FT=2 から FT=39 まで 1 時間間隔
逐次学習の有無	あり
説明変数	モデル面湿度、降水量、地上と 925 hPa との間の平均気温減率
目的变数	METAR, SPECI から算出した前 1 時間の最低 CIG 時の上空 38 層の雲量。目視観測がない時間帯は METAR AUTO を利用。
層別化処理の対象	作成対象地点(空港) 予報対象時刻( $1$ 時間毎 $1$ 日分) 季節(暖候期( $4 \sim 10$ 月) 寒候期( $11 \sim 3$ 月)) 寒候期・暖候期の係数切替前に $1$ か月の事前学習を行う。
備考	・ニューラルネットワークで空港上空の $38$ 層の雲量を求め、それを下から検索することによって $3$ 層の雲層を抽出する。 ・頻度バイアス補正の閾値は $0/8$ , $1/8$ , $3/8$ , $5/8$ , $8/8$ を使用。

### 表 4.10.2 雲底確率ガイダンスの仕様一覧

役 4.10.2 芸成惟平ガイノノ人の山脉 見					
作成対象	空港				
作成方法	ロジスティック回帰				
作成対象とするモデル	MSM				
予報対象時間単位	3 時間				
予報期間と間隔	間と間隔 FT=6 から FT=39 まで 3 時間間隔				
逐次学習の有無	なし				
説明変数	表 4.10.3 に示す説明変数から 6 つを選択する。				
目的变数	METAR, SPECI から算出した前 $3$ 時間の最低シーリングが $1000~\mathrm{ft}$ もしくは $600~\mathrm{ft}$ 未満 $(1)$ か否 $(0)$ か。目視観測がない時間帯は METAR AUTO を利用。				
層別化処理の対象	作成対象地点(空港)、季節(暖候期( $4\sim10$ 月)、寒候期( $11\sim3$ 月))、予報時間( $6$ 時間区切り)、予報対象時刻( $3$ 時間毎 $1$ 日分)				

### 表 4.10.3 雲底確率ガイダンスの説明変数一覧

必須選択	前 $3$ 時間における相対湿度の最大値 前 $3$ 時間における雲量 $(CVR)$ の最大値
二者択一	降水量の前 3 時間積算を周囲 20 km で平均した値 雪の降水量の前 3 時間積算を周囲 20 km で平均した値
3 つ選択	前 3 時間における地上と 925 hPa との間で算出した平均気温減率の平均値 前 3 時間における東西風の平均値 前 3 時間における南北風の平均値 前 3 時間における雲水量と雲氷量の和 (CWC) の最大値

層別化しているが、予測式の切り替えの際に予測値が 急変するのを軽減するために、予測に用いる 1 か月前 から予測式の更新を行う。

次に MSM のモデル面と気圧面から説明変数を作成する( )。説明変数は相対湿度、降水量、地上と 925~hPa との間の平均気温減率を利用する。相対湿度は予測対象である 38 層それぞれに最も近い高度のモデル面の $\pm 1$  層の相対湿度、降水量は前 3 時間降水量の正の平方根、平均気温減率は地上と 925~hPa の間で計算したものとする。なお、地上と 925~hPa の高度差が 100~m 以下である場合、平均気温減率の算出には 925~hPa の代わりに 850~hPa を用い、地上と 850~hPa の高度差が 100~m 以下である場合は  $5.0~K~km^{-1}$  とする。降水量と平均気温減率は空港を中心とした 20~km 領域の平均値を用いる。

と で得られた予測式と説明変数を用いてニューラルネットワークで 38 層の雲量を予測し、頻度バイアス補正で予測値を補正する。この 38 層の雲量を下の層から探索することによって、予測対象である 3 層の雲量を抽出する。3 層の抽出は空港の観測における雲の観測を模した次の手順で行う。

- 1. 38 層の雲量に対して、下から 1/8、3/8、5/8、7/8 以上となる層を最大で 4 層抽出する。
- 2. 抽出した層の高度差が高度の20%以下である場合、それら2つの層を一体として扱い、近接しないようにする。例えば $900~{\rm ft}$ で雲量1/8、 $1000~{\rm ft}$ で雲量4/8と抽出された場合、この $2~{\rm ft}$ で雲量 $0.900~{\rm ft}$ 0.
- 3. 下から3層の雲量とその雲底高度を予測値とする。

### (2) 雲底確率ガイダンス

雲底確率ガイダンスは前3時間内の最低 CIG が1000 ft もしくは600 ft 未満となる確率を予測対象とする。雲底確率ガイダンスの仕様を表4.10.2 に示す。目的変数の作成においては雲ガイダンスと同様に目視観測を優先して使い、目視観測のない時間帯には METAR AUTO を用いる。目的変数は前3時間内の最低 CIG が閾値未満となった場合に1、それ以外を0とする。

### 作成方法

雲底確率ガイダンスは以下の手順で作成する。

MSM 気圧面から説明変数を作成する。 ロジスティック回帰式で PoC10, PoC06 を求める。 PoC06 が PoC10 を上回る場合は、PoC06 を PoC10 で置き換える  $^9$ 。

説明変数は表 4.10.3 に示した説明変数から 6 つを 選択して利用する。候補となる変数は雲ガイダンスと 同様の要素に加え、冬季日本海側の降雪時における低 シーリングの予測精度向上を目的として前 3 時間の雪 の降水量を、海上等から移流してくる雲を考慮するた めに東西風及び南北風を、数値予報モデルで予測され る雲に関する量として雲量 (CVR) 及び雲水量と雲氷 量の和 (CWC) を用いる。相対湿度、東西風、南北風、 CVR、CWC は各地点のモデル標高から 1000 ft もし くは 600 ft の高度における値を利用する。回帰式に利 用する説明変数の選択方法は次の通りで、選択される 変数は地点や時刻などの層別ごとに異なる。なお、信 頼性の高い予測式の作成には一定数以上のデータを 重ねて利用して予測式を作成している。

- 1. 相対湿度と雲量は必須選択とする。
- 2. 前 3 時間降水量と前 3 時間の雪の降水量はどちら かを必ず選択する。
- 3. 必須選択と降水量の変数に加え、残り4つの変数 から3つを総当りで選んだ全組み合わせの回帰式 を作成する。
- 4. 全組み合わせのうち  ${
  m AIC}^{10}$  が最小となる組み合わせを採用し、その時の説明変数と予測式をガイダンスで利用する。

# 4.10.3 統計検証

雲ガイダンスと雲底確率ガイダンスの統計検証の結果を示す。検証には 2016 年 1 月から 12 月の 03, 09, 15, 21 UTC を初期時刻とする FT=6 から FT=27 の予測値を用いる。ただし予報時間別の検証である図 4.10.2 と図 4.10.6 には全初期値の FT=6 から FT=39 の予測値を用いる。また CIG においては値が小さい、つまり低

シーリングの予測精度が重要となるため、閾値未満の CIG が予測または観測されたときに「予測あり」または「観測あり」として検証する。検証は地域航空気象解説報でカテゴリー予想  $^{11}$  が発表される空港を対象とした。便宜上、以下では雲ガイダンスを Cld、雲底確率ガイダンスを PoC と記載することがある。なお、数値予報モデルには  $^{2018}$  年  $^{3}$  月現在運用している MSM の予測値を用いた。検証期間が現在運用している雲底確率ガイダンスの予測式作成期間 (後藤  $^{2017}$ ) と重なっているため、雲底確率ガイダンスの検証は  $^{1}$  か月抜きクロスバリデーション (第  $^{2}$   $^{2}$   $^{3}$ 

# (1) 雲ガイダンス

図 4.10.1 は雲ガイダンスの閾値別のエクイタブルスレットスコア (ETS) とバイアススコア (BI) である。低シーリングであるほど BI が小さく、ETS も低い傾向であることがわかる。これは低シーリングであるほど事例数が少ないことが影響していると考えられる。

図 4.10.2 と図 4.10.3 に閾値を 1000 ft とした予報時間別と予報対象時刻別の検証結果を示す。予報時間によらず ETS と BI は大きく変動しない。予報対象時刻別でみると、未明から朝にかけて事例数が多くなるものの、この時間の予測頻度及び精度は低い。これは放射霧の予測精度が十分でないことが影響している。

図 4.10.4 に月別の検証結果を示す。暖候期は BI は 1 に近く ETS も高いことがわかる。寒候期では暖候期に比べて事例数が少ないことや放射霧の予測精度が十分でない影響で、ETS が低い。特に 3 月と 11 月は ETS が他の月に比べて低いが、これは季節の変わり目であ

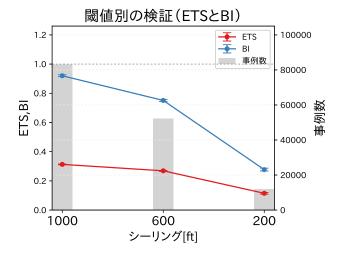


図 4.10.1 雲ガイダンスの閾値別の検証図。左軸折れ線グラフで ETS と BI を、右軸棒グラフで事例数を示す。ETS と BI のエラーバーは 95%信頼区間である。

 $<sup>^9</sup>$  ただし雲ガイダンスとの整合はとっていないため、例えば雲ガイダンスで CIG が  $400~\rm{ft}$  と予測されているのに  $\rm{PoC}06$  が低いといったことも起こりうる。

 $<sup>^{10}</sup>$  赤池情報量基準。詳しくは第 2.3.11 項 (1) を参照。

<sup>11</sup> 航空機の運航条件に対応した各種カテゴリー(VMC・IMC 及び周回進入条件及び雷の有無)を空港運用時間内の3時間 毎で予測して表形式で提供するもの(予報部予報課航空予報 室 2006)。

# 予報時間別の検証(ETSとBI) 1.2 - まTS(1000 ft) ま BI(1000 ft) ま TS(600 ft) ま BI(600 ft) 0.8 - 0.6 - 0.4 - ま TS(500 ft) ま BI(500 ft) ま BI(600 ft)

図 4.10.2 1000 ft と 600 ft を閾値とした雲ガイダンスの予報時間別の検証図。ETS と BI のエラーバーは 95%信頼区間である。

12 15 18 21 24 27 30 33 36 39

予報時間 [Hour]

ることや、暖候期と寒候期の予測式の切り替え月であるため、季節ごとの誤差特性を十分に修正出来ていないことが理由として考えられる。

図 4.10.5 は雲ガイダンスと雲底確率ガイダンスで閾値を 1000 ft とした場合の地点別検証である。事例の多い北日本において ETS-Cld が高い地点が多く、BI-Cld も 1 に近い。これに対し東日本や西日本、南西諸島の事例数の少ない地点では、予測頻度及び予測精度が低い。ETS-PoC と BI-PoC については後述する。

### (2) 雲底確率ガイダンス

ETS,BI

0.2

0.0

図 4.10.6 に雲底確率ガイダンスの予報時間別ブライアスキルスコア (BSS) を示す。PoC10 と PoC06 ともに予報時間が進むにつれ、徐々に BSS が低下する傾向がある。

図 4.10.7 と図 4.10.8 は雲底確率ガイダンスの信頼度曲線である。暖候期における信頼度曲線は PoC10, PoC06 ともに、図中点線で描画されている y=x の理想直線に近いが、50%以上の閾値で徐々に理想曲線から離れる。どちらも予測の発生確率が実況の相対出現頻度に比べて高いことから、雲底確率ガイダンスは高い確率値において、実況に比べてやや高い予測値となる傾向があることを示す。寒候期における PoC10, PoC06 ともに 30%程度から理想曲線から離れており、暖候期に比べて更に予測の確率が高く出やすい傾向がある。

次に雲ガイダンスとの比較として、図 4.10.9 と図 4.10.10 に 1000 ft 及び 600 ft を検証対象とした雲ガイダンスと雲底確率ガイダンスの ETS と BI を示す。検 証閾値 1000 ft 及び 600 ft ともに雲底確率ガイダンスの ETS(ETS-PoC) は約 30%で最大となり、同閾値の BI

# 予報対象時刻別の検証(ETSとBI)

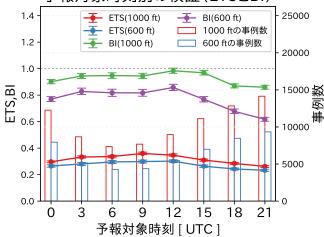


図 4.10.3 1000 ft と 600 ft を閾値とした雲ガイダンスの予報対象時刻別の検証図。左軸棒グラフで ETS と BI を、右軸棒グラフで各事例数を示す。ETS と BI のエラーバーは95%信頼区間である。

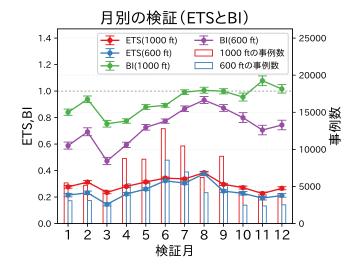


図 4.10.4 図 4.10.3 と同じ。ただし月別の検証結果を示す。

もほぼ1であることが分かる。ETS-PoC が最大となる 確率値付近では、雲ガイダンスの ETS(ETS-Cld) を上 回る精度となっている。このことは図 4.10.5 の Cld と PoC の比較でも確認出来る。多くの地点で ETS-PoC が ETS-Cld を上回っており、BI-PoC が BI-Cld より 1 に近い地点が多い。予報作業において、シーリングが 1000 ft もしくは 600 ft 未満となるか否かを判断する際 に、雲底確率ガイダンスの ETS が最大となる確率値を 目安とすることで、予報の精度が最も高くなることが 期待される。時系列の CIG を予測する際には雲ガイダ ンスを元に、雲底確率ガイダンスの確率値も参考に予 報を組み立てることが有効である。例えば雲ガイダン スの CIG の予測が 1000 ft 以上であっても、雲底確率 ガイダンスの予測確率が高い場合は雲ガイダンスを下 方修正するなど、雲底確率ガイダンスの予測も参考に しながら雲ガイダンスの予測を利用していただきたい。

# 雲ガイダンスと雲底確率ガイダンスの地点別検証(閾値1000 ft)

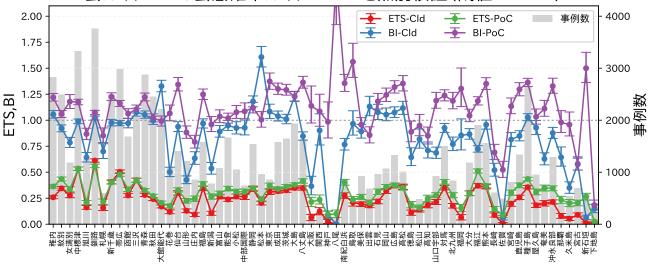


図 4.10.5 1000 ft を閾値とした雲ガイダンスと雲底確率ガイダンス ( PoC10、実況ありとする閾値は 30% ) の地点別検証図。 左軸折れ線グラフで ETS と BI を、右軸棒グラフで事例数を示す。ETS と BI のエラーバーは 95%信頼区間である。

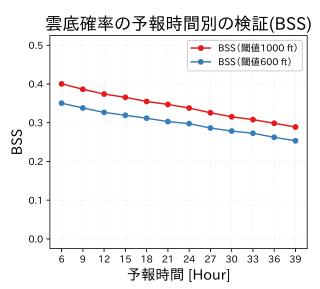


図 4.10.6 雲底確率ガイダンスの予報時間別の検証図。PoC10 と PoC06 の BSS を示す。

### 4.10.4 事例検証

雲ガイダンスと雲底確率ガイダンスの予測例として釧路空港で低シーリングとなった事例を示す。図 4.10.11 は 2016 年 1 月 13 日 21 時の日本域地上天気図、図 4.10.12 は翌日 14 日 3 時の 1 時間解析雨量である。日本海から東進する低気圧の影響で、釧路空港周辺では1 時ごろから降水域がかかり、断続的に降雪があった。

図 4.10.13 に釧路空港における CIG の観測値と 2016 年 1 月 12 日 18UTC を初期値とする雲ガイダンス、雲底確率ガイダンスの予測値を示す。釧路空港の観測では降雪の始まった 1 時ごろから CIG が急速に下がり、3 時には 1000 ft、7 時には最も低い 500 ft まで低下した。その後、9 時に降雪がやみ、CIG は 2000 ft となっ

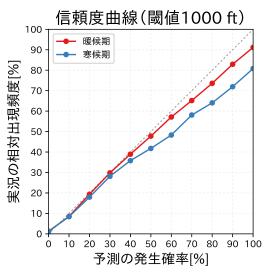


図 4.10.7 雲底確率ガイダンスの季節別信頼度曲線。検証対象は PoC10。暖候期は 4 月から 10 月、寒候期は 11 月から 3 月としている。

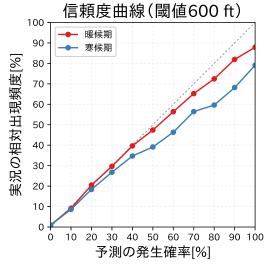


図 4.10.8 図 4.10.7 と同じ。ただし検証対象は PoC06。

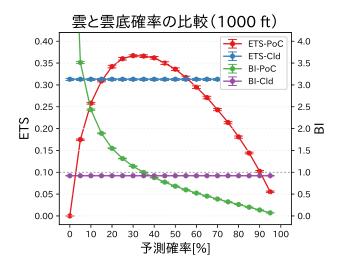


図 4.10.9 1000 ft を閾値とした雲ガイダンスと雲底確率ガイダンス (PoC10) の比較図。左軸で ETS を、右軸で BI を示す。エラーバーは 95%信頼区間である。

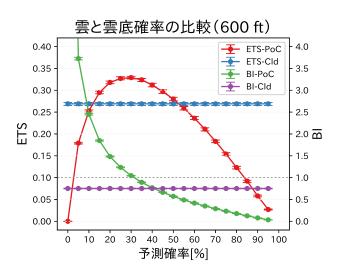


図 4.10.10 図 4.10.9 と同じ。ただし 600 ft を閾値として いる。

た。雲ガイダンスによる CIG の予測は実況からやや遅れてはいるが、6 時に 500 ft、8 時に 200 ft を予測し、実況と対応が良い。CIG の予測が実況に比べて全体的に 3 時間ほど遅れているのは、MSM の降水予測が実況に比べて 3 時間ほど遅れていたことが理由と考えられる(図略)。雲ガイダンスは 1 時間単位で予測が行われるため、数値予報モデルの予測の位置ずれの影響を大きく受ける。このため、降水分布などの数値予報モデルの予測と実況の位置ずれに留意して、雲ガイダンスの予測を修正していただきたい。

雲底確率ガイダンスでは実況で低シーリングとなる時間帯に PoC10 及び PoC06 を高く予測しており、実況を良く捕捉している。また雲ガイダンスでは 10 時頃まで実況にない低シーリングを予測しているが、雲底確率ガイダンスでは 12 時の確率値は低い。統計検証で示したとおり、利用している説明変数の違い(雪の降

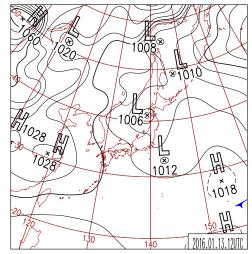


図 4.10.11 日本時間 2016 年 1 月 13 日 21 時の日本域地上 天気図。

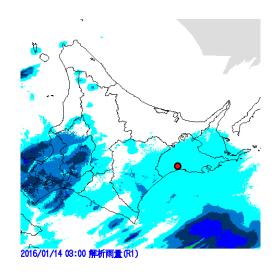


図 4.10.12 日本時間 2016 年 1 月 14 日 3 時の解析雨量。図 の赤丸は釧路空港の位置を示す。

水量や雲量)などの影響で、雲ガイダンスより雲底確率ガイダンスの方が統計的には精度が高い。 $1000~{\rm ft}$ もしくは $600~{\rm ft}$ 未満となるか否かを判断する際には雲底確率ガイダンスの確率値が有効である。

### 4.10.5 利用上の留意点と今後の展望

本節では雲ガイダンスと雲底確率ガイダンスの開発や改良、仕様、検証結果を示した。統計検証や事例検証で示したとおり、低シーリングの発生を予測する際には雲底確率ガイダンスが有効である。雲底確率ガイダンスの ETS が最大となる約 30%を目安として低シーリングが発生するか否かを判断することで、精度の高い予測となることが期待される。また雲ガイダンスは1時間単位で予測されるため、数値予報モデルの予測と実況の差に着目することで、予測の修正に役に立つ場合がある。

最後に雲ガイダンスと雲底確率ガイダンスの今後の

# 2016年1月14日の釧路の低シーリング事例

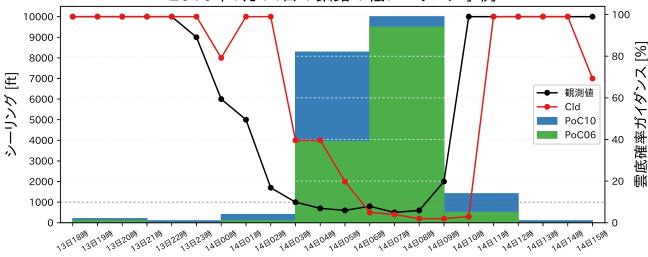


図 4.10.13 日本時間 2016 年 1 月 14 日の釧路空港のシーリングの観測値とガイダンス。ガイダンスは 2016 年 1 月 12 日の 18UTC 初期値で、左軸折れ線グラフで CIG の観測値と雲ガイダンスを、右軸棒グラフで雲底確率ガイダンス (PoC10 と PoC06) を示す。

展望を述べる。雲ガイダンスにおいてはニューラルネットワークの再構築を試みたい。特に説明変数の標準化など、第2.6節で紹介されているニューラルネットワークの学習を効率的に行うテクニックを利用し、精度向上を目指したい。説明変数についても雲底確率ガイダンスで利用されている雲量や雪の降水量などの追加を検討していきたい。また統計検証で示したとおり、季節の変わり目において精度が低い課題があるため、層別化の再検討も行いたい。

雲底確率ガイダンスについても層別化の見直しを行いたい。現在の層別化は雲ガイダンスを参考にしているが、より低シーリングの発生状況に適した層別化を検討する必要があるだろう。また説明変数は各空港のモデル標高から 1000 ft もしくは 600 ft 上空の高度の予測を利用しているが、予測対象としている閾値「未満」の高度の予測を含めていない。この「未満」をガイダンスに反映するために、地上から予測対象高度間におけるモデルで予測される物理量と観測値の関係性を調査し、有効な説明変数を検討したい。

### 参考文献

後藤尚親, 2017: MSM ガイダンスの特性の変化. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 56-60. 岩倉晋, 松本逸平, 大林正典, 2002: 航空ガイダンス. 平成 14 年度数値予報研修テキスト, 気象庁予報部, 46-51.

松本智文, 高田伸一, 2010: 雲底確率ガイダンスの開発、 視程ガイダンスの改良. 平成 22 年度数値予報研修テ キスト, 気象庁予報部, 88-94.

松澤直也, 藤枝鋼, 高田伸一, 古市豊, 2008: 航空気象予 報ガイダンスの検証. 平成 20 年度数値予報研修テキ スト, 気象庁予報部, 82-90.

大林正典, 2001: 航空気象予報支援資料. 平成 13 年度 数値予報研修テキスト, 気象庁予報部, 47-49.

大野久雄, 伊佐真好, 1984: GMS 観測に基づく総観スケール雲量と相対湿度の統計的関係. 天気, **31**, 493-495

高田伸一, 工藤淳, 新美和造, 2005: 航空気象予報. 平成 17年度数値予報研修テキスト, 気象庁予報部, 49-62. 高田伸一, 2007: 航空気象予報ガイダンス. 平成 19年 度数値予報研修テキスト, 気象庁予報部, 87-93.

予報部予報課航空予報室, 2006: 新しい航空気象予報業 務について. 測候時報、**73**, 23-49.

### 乱気流指数 1 4.11

### 4.11.1 はじめに

航空機の運航にとって、乱気流を回避することは安 全性や快適性にとって重要である。特に晴天乱気流は レーダーや目視での回避が困難であるため、予測に基 づいて事前に回避することが重要である。気象庁はシ グメット情報や国内悪天予想図 (FBJP) などにより乱 気流の予測情報を提供しており、これらの情報作成を 支援するために、数値予報課では乱気流指数(以下で はTBindexと書く)を開発し運用している。

乱気流とは航空機に影響を及ぼす乱流である。旅客機 に影響を与える乱流の水平スケールは数 10~2,000 m 程 度といわれている (中山 2010; Lane et al. 2012; Sharman and Lane 2016)。現在の現業数値予報モデルでは これを直接的に解像することはできないため、乱気流 の予測においてはその発生のポテンシャルを周囲の気 象状態から判断することになる。乱気流を発生させる 気象要因は様々あり、代表的なものとしては、晴天乱 気流の主な要因であるケルビン・ヘルムホルツ (KH)波 (Browning et al. 1973; Shapiro 1980; Lilly 1986) や活 発な対流雲 (MacCready 1964; MacPherson and Isaac 1977) およびその周辺での乱気流 (Trier et al. 2010; Lane et al. 2012)、風下や鉛直上方に伝播した山岳波 に伴う乱気流 (Gerbier and Berenger 1961; Lester and Fingerhut 1974; Scorer 1978; Shen and Lin 1999)、中 層雲の雲底下での固体降水の昇華冷却に伴うベナール 対流 (Kudo 2013; Kudo et al. 2015)、衛星画像で特 徴的なパターンとして見られるトランスバースバンド (気象衛星センター 2002; Knox et al. 2010; Trier and Sharman 2016) などが挙げられる。

従来の乱気流予測では、晴天乱気流については鉛直 シアーなどの数値予報から求めた指数を用いて予測し、 対流雲や山岳波などによる乱気流は安定度や降水量、下 層風の予測などから診断的に予測するという手法が用 いられていた。しかしながら鉛直シアーでは晴天乱気 流の予測が不十分な場合があることや、診断的な予測手 法では乱気流のポテンシャルを定量的に判断すること が難しいという問題があった。そこで、様々な要因で発 生する乱気流を総合的に予測する指数として TBindex を開発した。TBindex は乱気流の要因別に算出した複 数の指数をロジスティック回帰で組み合わせた指数で ある。これにより、様々な要因による乱気流を一つの 指標で予測することが可能になるとともに、単独の指 数での予測と比べて予測精度が大幅に向上した。また、 ロジスティック回帰で求めた確率値を調整することで、 高度に依らず一つの閾値で最適な予測を行うことが可 能となっている。

MSM の予測値から算出する国内航空悪天 GPV の

TBindex は、2010年2月に数値予報ルーチンでの運用 を開始し、2010年6月には係数の変更による改良を行 い(工藤 2010)、2011年3月からは空域予報現業での利 用が開始されている。また、2014年8月にはパラメー タの調整や説明変数の改良を行った (工藤 2014)。LFM の予測値に基づく LFM 航空悪天 GPV の TBindex は、 国内航空悪天 GPV の TBindex と同じ手法、同じ係数 で 2012 年 6 月から作成を開始しており、2014 年 3 月 からは下層悪天予想図と狭域悪天予想図での乱気流域 の予測に用いられている (金井ほか 2015)。また、2013 年には福岡 FIR 全域を対象とするために GSM 版の TBindex を開発し (三輪 2013)、パラメータの調整や 説明変数の改良を行った後に、2016年6月から数値予 報ルーチンでの運用と現業利用が開始されている (井 藤 2016)。

本節では乱気流指数の解説を行う。まず第4.11.2項 で TBindex の仕様と作成方法を述べ、第 4.11.3 項で各 航空悪天 GPV の TBindex の検証結果を示す。続いて 第 4.11.4 項で特徴的な予測事例を示した後、第 4.11.5 項で利用上の留意点等を述べる。なお、本節で使用す る空域予報に関連する用語の解説を付録 4.11.A にまと めたので適宜参照していただきたい。

# 4.11.2 仕様

モデル毎の TBindex の仕様を表 4.11.1 に示す。 FBJP では並 (MOD) 以上の乱気流を予測対象として いることから、目的変数は C-PIREP で報じられた並 (MOD) 以上の乱気流の有無としている。統計手法は 2 クラス分類に適したロジスティック回帰を用いてい る $^2$ 。また、C-PIREP で MOD 以上の乱気流が予測さ れる割合は1%以下(工藤2012)と低いため、頻度の少 ない現象に適した一括学習を用いている。乱気流は主 に高度によって発生要因が異なるため、フライトレベ ル (FL) 面の各高度で回帰式を層別化している。ロジス ティック回帰で得られる予測値は乱気流発生確率に相当 するが、C-PIREP はウィンドプロファイラや LIDEN などと異なり非定常的な観測であることから、算出さ れた予測値を確率と見なすことはできない。また、高 度によって予測に適切な閾値が変わってしまうと利用 する上で不便である。そこでロジスティック回帰で得 られた予測値を調整し、高度によらず最適な閾値が3.0 になるようにした値を TBindex として出力している。 これについては以下の(3)で述べる。

### (1) 説明変数の候補

TBindex に使用している説明変数を表 4.11.2 に示 す。乱気流は様々な要因で発生するが、従来から提案 されている乱気流予測のための指標は主に晴天乱気流

<sup>&</sup>lt;sup>2</sup> C-PIREP で報じられる乱気流強度の階級を目的変数とし た順序ロジスティック回帰も試したが、通常のロジスティッ ク回帰と比べて MOD 以上の乱気流に対する予測精度に差は 見られなかった。

<sup>1</sup> 工藤 淳

表 4.11.1 モデル毎の TBindex の仕様

	全球航空悪天 GPV	北太平洋航空悪天 GPV	国内航空悪天 GPV	LFM 航空悪天 GPV		
作成対象モデル		GSM	MSM	LFM		
作成方法		ロジスティッ	クロ帰			
格子間隔	1.25 度	0.5 度	40 km	10 km		
格子系	等緯原	度経度座標	ポーラース	ステレオ座標		
対象領域	全球	$10^{\circ}\text{S}-65^{\circ}\text{N}, 80^{\circ}\text{E}-110^{\circ}\text{W}$	日本及びその周辺	日本周辺		
鉛直層	700 hPa ~ 100 hPa の 9 層の等圧面	$\mathrm{FL010} \sim \mathrm{FL550}  \mathcal{O}  2,0$	000 ft <b>毎の</b> FL <b>面</b>	FL010 ~ FL450 の 1,000 ft 毎の FL 面		
予報対象時間	6 時間毎の正時	3 時間毎の正時	毎正時	毎正時		
予報時間	$FT = 6 \sim 36$	$FT = 3 \sim 36$	$FT = 1 \sim 39$	$FT = 1 \sim 9$		
予報時間間隔	6 時間	6 時間 3 時間 1 時間				
目的変数	C-PIREP で通報された並 (MOD) 以上の乱気流の有無					
層別化	各 FL 面(全球航空悪天 GPV では FL 面で作成した値を等圧面に内挿している)					
調整手法	スキルスコアが最大となる閾値が 3.0 になるようにロジスティック回帰による予測値を変換する					
逐次学習	なし					

表 4.11.2 TBindex に用いている説明変数とその略号および意味。 ${\rm G/P}$  の番号は全球航空悪天  ${\rm GPV}$  および北太平洋航空悪天  ${\rm GPV}$  で用いている説明変数、 ${\rm M/L}$  の番号は国内航空悪天  ${\rm GPV}$  および LFM 航空悪天  ${\rm GPV}$  に用いている説明変数を表し、数字は寄与最大説明変数(第 4.11.2 項 (4) を参照)として割り振られる番号を、空欄は各航空悪天  ${\rm GPV}$  には用いていない説明変数であることを示す。

G/P	M/L	名称	略号	意味
7	1	斜方シアー	SWS	斜め方向に算出した風のシアー
1		鉛直シアー	VWS	鉛直方向の風のシアー
2	2	対流雲中乱気流指数	CONV	対流雲の雲中および雲頂での強い上昇流や VWS の強化により発生する乱気流を予測する指数
3	3	風下山岳波指数	MTW1	風下山岳波に伴い山頂付近の安定層下で発生する乱気流を予 測する指数
4	4	鉛直伝播山岳波指数	MTW2	鉛直上方に伝播した山岳波の砕波による乱気流を予測する指数
5	5	中層雲底乱気流指数	BASETB	中層雲底下でのベナール対流による乱気流を予測する指数
6	6	トランスバースバンド指数	TRAV	トランスバースバンドに伴う乱気流を予測する指数
	7	水平シアー	HWS	水平方向の風のシアー
8	8	VHS	VHS	VWS × (流れに沿った方向の風の水平シアー)× (風速)
9	9	リチャードソン数	RI	KH 不安定の発生を予測する指数
10	10	気温の水平傾度	GRADT	等フライトレベル面で算出した気温の水平傾度の大きさ

を対象とした指数であった。そこで晴天乱気流以外の乱気流を予測するための指数として、風下山岳波指数 (MTW1)、鉛直伝播山岳波指数 (MTW2)、中層雲底乱気流指数 (BASETB)、トランスバースバンド指数 (TRAV)、対流雲中乱気流指数 (CONV)を開発した。また、鉛直シアーだけでは晴天乱気流の予測に不十分な場合があることから、斜方シアー (SWS) と VHS を開発した。これらの指数に従来から用いられている鉛直シアー (VWS) やリチャードソン数 (RI) などを説明変数の候補に加え、ロジスティック回帰で係数を求める。説明変数の候補として使用される各種指数の算出方法は付録 4.11.B にまとめた。

説明変数の候補の中には変数間の相関が強いものも 含まれている。これらも全て使用して回帰式を作成す れば予測精度としては多少良くなる可能性はあるが、 係数が不安定になったり予測結果の解釈が難しくなったりする。そこで初めに説明変数間の相関を調べ、相 関の強い変数は予め用いないようにしている。

表 4.11.3 に 2016 年の国内航空悪天 GPV から算出した各種指数間の標本相関係数を示す。表では、相関が強いもの (絶対値が 0.6 以上 )を赤、中程度のもの (絶対値が  $0.4 \sim 0.6$  )を黄色、弱いもの (絶対値が  $0.4 \sim 1.6$  )を青で色分けしている。表に示したように、MTW1, MTW2, BASETB, TRAV は他のどの指数と比べても相関が弱く、独立に近い指数であるといえる。また CONV と GRADT も概ねほかの指数との相関は弱い。鉛直シアーに関連した指数である SWS, VWS, VHS, 水平シアー (HWS), Ellrod の指数 (TI1, TI2) は変数間の相関が強いため、予測精度を考慮した上で、全球および北太平洋航空悪天 GPV ではこれらの中から SWS.

表 4.11.3 国内航空悪天 GPV から算出した各種指数間の標本相関係数の例。2016 年の 1 年間の国内航空悪天 GPV のデータを元に、2,000 ft 毎の高度別に求めた相関係数を全高度で平均した値。相関係数の絶対値が 0.6 以上を赤、0.4 以上 0.6 未満を黄、0.4 未満を青で示している。

	SWS	VWS	CONV	MTW1	MTW2	BASETB	TRAV	HWS	VHS	RI	GRADT	TI1	TI2
SWS	1.00	0.84	0.44	-0.04	0.13	0.06	0.20	0.64	0.64	-0.48	0.47	0.76	0.63
VWS	0.84	1.00	0.28	-0.05	0.08	0.07	0.21	0.27	0.61	-0.63	0.47	0.71	0.59
CONV	0.44	0.28	1.00	-0.05	0.05	0.05	0.15	0.42	0.40	-0.23	0.22	0.36	0.20
MTW1	-0.04	-0.05	-0.05	1.00	-0.08	-0.00	0.02	-0.03	-0.02	-0.00	-0.01	-0.05	-0.03
MTW2	0.13	0.08	0.05	-0.08	1.00	-0.03	-0.09	0.13	-0.04	-0.07	0.10	0.10	0.08
BASETB	0.06	0.07	0.05	-0.00	-0.03	1.00	-0.03	0.02	0.01	-0.01	0.14	0.06	0.07
TRAV	0.20	0.21	0.15	0.02	-0.09	-0.03	1.00	0.05	0.24	-0.11	0.07	0.15	0.09
HWS	0.64	0.27	0.42	-0.03	0.13	0.02	0.05	1.00	0.49	-0.14	0.29	0.69	0.60
VHS	0.64	0.61	0.40	-0.02	-0.04	0.01	0.24	0.49	1.00	-0.31	0.40	0.68	0.58
RI	-0.48	-0.63	-0.23	-0.00	-0.07	-0.01	-0.11	-0.14	-0.31	1.00	-0.19	-0.37	-0.29
GRADT	0.47	0.47	0.22	-0.01	0.10	0.14	0.07	0.29	0.40	-0.19	1.00	0.42	0.37
TI1	0.76	0.71	0.36	-0.05	0.10	0.06	0.15	0.69	0.68	-0.37	0.42	1.00	0.87
TI2	0.63	0.59	0.20	-0.03	0.08	0.07	0.09	0.60	0.58	-0.29	0.37	0.87	1.00

表 4.11.4 国内航空悪天 GPV での各高度の説明変数の寄与量の例。ロジスティック回帰で求めた回帰係数に各高度における説明変数の代表値を掛けた値を示す。切片に対して 10%以上の値を青で表示している。空欄は係数が 0 であることを示す。2011年から 2012年の 2年間のデータを使用。

高度	切片	SWS	CONV	MTW1	MTW2	BASETB	TRAV	HWS	VHS	RI	GRADT
FL410	-7.58	0.70	0.33				0.14	1.18	0.29	-0.02	
FL390	-7.11	0.86	0.42		0.95		0.24	0.81	0.37	-0.02	
FL370	-6.80	0.77	0.45		1.86		0.33	0.66	0.46	-0.02	
FL350	-6.74	1.27	0.39		1.82		0.41	0.51	0.44	-0.01	
FL330	-6.67	2.06	0.61		1.06		0.30	0.29	0.44		
FL310	-6.47	2.46	0.87		0.80		0.33		0.52		
FL290	-6.51	2.74	0.88		0.97		0.56		0.46	0.00	
FL270	-6.39	2.93	0.92		0.86		0.70		0.27	-0.01	
FL250	-6.17	2.72	1.07		0.79	0.79	0.69	0.25	0.19	-0.01	-0.18
FL230	-6.27	2.55	1.06		0.58	1.01	0.29	0.21	0.23		
FL210	-6.22	2.47	1.28		0.68	1.70	0.19	0.24	0.25	0.00	0.16
FL190	-5.84	2.23	1.20	1.15	0.62	1.96		0.32	0.23		
FL170	-5.68	2.35	1.20	0.51	0.98	1.99		0.34			
FL150	-5.26	2.37	1.09	1.65	0.76	1.68		0.37	-0.11		0.27
FL130	-4.54	1.85	0.83	1.60	0.32	1.01		0.48			0.22
FL110	-4.40	1.55	0.70	1.19	0.36	0.71		0.53	0.08	-0.01	0.25
FL090	-4.40	1.40	0.63	0.62				0.46	0.19	-0.01	0.26
FL070	-4.35	1.22	1.00	0.97				0.36	0.22	-0.01	0.13
FL050	-4.33	1.23	1.30	1.24				0.38	0.17	-0.01	
FL030	-4.23	1.54	1.05	1.02					0.21	-0.02	

VWS, VHS を、国内および LFM 航空悪天 GPV では SWS, HWS, VHS を使用している。各航空悪天 GPV で使用している説明変数の候補は、表 4.11.2 の左の 2 列に数字が書かれている指数となる。

### (2) 回帰式の作成

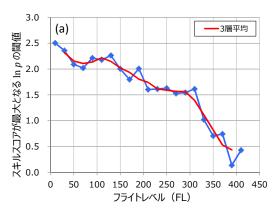
ロジスティック回帰で予測式を作成する際、表 4.11.2 で挙げた説明変数の候補の中から、赤池情報量基準 (AIC) に基づいて説明変数を選択する。目的変数は、C-PIREP で MOD 以上の乱気流が通報された場合を 1、MOD 未満を 0 とする。C-PIREP では晴天や雲中など、飛行中の雲の状態が報じられることもあるが、TBindex では全ての乱気流を予測対象とするため、C-PIREP でも雲の状態に関わらず全ての通報を利用して

回帰式を作成する。

乱気流の発生要因は高度によって異なる。例えば風下山岳波は主に下層で発生し上層では発生しない。またトランスバースバンドに伴う乱気流は主に上層で発生し下層では発生しない。高度による違いを適切に表現するため、回帰式は 2,000 ft 毎の FL 面で層別化する。

表 4.11.4 に国内航空悪天 GPV での各高度における 説明変数の代表的な寄与量  $^3$  を示す。表では切片に対して 10%以上の寄与量を持つ説明変数の背景を青で塗りつぶしている。空欄は、AIC に基づく判定でその高

<sup>&</sup>lt;sup>3</sup> 寄与量は説明変数の値に回帰係数を掛けて求められる値である。ここでは説明変数の有効性を比較するため、各高度・説明変数での代表的な寄与量として、スキルスコアが最大または極大となる説明変数の値を用いて寄与量を算出した。



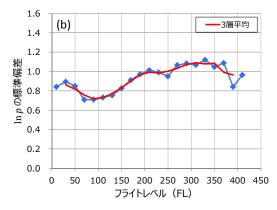


図 4.11.1 2011 年から 2012 年の国内航空悪天 GPV のデータを用いて作成した (a) スキルスコアが最大となる  $\ln p$  の閾値、 (b)  $\ln p$  の標本標準偏差、および上下 3 層で平均した値。

度では説明変数として採用されなかったことを示している。SWS と CONV は全ての高度で採用され、上層の一部を除いては寄与量が大きくなっている。MTW1, MTW2, BASETB はそれぞれ、下層、中・上層、中層を中心に説明変数として採用されており、採用されている高度では寄与量が大きい。TRAV は FL210 以上の全ての高度で説明変数として採用されているが、寄与量が大きいのは一部の高度のみである。HWS と VHS は多くの高度で採用されているが、一部の高度を除いて寄与量は小さい。RI と GRADT は全般に寄与量が小さい。

### (3) 確率値の調整

TBindex では、ロジスティック回帰で求めた確率値を調整することで、MOD 以上の強度の乱気流に対して全ての高度で同じ閾値 (TBindex = 3.0) を用いて最適な予測が行えるようにしている。各格子点におけるTBindex は、次の式で求められる。

TBindex = 
$$\frac{\ln p(i, j, k) - \mu(k)}{\sigma(k)} + 3.0$$
 (4.11.1)

ここで i,j,k は格子点の水平 (i,j) および鉛直 (k) 座標、p はロジスティック回帰で求めた  $0\sim1$  の確率値に 100 を掛けて%で表した確率である。また  $\mu(k)$  と  $\sigma(k)$  は、 $\ln p$  の値を変えながら高度別にスキルスコアを求めた際に、スキルスコアが最大になるときの  $\ln p$  の値と、 $\ln p$  の標本標準偏差である。右辺第一項は  $\ln p$  に対するスキルスコアの分布を正規分布とみなした場合に、平均 0、標準偏差 1 の正規分布になるように標準化する操作である。これに 3.0 を加えることで、多くの格子点で TBindex が正の値を持つようになる。また 3.0 を閾値として TBindex を利用することで、高度に依らずスキルスコアが最大となる予測を行うことが可能になると期待される。

 $\mu$  と  $\sigma$  は学習期間のデータから得られる値であるが、そのまま用いると学習データに過度に適合してしまい、予測には適さない可能性がある。そこで、学習データから得られた層別化した各高度での  $\mu$  と  $\sigma$  を上下 3 層

で平均した値を用いている。図 4.11.1 は、2011 年から 2012 年の国内航空悪天 GPV のデータを用いて高度別 に求めた  $\mu,\sigma$  と上下 3 層での平均値である。平均値を 用いることで層毎の変化が滑らかになっている。

# (4) 寄与最大説明変数の算出

乱気流の予測や実況監視において、着目すべきポイントは発生要因により異なる。このため空域予報作業では、単に乱気流が発生するか否か(TBindex が大きいか小さいか)だけではなく、その発生要因(どの説明変数の寄与が大きいか)を知ることが重要となる。そこでTBindexへの寄与が最も大きい説明変数を示す要素として寄与最大説明変数を算出している。具体的には、各説明変数に表 4.11.2 の左の 2 列に示した番号を割り振っておき、各格子点で寄与量(係数 × 説明変数値)が最大となった説明変数の番号を寄与最大説明変数をりが最大となった説明変数の番号を寄与最大説明変数として出力している。寄与最大説明変数を利用することでTBindexが大きくなった理由を直接的に知ることができる。

### 4.11.3 統計検証

ここでは北太平洋航空悪天 GPV、国内航空悪天 GPV、および LFM 航空悪天 GPV の TBindex の検 証結果を示す。また、比較のために従来の乱気流予 測で主に用いられていた VWS の検証結果も示す。 以下では表記を簡単にするため、北太平洋航空悪天 GPV の TBindex と VWS をそれぞれ P\_TBindex と P\_VWS、国内航空悪天 GPV の TBindex と VWS を それぞれ M\_TBindex と M\_VWS、LFM 航空悪天 GPV の TBindex と VWS をそれぞれ L\_TBindex と L\_VWS と書く。実況データとしては、学習データと同様に C-PIREP を使用し、天候状態(晴天、雲中等)による区別 はしない。実況があった地点の最近接格子での TBindex および VWS の値を予測値として検証する。時刻につ いては、北太平洋航空悪天 GPV では予報時間の前後 90 分以内、国内航空悪天 GPV と LFM 航空悪天 GPV では予報時間の前後30分以内を観測時刻とした実況と

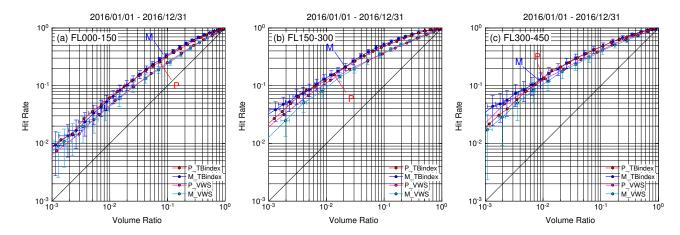


図 4.11.2 2016 年 1 月から 12 月の MOD 以上の乱気流に対する TBindex と VWS の高度別・閾値別の捕捉率 (Hit Rate) と体積率 (Volume Ratio)。北太平洋航空悪天 GPV (00, 06, 12, 18 UTC 初期値)の FT=12, 15 と国内航空悪天 GPV (03, 09, 15, 21 UTC 初期値)の FT=8~13 を比較。(a) FL000-FL150, (b) FL150-FL300, (c) FL300-FL450。P\_TBindex と P\_VWS はそれぞれ北太平洋航空悪天 GPV の TBindex と VWS の検証結果を、M\_TBindex と M\_VWS はそれぞれ国内航空悪天 GPV の TBindex と VWS の検証結果を示す。TBindex は 0 から 0.2 毎に、VWS は 0 から 1 kt/1000ft 毎に検証した結果を示している。図中の P と M は、P\_TBindex と M\_TBindex が 3.0 の場合の検証結果。MSM の TBindex と VWS のみにプートストラップ法で求めた 95%信頼区間を示すエラーバーを付加している。

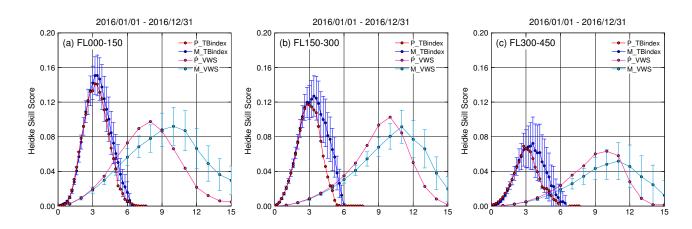


図 4.11.3 2016 年 1 月から 12 月の MOD 以上の乱気流に対する TBindex と VWS の高度別・閾値別のスキルスコア。北太平洋航空悪天 GPV (00, 06, 12, 18 UTC 初期値)の FT=12, 15 と国内航空悪天 GPV (03, 09, 15, 21 UTC 初期値)の FT=8~13 を比較。(a) FL000-FL150, (b) FL150-FL300, (c) FL300-FL450。P.TBindex と P.VWS はそれぞれ北太平洋航空悪天 GPV の TBindex と VWS の検証結果を、M.TBindex と M.VWS はそれぞれ国内航空悪天 GPV の TBindex と VWS の検証結果を下す。MSM の TBindex と VWS のみにプートストラップ法で求めた 95%信頼区間を示すエラーバーを付加している。

比較を行う。なお、乱気流の遭遇地点については、「地点 A で揺れた」と通報される場合と、「地点 A から地点 B まで揺れた」と通報される場合の 2 つのパターンがある。後者の場合は 2 地点間の中点を実況があった場所とするが、2 地点間の距離が水平 240 km 以上または鉛直 6,000 ft 以上離れていた場合は中点に代表性がないと判断し、検証には使用しない。検証対象領域は日本周辺の北緯 20 度から 45 度、東経 110 度から 150 度の範囲とする。検証期間は 2016 年 1 月から 12 月の 1 年間で、北太平洋航空悪天 GPV は 2017 年 5 月 25 日に変更された GSM の試験データから作成し、国内航空悪天 GPV は 2017 年 2 月 28 日に変更された MSM

の試験データから作成している。本期間にLFMの予測精度に大きく影響を与えるような変更は行われなかったため、LFM 航空悪天 GPV はルーチンのデータを用いて検証を行った。

図 4.11.2 と図 4.11.3 に  $P_TBindex$  と  $M_TBindex$  および、 $P_TVWS$  と  $M_TVWS$  の検証結果を示す。 対象時刻を揃えるため、北太平洋航空悪天 GPV は 00,06,12,18 UTC 初期値の FT=12,15、国内航空悪天 GPV は 03,09,15,21 UTC 初期値の  $FT=8\sim13$  の予測値を用いた。

図 4.11.2 は TBindex の閾値別に MOD 以上の乱気流を予測した場合の捕捉率 (Hit Rate) と体積率 (Volume

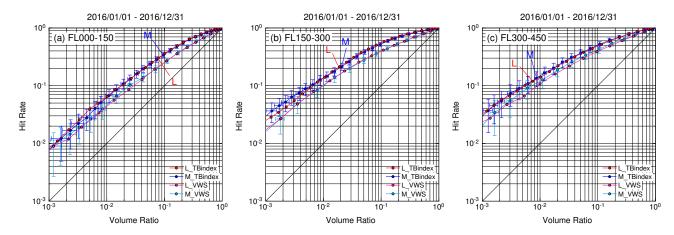


図 4.11.4 図 4.11.2 と同じ。ただし、FT=4~9 について、国内航空悪天 GPV (03, 09, 15, 21 UTC 初期値)と LFM 航空悪天 GPV (03, 09, 15, 21 UTC 初期値)を比較した結果。M.TBindex と M.VWS はそれぞれ国内航空悪天 GPV の TBindex と VWS の検証結果を、L.TBindex と L.VWS はそれぞれ LFM 航空悪天 GPV の TBindex と VWS の検証結果を示す。図中の M と L は、M.TBindex と L.TBindex が 3.0 の場合の検証結果。

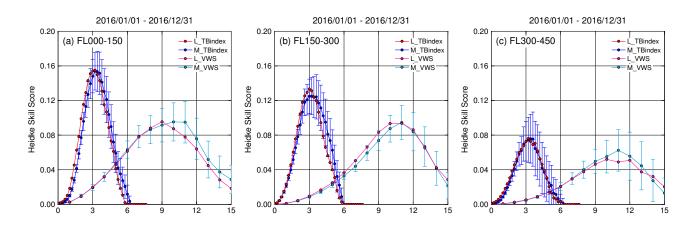


図 4.11.5 図 4.11.3 と同じ。ただし、FT=4~9 について、国内航空悪天 GPV (03, 09, 15, 21 UTC 初期値)と LFM 航空悪天 GPV (03, 09, 15, 21 UTC 初期値)を比較した結果。M\_TBindex と M\_VWS はそれぞれ国内航空悪天 GPV の TBindex と VWS の検証結果を、L\_TBindex と L\_VWS はそれぞれ LFM 航空悪天 GPV の TBindex と VWS の検証結果を示す。

Raito)の関係である。体積率が等しい予測を比べた場合には捕捉率が大きいほど乱気流を多く捕らえた良い予測であり、捕捉率が等しい予測を比べた場合には体積率が小さいほど予測領域を限定した良い予測といえるため、図では線が左上にあるほど予測精度が高いといえる。図を見ると、どの高度でもP\_TBindex と M\_TBindexの予測精度は同等であり、VWSの予測精度を上回っている。TBindex が 3.0 の場合の検証結果を比べると、FL000-FL150と FL150-FL300 では M\_TBindex の方が体積率が大きく、それに伴って捕捉率も高くなっている。

図 4.11.3 は TBindex の閾値別に MOD 以上の乱気流を予測した場合の高度別のスキルスコアである。スキルスコアが最大となる TBindex の閾値を比べると、P\_TBindex ではどの高度でも概ね3.0 付近で最大となっており、指数の調整が適切に機能しているのに対し、

M.TBindex では分布がやや高閾値側にシフトしている。これは本検証期間では M.TBindex の予測がやや過大であったことを表している。スキルスコアの最大値を比較すると、FL000-FL150 と FL150-FL300 では M.TBindex の方が P.TBindex よりも最大値がやや大きく、予測精度はやや高い。FL300-FL450 では同等である。各 TBindex のスキルスコアの最大値は VWS よりも概ね大きくなっており、VWS と比べて予測精度が高いことがわかる。

図 4.11.4 と図 4.11.5 に  $M_{-}$ TBindex と  $L_{-}$ TBindex および、 $M_{-}$ VWS と  $L_{-}$ VWS の検証結果を示す。国内航空悪天 GPV、LFM 航空悪天 GPV とも、03, 09, 15, 21 UTC 初期値の  $FT=4\sim9$  の予測値を用いた。捕捉率—体積率の関係(図 4.11.4)を見ると、どの高度でも $M_{-}$ TBindex と  $L_{-}$ TBindex の予測精度は同等であり、VWS の予測精度を上回っていることがわかる。また、

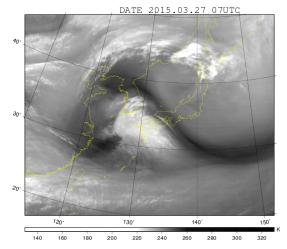


図 4.11.6 2015 年 3 月 27 日 07 UTC の衛星水蒸気画像

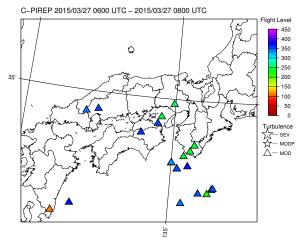


図 4.11.7 2015 年 3 月 27 日 07 UTC の前後 1 時間以内に C-PIREP で通報された MOD 以上の乱気流の実況。 $\triangle$  は MOD の乱気流を、色は高度 (FL) を示す。

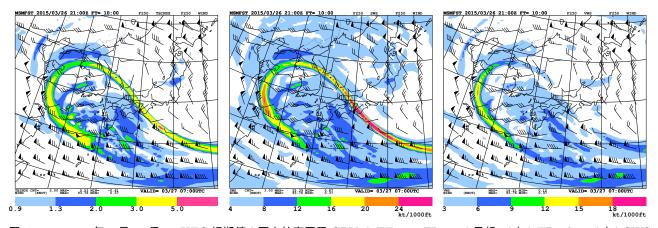


図 4.11.8 2015 年 3 月 26 日 21 UTC 初期値の国内航空悪天 GPV の FT=10, FL250 の予想。(左) TBindex、(中) SWS、(右) VWS。矢羽は風向・風速 [ノット] を表す。

TBindex が 3.0 の場合の検証結果を比べると、全般に  $M_{-}$ TBindex の方がやや体積率が大きく、それに伴って 捕捉率も高くなっている。スキルスコア (図 4.11.5)を 見ると、 $L_{-}$ TBindex ではどの高度でも概ね 3.0 付近で 最大となっており、指数の調整が適切に機能しているのに対し、 $M_{-}$ TBindex では分布がやや高閾値側にシフトしており、本検証期間では  $M_{-}$ TBindex の予測がや や過大であったことを表している。スキルスコアの最大値を比較すると、各 TBindex は VWS よりも大きく なっており、VWS と比べて予測精度が高いことがわ かる。

# 4.11.4 事例検証

ここでは国内航空悪天 GPV の TBindex による乱気流の予測事例を示す。初めに VWS では予測が不十分な乱気流について、TBindex では予測できている典型的な事例として、(1) 斜方シアーが強く予想された領域での乱気流事例(2015年3月27日)と(2) 中層雲底乱気流の事例(2015年6月9日)を、続いて TBindexでの予測が不十分だった事例として、(3) 上層雲先端

付近での乱気流事例 (2017年4月16日)を示す。これらの事例は前節で述べた統計検証の期間とは異なるが、いずれも統計検証に用いたものと同じバージョンの MSM と国内航空悪天 GPV による予測事例である。

(1) 斜方シアーが強く予測された領域での乱気流事例 2015 年 3 月 27 日には SWS が強く予測された領域で乱気流が多数発生した。図 4.11.6 は 27 日 07 UTC の衛星水蒸気画像である。東シナ海には低気圧があり、前面の雲域が九州に掛かっている。広域雲解析情報図 (TSAS1) によれば、この雲域の雲頂は FL320-FL330付近であった(図略)。衛星画像ではこの雲域の先の中国地方から近畿地方で暗域が明瞭である。図 4.11.7 は、27 日 07 UTC の前後 1 時間以内に C-PIREP で報じられた MOD 以上の乱気流実況である。暗域にあたる中国・四国・近畿地方の上空 FL350-FL370付近および、FL210-FL270付近では MOD の乱気流が多数通報されており、報じられた天候状態は主に晴天 (CLR)であった。図 4.11.8 は、26 日 21 UTC 初期値の FT=10, FL250 の国内航空悪天 GPV の予測図である。乱気流

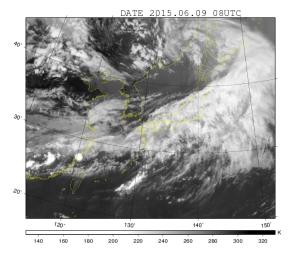


図 4.11.9 2015 年 6 月 9 日 08 UTC の衛星赤外画像

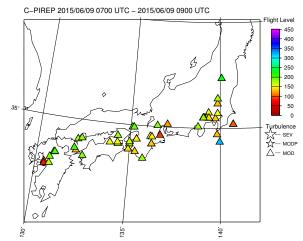


図 4.11.10 2015 年 6 月 9 日 08 UTC の前後 1 時間以内に C-PIREP で通報された MOD 以上の乱気流の実況。 $\triangle$  は MOD の乱気流を、色は高度 (FL) を示す。

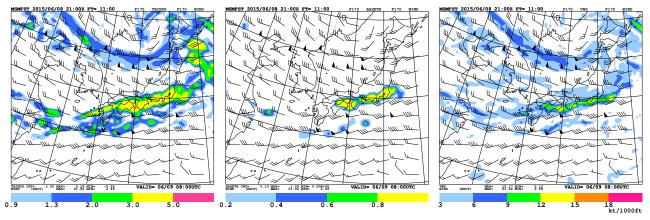


図 4.11.11 2015 年 6 月 8 日 21 UTC 初期値の国内航空悪天 GPV の FT=11, FL170 の予想。(左) TBindex、(中) BASETB、(右) VWS。矢羽は風向・風速 [ノット] を表す。

発生領域の VWS の予測値は  $3\sim6~kt/1000ft$  であり、 国内航空悪天 GPV で晴天乱気流を予測する目安(気象庁予報部 2017)である 12~kt/1000ft と比べるとかなり小さい。一方 SWS では大きな値が予測されており、これによって TBindex では MOD の乱気流を予測する目安である 3.0~e越える値が予測できていた。

#### (2) 中層雲底乱気流の事例

2015年6月9日には西日本から東日本の広い範囲で中層雲底乱気流が多数発生した。図4.11.9は9日08UTCの衛星赤外画像である。9日06UTCの地上天気図(図略)によれば、この時三陸沖に低気圧があり、寒冷前線が関東から九州の南岸にのびていた。衛星画像より、西日本から東日本は前線に伴う雲に広く覆われていることがわかる。図4.11.10は、9日08UTCの前後1時間以内にC-PIREPで報じられたMOD以上の乱気流実況である。前線の北側に当たる西日本から東日本の陸上のFL100-FL220付近でMODの乱気流が多数通報されており、報じられた天候状態は雲底(BSE)が多かった。図4.11.11は、8日21UTC初期値のFT=11、

FL170の国内航空悪天 GPV の予測図である。VWS の予測を見ると、前線近傍に当たる関東から東海、四国の沿岸部で最大で 12 kt/1000ft 程度の予測がされているが、乱気流が発生した陸上では  $3{\sim}6 \text{ kt}/1000\text{ft}$  程度しか予測されていない。一方、前線の北側の領域では BASETB の値が大きく予測されており、これによって TBindex では MOD の乱気流を予測する目安である 3.0を越える値が予測できていた。

#### (3) 上層雲先端付近での乱気流事例

2017 年 4 月 16 日には西日本周辺で乱気流が多数発生した。図 4.11.12 は 16 日 05 UTC の衛星赤外画像である。九州の西の海上には CB クラスターがあり、海上を中心に多数の雷が検知されていた(図略)。この CB クラスターから北東方向には上層雲が吹き出しており九州南部に掛かっている。一方、この時の毎時大気解析(図略)によれば、中国東北区から日本海中部、東海地方にはジェット気流が流れており、関東の南にあるトラフに向かってジェットの減速場となっている。衛星画像ではこのジェットに伴う雲域が中国東北区から日本海

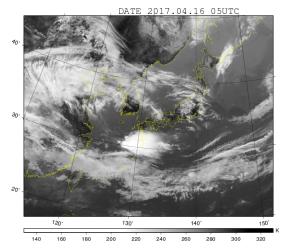


図 4.11.12 2017 年 4 月 16 日 05 UTC の衛星赤外画像

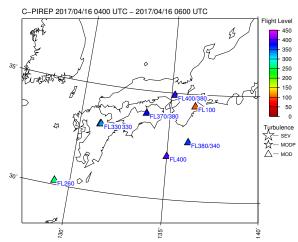


図 4.11.13 2017 年 4 月 16 日 05 UTC の前後 1 時間以内に C-PIREP で通報された MOD 以上の乱気流の実況。△ は MOD の乱気流を、色は高度 (FL) を示す。

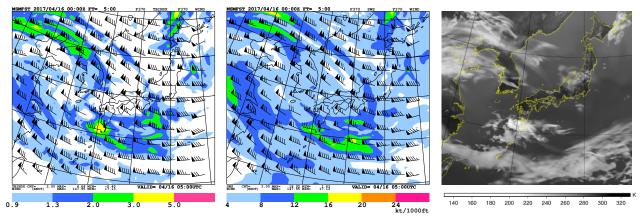


図 4.11.14 2017 年 4 月 16 日 00 UTC 初期値の国内航空悪天 GPV の FT=5, FL370 の予想。(左) TBindex、(中) SWS、(右) 同初期値と FT での MSM からの予想衛星赤外画像。矢羽は風向・風速 [ノット] を表す。

中部に見られるほか、朝鮮半島から山陰にもトランス バースバンド状の雲域が見られる。このトランスバー スバンド状の雲域はジェット気流から分流した流れの 存在を示唆している。図 4.11.13 は、16 日 05 UTC の 前後1時間以内に C-PIREP で報じられた MOD 以上 の乱気流実況である。CB クラスターからの吹き出し と、朝鮮半島から山陰に向かうジェットから分流した流 れの中間に当たる四国とその周辺の FL330-FL400 で MOD の乱気流が多数通報された。図 4.11.14 は、16 日 00 UTC 初期値の FT=5, FL370 の国内航空悪天 GPV の予測図と MSM による予想衛星赤外画像である。CB クラスターの中心付近では、CONV が強いこと(図略) により TBindex が3.0 以上の領域が予想されているが、 MOD の乱気流が発生した四国とその周辺では小さな 値しか予想されていない。これは晴天乱気流を予測す る主な指数である SWS の予測が不十分であったこと が一つの要因である。

国内航空悪天 GPV で SWS の予測が不十分であったのは、MSM の予測が不十分であったためと考えられる。MSM からの予想衛星画像 (大和田 2006, 2013) を

見ると、CB クラスターの予想はできているものの、そこからの吹き出しの予測は実況と比べて不十分である。また、朝鮮半島から山陰に向かう雲域も不明瞭である。実況の衛星画像より、本事例の乱気流は、CB クラスターからの吹き出しの南西風と、山陰に向かう流れの北西風によって生じた強い鉛直シアーによって発生したものと考えられる。実際、0500 UTC の鹿児島のウィンドプロファイラでは FL380 で 80 ノットの南西風が観測されていた(図略)が、FL370 の予測ではこの南西風を予測できていなかった。また、0500 UTC 頃の航空機自動観測では、中国地方の上空 FL380 付近で 115 ノットの北西風が観測されていた(図略)が、予測では 100 ノット程度であり実況よりも弱かった。このように、実況と比べて MSM の予測が不十分であったため、TBindex も不十分になったと考えられる。

本事例のように、数値予報モデルでは CB クラスターからの吹き出しや、ジェット気流から分流した流れなど、上層雲先端付近の表現は不十分である場合が多い。このような場合、VWS や SWS の予測が不十分になり、TBindex の予測も不十分になる。

#### 4.11.5 利用上の留意点と今後の展望

TBindex は様々な要因で発生する乱気流を総合的に予測する指数であるが、全ての乱気流を予測対象としているわけではない。対象としている乱気流は、発生要因に関する調査・研究が進み、予測手法がある程度確立された要因による乱気流のみである。また、後方乱気流や衝撃波失速など、気象以外の要因による揺れも予測対象としていない。

TRAV はトランスバースバンドに伴う乱気流を対象とした指数であるが、トランスバースバンドの発生メカニズムに関する研究は世界的に見てもまだ十分に行われていない (Knox et al. 2010) ことや、数値予報モデルの上層雲の表現が不十分であることなどから、TBindexへの寄与は大きくない (表 4.11.4)。現状ではトランスバースバンドに伴う乱気流の予測精度は不十分であるといえる。また、当然ではあるが、第 4.11.4 項 (3) で述べたような、数値予報モデルの表現が不十分な場合にも TBindex の予測は不十分になる。他のガイダンスと同様に、TBindex を利用する場合にも数値予報モデルの特性を把握し、実況と予測を比較しながら利用することが重要である。

TBindex は係数固定型のガイダンスであり、数値予報モデルの更新時には予測精度が低下する可能性がある。数値予報モデルの更新時には予測精度への影響を評価するとともに、精度の低下が見込まれる場合には係数を再作成するなどの対処が必要になる。また精度低下がない場合であっても、係数を適切な状態に保つために、適宜係数を更新する必要がある。

今後の展望としては、乱気流の発生要因に関する調査を進め、現在のTBindexでは予測が不十分な乱気流の予測精度を向上させたいと考えている。また、新たな予測手法を導入するなど、手法自体の改善にも努めていきたい。

#### 付録 4.11.A 空域予報に関連する用語の解説

ここでは第 4.11 節 ~ 第 4.13 節で共通で利用される 空域予報に関する用語を解説する。

#### **ARS**

特別機上観測 (Special Air-Report)。航空気象官署がパイロット等から収集した航空機気象観測報告を気象庁予報部宛てに行う通報。並 (MOD) 以上の強度の乱気流や着氷のほか、雷電、火山灰、天候状態などが報じられる。

#### C-PIREP

Common-PIREP。従来は各航空会社内でのみ利用されていた PIREP(カンパニー PIREP)を、国土交通省航空局のシステムで集約して形式を統一したPIREP。乱気流、着氷、天候状態などが報じられる。乱気流の強さは、静穏 (SMTH)、弱 $^-$  (LGTM)、弱(LGT)、弱 $^+$  (LGTP)、並 (MOD)、強 (SEV) の

6 段階、着氷の強さは、弱 (LGT)、並 (MOD)、強 (SEV) の 3 段階で報じられる。天候状態は、晴天 (CLR)、雲中 (INC)、雲頂 (TOP)、雲底 (BSE) などが報じられる。MOD や SEV の乱気流だけではなく、 $SMTH \sim LGTP$  の乱気流も数多く通報されることが特徴。

#### FIR

飛行情報区 (Flight Information Region)。国際民間 航空機関 (ICAO) により制定された航空機の運航に 必要な各種情報の提供または捜索救難活動が行われ る空域。日本の担当空域は福岡 FIR。

#### フライトレベル

標準大気の気圧高度 [ft] を 100 で割った高度。 「FL100」など "FL" に続けて 100 ft 単位の数値で表される。 $FL100=10{,}000$  ft、1 ft  $\simeq 0.3048$  m。

#### 国内悪天予想図 (FBJP)

乱気流、着氷、雷電など航空機の運航に影響を及ぼす悪天域のほか、ジェット気流、熱帯低気圧、地上の高・低気圧、前線、FL050 と FL100 の 0 °C の等温線などの予想を図示したプロダクト。概ね 5 時間30 分先の予想図を一日 4 回 6 時間ごとに本庁空域予報班が作成している。

#### 国内悪天 12 時間予想図 (FBJP112-412)

国内航空悪天 GPV の予測値を元に抽出したジェット 気流、晴天乱気流、積乱雲域、地上気圧、前線、着 氷域、雲などに関する予想図を 4 面図に表示したプ ロダクト。初期値時刻から 12 時間後の予想図を 3 時 間ごとに 1 日 8 回自動作成している。

#### 航空悪天 GPV

空域予報作業を支援するための格子点資料。風、気温、湿度などの一般的な気象要素のほか、乱気流、着氷、積乱雲量・雲頂高度、圏界面等の空域予報作業に必要な予測値を等フライトレベル面で算出している。

#### 狭域悪天予想図 (FBTT, FBGG, FBBB)

国内有数の航空交通量を持つ東京、中部、関西の各進入管制区における航空機の運航支援を目的とし、進入管制区及びその周辺領域における悪天の予想を、LFM 航空悪天 GPV を元に自動算出し、図にまとめて表示するプロダクト。3時間ごとに一日8回自動で作成している。

#### **PIREP**

操縦士報告 (Pilot Report)。パイロットが飛行中に 乱気流や着氷などの現象に遭遇した場合に管制機関 や航空会社、気象官署等に行う通報。または、管制 機関等がパイロットから入手した乱気流や着氷など の情報に基づいて行う通報。

#### シグメット (SIGMET) 情報

FIR を対象に発表される空域気象情報。航空機の運航の安全に影響する航空路上の現象(乱気流、着氷、雷電、台風、火山の噴煙、放射性物質の状況など)が発現または予想された場合に発表される。日本では気象庁本庁が福岡 FIR を対象に発表している。

# 付録 4.11.B TBindex の説明変数として利用する各種指数の算出方法

ここでは TB index の算出に用いられている説明変数 と、乱気流予測に用いられている代表的な指数の算出 方法を述べる。本項で用いる主な変数は以下の通りで ある。水平座標 x,y、鉛直座標 z、風の水平成分 u,v、風の鉛直成分 w、風速 V、気温 T、相対湿度 RH。このほかの変数は各変数の説明の中で述べる。

#### (1) 鉛直シアー (VWS)

鉛直シアー (VWS) は水平風の鉛直微分の大きさで、 以下の式で定義される。

$$VWS = \sqrt{\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2}$$
 (4.11.2)

晴天乱気流の主な要因である KH 波は VWS が大きいか静的安定度が弱い場合に発生する。KH 波による鉛直突風の強さは VWS の大きさに比例して強くなるため、VWS が大きいほど強い乱気流が発生しやすくなる。また VWS は、晴天乱気流だけではなく、雲頂、雲中、雲底 (Lane and Sharman 2008; Trier and Sharman 2009; Kudo 2013) で発生する乱気流の一つの要因でもある。このようなことから VWS は乱気流を予測するための指標として優れており、単独の指標としての予測精度は高い (工藤 2008, 2010)。

VWSの計算は鉛直差分によって行うため、差分間隔の取り方によって値が変化することに注意しなければならない。国内航空悪天 GPV では以下のように VWSを算出することで、高度によってモデル面の鉛直間隔が異なる影響を緩和している (工藤 2007)。

- 1. MSM モデル面から、水平 5 km、鉛直 1,000 ft 間隔のフライトレベル面に、風と高度を内挿する
- 2. 2,000 ft 間隔の各フライトレベル面の上下 1,000 ft の風と高度から VWS を算出する
- 3. 国内航空悪天 GPV (水平 40 km 間隔)の各格子 点の近傍 5 × 5 格子の VWS の平均をその格子の VWS とする
- 4. VWS の単位を kt/1000ft に変換する

国内航空悪天 GPV 以外の航空悪天 GPV でも同様の方法で VWS を算出しているが、全球航空悪天 GPV と北太平洋航空悪天 GPV では手順 3 の平均操作は行っていない。算出している VWS の単位は kt/1000ft であるが、これは標準大気の気圧高度 [ft] で高度差が 1,000ft の場合のシアーに相当する値ではなく、モデルの高度

差 [m] から算出した VWS を単位変換した値である。

#### (2) 水平シアー (HWS)

水平シアー (HWS) は水平風の水平微分の大きさで、 以下の式により算出している。

$$HWS = \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2}$$
(4.11.3)

水平方向には 2 次元の自由度があるため微分を取る方向は様々考えられるが、航空悪天 GPV では上式により求めている。VWS と同様に水平差分の取り方によって値が変わることに注意が必要である。航空悪天 GPV では水平方向に概ね  $80~{\rm km}$  間隔(格子点を中心に  $\pm 40~{\rm km}$  間隔)で HWS(単位は  ${\rm s}^{-1}$ )を算出している。

#### (3) 斜方シアー (SWS)

VWS が鉛直方向のシアー、HWS が水平方向のシアーであるのに対して、斜方シアー (SWS) は斜め方向のシアー(単位は kt/1000ft) である。SWS は以下の式で算出している。

$$SWS = \sqrt{SWS_x^2 + SWS_y^2}$$
 (4.11.4)

ここで  $\mathrm{SWS}_x$ と  $\mathrm{SWS}_y$  は  $\mathrm{SWS}$  の x,y 方向成分で、例えば  $\mathrm{SWS}_x$  は図 4.11.15 で示すように x-z 平面上の 2 つの斜め方向の水平風速のシアーの大きさ  $\mathrm{SWS}_{x1}$ と  $\mathrm{SWS}_{x2}$  の平均として算出している。 $\mathrm{SWS}_{x1}$  を図の変数を用いて書くと

$$SWS_{x1} = \frac{\sqrt{(u_2 - u_1)^2 + (v_2 - v_1)^2}}{z_2 - z_1}$$
(4.11.5)

となる。 $SWS_{x1}$ ,  $SWS_{x2}$  の分子は斜め方向の風速差であるのに対し、分母は斜め方向の距離の差ではなく、鉛直方向の差として定義している。このような方法でシアーを計算することに物理的な意味はないが、以下の理由から数値予報モデルを利用するという観点では意味があると考える。

KH 波に伴う乱気流は風のシアーをエネルギー源として生成される。通常の大気では HWS は VWS と比

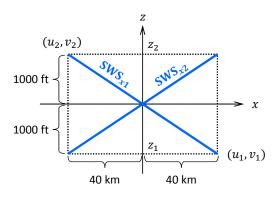


図 4.11.15 斜方シアーの算出

べて非常に小さいため、純粋な HWS だけで航空機に 影響を与えるような、ある程度大きなエネルギーを持っ た KH 波を発生させるとは考えにくい。実際、HWS に よる乱気流予報の精度は VWS と比べると非常に悪い (例えば工藤 (2010)の図 3.5.2)。 しかしながら、数値 予報モデル上で HWS が相対的に大きくて VWS は小 さい(といっても大きさとしては VWS の方が大きい) 領域で乱気流が発生することがある。このことは、純 粋な HWS によって乱気流が発生したと考えるよりは、 HWS が大きい領域で局所的に VWS が大きくなって 乱気流が発生したと解釈する方が自然であると考えて いる。なぜならば、HWS が大きく VWS が小さい場合 に、何らかのきっかけで上下の流れに水平方向のずれ が生じると、大きな HWS は大きな VWS に変換され うるからである。つまり、HWS が大きい状態はHWS が小さい状態と比べて VWS が局所的に大きくなる可 能性が高く、乱気流が発生する可能性も高いと考えら れる。SWS は VWS が大きくなる可能性を表現した指 数であり、上下の流れにずれが生じた場合の VWS の 大きさを表している。SWS は VWS と比べて乱気流の 予測精度はやや高い(例えば工藤(2010)の図3.5.2)。

#### (4) 風下山岳波指数 (MTW1)

風下に伝播した山岳波に伴う乱気流には、山頂高度付近に安定層がある場合に風下波に伴って発生する下層乱流域 (Lower Turbulent Zone, LTZ) での乱気流 (Lester and Fingerhut 1974) と、山岳波の波によって山頂高度付近の安定層内で KH 不安定が強まって発生する乱気流 (Scorer 1978) の 2 つのタイプがある。このうち、後者については KH 不安定を対象とした指数によってある程度表現できるものと考え、ここでは LTZ での乱気流を予測する指数を MTW1 とする。

LTZ を予測する上では以下の3点がポイントとなる。

- 山岳波が発生するために、山脈に直交する山頂付 近の風がある程度強い
- 地表面と山頂高度付近の安定層までの間の成層状態は中立に近い
- 上記の中立に近い成層中で、ある程度風速が強い

これらを踏まえて以下の手順で MTW1 を算出する。

- 1. まず、地表面から最初の安定層までの間で相当温 位の鉛直傾度が 3 K/km 以下で、風速が 20 ノッ ト以上である格子点を探す
- 2. 次に、その格子点をスタート地点として、等フライトレベル面内で最大で 4 時間分の後方流跡線解析を行い、上記 1 の条件を満たす領域の外に出ることなしに山頂(ここではモデル標高が  $200~{\rm m}$  以上の格子点とする)に達した場合、到達点での地表面高度  $z_s$  の傾き  $-\nabla z_s/|\nabla z_s|$  と到達点での地表面付近の風速ベクトル  $v_s[{\rm m/s}]$  の内積を  $10~{\rm cm}$

た値を、スタート地点における MTW1 とする。式で書くと次のようになる。

$$MTW1 = \frac{-\nabla z_s}{|\nabla z_s|} \cdot \frac{\boldsymbol{v}_s}{10}$$
 (4.11.6)

ただし、MTW1 が 1.5 以下の場合、すなわち山脈 に直交する風が弱い場合には山岳波は発生しない ものとし、MTW1=0 とする。

このようにして全格子で MTW1 を算出した後に周囲の格子で平滑化する。

#### (5) 鉛直伝播山岳波指数 (MTW2)

鉛直方向に伝播した山岳波は、臨界高度(内部重力波の位相速度と周囲の風速が等しくなる高度)やその付近での温位の逆転、または強い逆シアー等により波が砕波することで乱気流を発生させる(Gerbier and Berenger 1961; Shen and Lin 1999; Lane and Sharman 2008)。この内、強い逆シアー等により波が砕波することで発生する乱気流は KH 不安定を対象とした指数である程度表現されるものと考え、ここでは臨界高度付近での山岳波の砕波による乱気流を予測することを考える。

まず山岳波の発生条件として、山頂付近での風速がある程度強いことは必要である。そしてその風速が強いほど乱気流の発生確率や強度も大きくなりやすいと考えられる。また山頂付近の風速が一定であれば山岳波の位相速度は0であるため、山岳波の臨界高度は周囲の風速が0になる高度である。これらの状況を表現する指数として、次のような指数を MTW2 とする。

モデル上の標高が  $200~\mathrm{m}$  以上である格子において、地表面付近の風速の大きさを  $V_s$  [ J ット]、その上方のある高度 z での風速の大きさを  $V_z$  [ J ット] とする。このとき、 $V_s > V_z$  かつ  $V_s > 30$  J ットの場合、

$$MTW2 = \frac{V_s - V_z}{10} \tag{4.11.7}$$

を高度 z における MTW2 とする。上記の条件が満たされない場合は MTW2 = 0 とする。MTW1 と同様に、全格子で MTW2 の値を求めた後に平滑化する。MTW2 は山頂付近の風速が強いか、その上方で風速が0 に近いほど値が大きくなる指数である。鉛直方向に伝播した山岳波は、理論的には山頂付近の風向に平行な風の成分が0 になったときに砕波するが、指数としては山頂付近の風向に平行な成分が0 に近いとするよりも、単に風速が0 に近いとする方が乱気流の予測精度が良かったため、(4.11.7) 式を用いることとした。

# (6) 中層雲底乱気流指数 (BASETB)

前線面上に広がる中層雲の雲底付近では、中層雲からの固体降水の昇華冷却によって雲底下に絶対不安定な成層が形成され、ベナール対流が発生することで乱気流(中層雲底乱気流)が生じる(Kudo 2013; Kudo et al. 2015)。中層雲底下での乱気流を予測する上では以下の3点がポイントとなる。

- 中層雲から地上に達しない固体降水が十分にある こと(雲底気温が0°C以下であることが必要条件)
- 中層雲の雲底下が十分に乾燥していること
- 固体降水が昇華することで雲底下に絶対不安定層 が形成されること

この中で、特に絶対不安定が形成されるか否かが重要であり、絶対不安定な成層が形成された場合に中層雲底乱気流が発生しうる。これらを踏まえて、以下の手順でBASETBを算出する。

- 1. 予報対象時刻の前 1 時間の降水量が 0.5 mm/h 未満のカラム(鉛直方向の格子)に対して、最上層の格子から下に向かって相対湿度 RH [%] が設定した閾値  $c_{md1}$  [%] 以上である格子(以下では飽和格子と書く)を探す。 $c_{md1}$  はモデルに応じて決めるパラメータで 100 に近い値を設定する。
- $2. \ c_{md1}$  を上回った分の相対湿度に相当する混合比 [
  m kg/kg] を求める。
- 3. さらに下に向かって同様に飽和格子を探し、 $RH < c_{md1}$  となる格子 (以下では未飽和格子と書く)が 出現するか、気温が  $0\,^{\circ}$ C 以上になるまで閾値を上回った分の混合比を積算する。このようにして積 算された混合比  $R_{ma}$  [kg/kg] が地上に達しない固体降水の総量であると考える。
- 4. 飽和格子の下方に未飽和格子がある場合、未飽和格子の混合比を  $R_{m1}$  [kg/kg]、相対湿度が  $RH+c_{md2}(c_{md1}-RH)$  [%] となるときの混合比を  $R_{m2}$  [kg/kg] とする。ここで  $c_{md2}$  は乾燥域で降水を蒸発させる割合で、モデルに応じて決める  $0\sim1$  のパラメータ (0.5 程度を使用)である。
- $5.~R_{m2}$ と $R_{m1}$ の差を蒸発分の混合比 $R_m$  [kg/kg]とする。ただし $R_{ma}$ を上限とする。

$$R_m = \min(R_{m2} - R_{m1}, R_{ma}) \tag{4.11.8}$$

 $R_m$  は  $R_{ma}$  からの蒸発分であるから、 $R_{ma}$  から $R_m$  を差し引く。

- 6. 固体降水の昇華による雲底下での気温の低下量  $\Delta T$   $[{
  m K}]$  を  $\Delta T=L\,R_m/C_p$  により求める。ここで L は氷の昇華潜熱  $[{
  m J}\,{
  m Kg}^{-1}]$ 、 $C_p$  は乾燥空気の定圧比熱  $[{
  m J}\,{
  m K}^{-1}{
  m kg}^{-1}]$  である。
- 7. 昇華冷却による気温低下を考慮した場合の気温と 1 層下の層の気温 ( 次の手順で示す気温低下を考慮する前の気温 ) との間で気温減率  $\Gamma$  [K/km] を求める。
- 8. この $\Gamma$ が乾燥断熱減率 $\Gamma_d$   $[{
  m K/km}]$  よりも大きい場合、すなわち絶対不安定である場合、

$$BASETB = \Gamma - \Gamma_d \tag{4.11.9}$$

とする。ただしBASETBの最小値は0、最大値は 1 とする。 9. さらに下に向かって手順  $4\sim8$  を繰り返し、各高度での BASETB の値を求める。最下層に到達する前に再び 0 °C 以下で飽和格子が現れた場合には手順 2 へ戻る。

BASETB では上限を 1、すなわち、 $\Gamma - \Gamma_d$  の上限を  $1.0~{\rm K/km}$  とした。これは、現実の大気中では固体降水 は徐々に昇華しながら絶対不安定な成層を形成し、その後は比較的速やかに不安定が解消されると考えられるため、極端に大きな気温減率になることはないと考えたためである。

#### (7) トランスバースバンド指数 (TRAV)

ジェット気流に伴う雲バンドの中で発生するトランス バースバンドは、乱気流発生の重要な指標の一つである (気象衛星センター 2002)。大野・三浦 (1982) は、トランスバースバンドは圏界面下で励起された KH 波であり、発生条件は「圏界面直下の低安定度層厚が  $1~\rm km$  を超え、その中で最小リチャードソン数がおよそ  $0.25~\rm k$  り小さく、最大風速が  $50~\rm m/s$  をこえる」と述べている。ただしこれは  $1~\rm km$  次元方程式の数値解であり、  $3~\rm km$  元の現象であるトランスバースバンドの発生条件と一致しているとは限らない。現実の大気中では、トランスバースバンドは、ジェット気流南側の弱安定層の上部に位置する圏界面付近で、風速がおよそ  $100~\rm lm$  (約  $50~\rm lm/s$ ) 以上で VWS がやや大きい領域で発生していると考えられる。これを表現する指標をトランスバースバンド指数とし、以下の手順で算出する。

- 1. あるカラムで、相当温位の鉛直変化率が  $2.0~{
  m K/km}$  以下の層が  $4{,}000~{
  m ft}$  以上続いている場合、その領域の上端高度の鉛直層の番号を  $k_1$  とする。
- $2. \ k_1$  番目の層から上に相当温位の鉛直変化率が  $2.0 \ {
  m K/km}$  よりも大きい層が続いている場合、その中のあるフライトレベル面(層番号  $k_2$ )におけるトランスバースバンド指数  ${
  m TRAV}$  を次式で定義する。

$$TRAV(k_2) = VWS(k_2) \cdot F(k_1, k_2) \cdot S(k_2)$$
(4.11.10)

$$F(k_1, k_2) = 2^{[1.5 - {FL(k_2) - FL(k_1)}/20]}$$
(4.11.11)

$$S(k_2) = \frac{1}{1 + \exp[-20\{V(k_2) - 90\}/90]}$$
 (4.11.12)

ここで VWS  $(k_2)$  は第  $k_2$  層の VWS [kt/1000ft]、FL(k) は第 k 層のフライトレベル、 $V(k_2)$  は第  $k_2$  層の風速 [ ノット] である。 $F(k_1,k_2)$  は、 $k_2$  が弱安定層の上端 高度である  $k_1$  に近いほど大きくなる関数で、 $k_1$  から 1,000 ft 離れる毎に、 $2,\sqrt{2},1$  というように、 $1/\sqrt{2}$  倍に なっていく。これは弱安定層の上端高度がトランスバースバンドの雲頂に相当し、その雲頂に近いほど乱気流が発生しやすいと考えられることを表している。 $S(k_2)$  は風速に応じて  $0\sim1$  に変化するロジスティック関数で、風速が 90 ノットのときに 0.5 となる(図 4.11.16)。こ

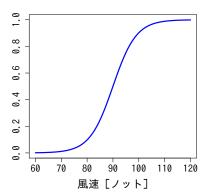


図 4.11.16 トランスバースバンド指数の算出に用いられる ロジスティック関数

れは風速が弱いと0で100 ノットを越えると連続的に1に近づく様子を表している。TRAV は風速がおよそ100 ノット以上の弱安定層の上端付近でVWS が強いほど値が大きくなる指数である。

#### (8) 対流雲中乱気流指数 (CONV)

対流雲の雲中では強い上昇・下降流によって鉛直方向の突風が発生することにより乱気流が発生する。また対流雲中や雲頂に VWS の大きな層がある場合には、上昇流によって VWS が強化されて KH 波が発生し、乱気流が発生する (Lane and Sharman 2008)。対流雲中乱気流指数 (CONV) はこのような乱気流を予測することを目的とし、以下の式で求める。

$$CONV(k_1) = W_{mean}(k_1) \cdot VWS(k_2) \qquad (4.11.13)$$

ここで  $W_{\rm mean}$  は水平  $80~{\rm km}$  格子で平均した鉛直風速  $[{\rm m/s}]$ 、  ${\rm VWS}(k_2)$  は第  $k_2$  層の鉛直シアー  $[{\rm kt/1000ft}]$ 、 $k_1$  は  $W_{\rm mean}>0$  の格子点の鉛直層の番号、 $k_2$  は  $k_1$  から上方に  $4{,}000~{\rm ft}$  以内で  $W_{\rm mean}>0$  領域が継続している層の中で、  ${\rm VWS}$  が最大となる鉛直層の番号を示す。  $W_{\rm mean}(k_2)\leq 0$  の場合は  ${\rm CONV}(k_1)=0$  とする。山岳波などの対流雲以外での強い上昇流の影響を除去するために、対象とする地点での前 1 時間降水量が  $0.1~{\rm mm/h}$ 未満である場合はその上空の全ての層で  ${\rm CONV}=0$  とする。

#### (9) VHS

ジェット気流の加速場や減速場では乱気流が発生しやすいことを考慮した指標として VHS (Vws-Hws-Speed)を以下の式で定義する。

$$VHS = VWS \times HWS_{\parallel} \times V \times 10^{5}$$
 (4.11.14)

ここで VWS は鉛直シアー (単位は  ${
m s}^{-1}$ )、 ${
m HWS}_{\parallel}$ は流れに沿った方向の水平シアー  $[{
m s}^{-1}]$ 、V は風速  $[{
m m/s}]$ 、 $10^5$ は VHS のオーダーを 1 程度にするための係数である。 VHS は鉛直シアーが強いか、流れに沿って風速が変化する場合や流れが曲率を持つか、風速が強い場合に大きな値を持つ指数である。

# (10) リチャードソン数 (RI)

リチャードソン数 (RI) は次式で定義される。

$$RI = \frac{N^2}{VWS^2}$$
 (4.11.15)

ここで N はブラント・バイサラ振動数  $[\mathbf{s}^{-1}]$ 、VWS の 単位は  $\mathbf{s}^{-1}$  である。リチャードソン数が 1/4 以下になると KH 波が発生する。

#### (11) 気温の水平傾度 (GRADT)

気温の水平傾度 (GRADT) は次式で定義される。

GRADT = 
$$\sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2}$$
 (4.11.16)

温度風の関係から、等圧面(等フライトレベル面)での気温の水平傾度は地衡風の鉛直シアーの大きさに比例することから、気温の水平傾度は晴天乱気流を予測する指標として利用される。

#### (12) Ellrod の指数 (TI1, TI2)

Ellrod and Knapp (1992) による指数 TI1, TI2 は次式で定義される。

$$TI1 = VWS \times DEF \tag{4.11.17}$$

$$TI2 = VWS \times (DEF + CVG) \tag{4.11.18}$$

ここで VWS の単位は  $\mathrm{s}^{-1}$ 、DEF は変形場  $[\mathrm{s}^{-1}]$ 、CVG は水平収束  $[\mathrm{s}^{-1}]$  を表す。 $\mathrm{TI1}$  と  $\mathrm{TI2}$  は単独の指標としては乱気流の予測精度が高いといわれていることから世界的には広く利用されている。ただし C-PIREPを用いた日本周辺空域での検証では、VWS と比べて予測精度に大きな違いはない ( 例えば工藤 (2010) の図 3.5.2 )。

#### (13) Dutton の指数 (DI)

Dutton (1980) による指数 DI は次式で定義される。

$$DI = 1.25 \times 10^5 \; HWS_{\perp} + 2.5 \times 10^5 \; VWS^2 + 10.5 \endaligned (4.11.19)$$

ここで  $HWS_{\perp}$  は流れに直交する方向の水平シアーで単位は  $s^{-1}$ 、VWS の単位は  $s^{-1}$  である。 Dutton (1980) は 1976 年の春に北大西洋域で収集した乱気流の実況と当時の数値予報モデルの出力値を利用して、乱気流の発生確率を表す式として (4.11.19) 式を線形重回帰で求めた。このことから分かるように、上式をそのまま乱気流発生確率として他の数値予報モデルに適用することはできないが、DI を HWS と VWS を組み合わせた指数と見た場合には VWS と同程度の予測精度を持つ (例えば工藤 (2010) の図 3.5.2)。

#### 参考文献

Browning, K. A., G. W. Bryant, J. R. Starr, and D. N. Axford, 1973: Air motion within Kelvin-Helmholtz billows determined from simultaneous

- Doppler radar and aircraft measurements. Quart. J. Roy. Meteor. Soc., 99, 608–618.
- Dutton, M. J. O., 1980: Probability forecasts of clearair turbulence based on numerical model output. *Meteorological Magazine*, 109, 293–310.
- Ellrod, G. P. and D. I. Knapp, 1992: An objective clear-air turbulence forecasting technique: Verification and operational use. *Wea. Forecasting*, 7, 150–165.
- Gerbier, N. and M. Berenger, 1961: Experimental studies of lee waves in the French alps. Quart. J. Roy. Meteor. Soc., 87, 13–23.
- 井藤智史, 2016: GSM 乱気流指数の改良. 平成 28 年度 数値予報研修テキスト, 気象庁予報部, 69-76.
- 金井義文, 満満男, 工藤淳, 2015: 下層悪天予想図及び 新しい狭域悪天予想図. 航空気象ノート第 77 号, 気 象庁航空気象管理官.
- 気象庁予報部, 2017: 航空気象予報作業指針 (2017年 8月版).
- 気象衛星センター, 2002: 気象衛星画像の解析と利用 -航空気象編 - . 223 pp.
- Knox, J. A., A. S. Bachmeier, W. M. Carter, J. E. Tarantino, L. C. Paulik, E. N. Wilson, G. S. Bechdol, and M. J. Mays, 2010: Transvere cirrus bands in weather systems: a grand tour of an enduring enigma. Weather, 65(2), 35–41.
- 工藤淳, 2007: 航空悪天 GPV. 平成 19 年度数値予報研修テキスト, 気象庁予報部, 82-86.
- 工藤淳, 2008: 国内航空悪天 GPV. 平成 20 年度数値予報研修テキスト, 気象庁予報部, 92-98.
- 工藤淳, 2010: 乱気流指数の開発. 平成 22 年度数値予報研修テキスト, 気象庁予報部, 95-108.
- 工藤淳, 2012: C-PIREP 10 年の統計. 航空気象ノート 第73号, 気象庁航空気象管理官, 12-19.
- 工藤淳, 2014: 乱気流指数の改良. 平成 26 年度数値予報研修テキスト, 気象庁予報部, 87-99.
- Kudo, A., 2013: The generation of turbulence below midlevel cloud bases: The effect of cooling due to sublimation of snow. J. Appl. Meteor. Climat., 52(4), 819–833.
- Kudo, A., H. Luce, H. Hashiguchi, and R. Wilson, 2015: Convective instability underneath midlevel clouds: Comparisons between numerical simulations and VHF radar observations. J. Appl. Meteor. Climat., 54(11), 2217–2227.
- Lane, T. P. and R. D. Sharman, 2008: Some influences of background flow conditions on the generation of turbulence due to gravity wave breaking above deep convection. *J. Appl. Meteor. Climat.*, 47(11), 2777–2796.

- Lane, T. P., R. D. Sharman, S. B. Trier, R. G. Fovell, and J. K. Williams, 2012: Recent advances in the understanding of near-cloud turbulence. *Bull. Amer. Meteor. Soc.*, 93, 499–515.
- Lester, P. F. and W. A. Fingerhut, 1974: Lower turbulent zones associated with mountain lee waves. J. Appl. Meteor., 13(1), 54–61.
- Lilly, D. K., 1986: Instabilities. Mesoscale Meteorology and Forecasting. P. S. Ray Ed., Amer. Meteor. Soc., 259-271.
- MacCready, P. B., 1964: Standardization of gustiness values from aircraft. J. Appl. Meteor., 3, 439–449.
- MacPherson, J. I. and G. A. Isaac, 1977: Turbulent characteristics of some Canadian cumulus clouds. J. Appl. Meteor., 16, 81–90.
- 三輪剛, 2013: GSM 乱気流指数の開発. 平成 25 年度数値予報研修テキスト, 気象庁予報部, 63-70.
- 中山章, 2010: 飛行機と気象. 成山堂書店, 153 pp.
- 大野久雄, 三浦信男, 1982: 圏界面直下におけるケルビンへルムホルツ波の励起. 天気, **29**, 1235–1241.
- 大和田浩美, 2006: 予想衛星画像. 気象研究ノート, **212**, 105–119
- 大和田浩美, 2013: 予想衛星画像. 数値予報課報告・別冊第 59 号, 気象庁予報部, 76-80.
- Scorer, R. S., 1978: *Environmental aerodynamics*. Ellis Horwood, 488 pp.
- Shapiro, M. A., 1980: Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the stratosphere and troposphere. J. Atmos. Sci., 37, 994–1004.
- Sharman, R. D. and T. P. Lane, 2016: Aviation Turbulence. Springer, 523 pp.
- Shen, B. and Y. Lin, 1999: Effects of critical levels on two-dimensional back-sheared flow over an isolated mountain ridge on an f plane. *J. Atmos. Sci.*, **56(18)**, 3286–3302.
- Trier, S. B. and R. D. Sharman, 2009: Convection-permitting simulations of the environment supporting widespred turbulence within the upper-level outflow of a mesoscale convective system. *Mon. Wea. Rev.*, **137(6)**, 1972–1990.
- Trier, S. B. and R. D. Sharman, 2016: Mechanisms influencing cirrus banding and aviation turbulence near a convectively enhanced upper-level jet stream. *Mon. Wea. Rev.*, **144**, 3003–3027.
- Trier, S. B., R. D. Sharman, R. G. Fovell, and R. G. Frehlich, 2010: Numerical simulation of radial cloud bands within the upper-level outflow of an observed mesoscale convective system. *J. Atmos.* Sci., 67, 2990–2999.

#### 4.12 着氷指数 1

#### 4.12.1 はじめに

着氷とは、空気中の水蒸気や過冷却水滴が物体の表 面に凍りつく現象である。着氷が航空機に与える影響 は大きく、翼に着氷した場合は、揚力の低下、抵抗の 増加、失速速度の増加に繋がり、ピトー管2への着氷 は速度や高度の誤認を招く。エンジンの吸気系統へ着 氷した場合は、混合ガスや空気が流れにくくなるため エンジンの性能が低下し、最悪の場合エンジンが停止 することもある。近年は、機体各部に氷が付着しない ようにするための防氷装置や、付着した氷を取り除く ための除氷装置の発達により、ある程度の着氷は防ぐ ことができるようになったが、気象条件によってはこ れらの装置を使用しても着氷を防ぐことができない場 合がある。そのため、航空機の安全な運航のためには、 事前に着氷域を予測し、その空域を回避することが重 要となる。参考として、FAA (2017) で定義されている 着氷の強さを以下に記す。

- 弱: 長時間の飛行ならば氷の蓄積は問題となるが、 防氷装置、除氷装置を時々作動させれば、氷 の蓄積を防げる状態。
- 並: そのまま氷が蓄積すると短時間でも危険になる可能性があり、防氷装置、除氷装置を使用するかその空域を回避する必要がある状態。
- 強: 防氷装置、除氷装置を使用しても着氷を防ぐ ことができず、直ちにその空域を回避する必 要がある状態。

着氷は主に、氷点下の機体表面に水蒸気が昇華してできる着氷、過冷却水滴が機体に衝突して凍結する着氷、過冷却状態にある雨(着氷性の雨)が機体に衝突して凍結する着氷に大別される3。日本付近で発生する着氷は過冷却水滴が主な原因である(中山 1996)。このため、気象庁の空域予報現業では過冷却水滴による着氷を予測する手法として、-8D法(Godske et al. 1957)や過去の文献や調査により決定した-20~0°Cを目安に着氷が発生すると判別する方法が用いられてきた。-8D法は、雲の中に過冷却水滴が存在するとき、周囲の湿潤空気が氷に対して過飽和、水に対して飽和の状態になると着氷が起こるとの考えに基づいている。気

温を T [ $^{\circ}$ C]、露点温度を  $T_d$  [ $^{\circ}$ C] とすると、霜点温度  $T_f$  [°C] が近似的に  $T_f = -8(T - T_d)$  と表せることか ら、 $T < -8(T - T_d)$  のときは氷面に対して過飽和であ るとし、着氷が発生すると判定する。しかし、Godske et al. (1957) でも述べられている通り、近似の精度は 気温 -13°C 程度までは良いものの、それ以下では悪く なるという問題がある。また、 $-20\sim0$  °C を目安にす る方法に関しても、工藤 (2006) による調査で、−20°C 以下であっても無視できない頻度で着氷が発生してい ることが分かった。このため、工藤 (2008) は、着氷域 を予測するために着氷指数を開発し、MSM の予測値 から算出する国内航空悪天 GPV で 2008 年 8 月から着 氷指数の作成を開始した。その後、LFM の予測値から 算出する LFM 航空悪天 GPV では 2012 年 8 月から、 GSM の予測値から算出される北太平洋航空悪天 GPV では2017年5月から着氷指数の作成が開始された(井 藤 2017)。空域予報現業では、国内悪天予想図 (FBJP) の作成時に着氷指数を利用しているほか、自動作成さ れる狭域悪天予想図 (FBTT, FBGG, FBBB) にも着氷 指数が利用されている。

本節では、着氷指数について解説する。まず、第4.12.2 項で着氷指数の作成方法について、第4.12.3 項で着氷指数の仕様について述べ、第4.12.4 項で統計検証、第4.12.5 項で事例検証を示す。最後に、第4.12.6 項でまとめと利用上の留意点、第4.12.7 項で課題と今後の展望を述べる。以下、北太平洋航空悪天 GPV の着氷指数を GSM 着氷指数、国内航空悪天 GPV の着氷指数を MSM 着氷指数、LFM 航空悪天 GPV の着氷指数を LFM 着氷指数と呼ぶ。また、空域予報に関連する用語については付録4.11.A にまとめているので、そちらを参照して頂きたい。

#### 4.12.2 作成方法

過冷却水滴による着氷の正確な予測には、気温、相対湿度、水滴のサイズの予測が重要となる。一般に過冷却水滴は -20~0°C で多く存在し、相対湿度が高いと水滴の量が多くなり、氷がより早く機体表面に蓄積するので危険である。また、大きな水滴は小さな水滴よりも航空機周辺の気流の影響を受けにくいので、水滴のサイズが大きいほど機体表面に衝突しやすくなる。しかし、水滴のサイズを正確に予測することは現状では難しいため、着氷指数では気温と湿数の予測と並以上の着氷の観測頻度に基づいた指数を作成することで、並以上の着氷に対する予測を行う。

各航空機観測 (PIREP, ARS, C-PIREP) で報じられた着氷の実況と、直近の初期値の数値予報モデルの気温および湿数の予測値を用いてデータにフィッティングする関数 (当てはめ関数)を求め、この当てはめ関数から着氷指数を計算する。なお、着氷は基本的に雲の中で発生するので、晴天 (Clear) と報じられた通報は誤りであることが多いため除外する。

<sup>1</sup> 井藤 智史

<sup>&</sup>lt;sup>2</sup> 航空機の速度を計測する装置。航空機の進行方向正面からくる空気の圧力(全圧)と、その横を通り過ぎる周囲の空気の圧力(静圧)の差を測定することで速度を求める。静圧からは飛行高度も知ることができる。

 $<sup>^3</sup>$  他にも、氷晶がエンジンの熱により部分的に溶融し、エンジン内部に付着して表面温度を下げ再凍結する氷晶着氷 (ICI:Ice Crystal Icing) と呼ばれる着氷がある。過冷却水滴による着氷は  $-20\sim0$  °C で多く発生し、-30 °C 以下の低温ではほとんど発生しないが、ICI は -30 °C 以下でも発生する場合がある (Mason 2007)。

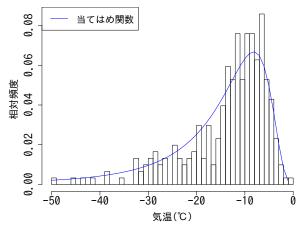


図 4.12.1 2013 年 12 月から 2015 年 11 月の並以上の着氷 頻度と直近の GSM で予測された気温の関係。棒グラフが 着氷の相対頻度、青線が当てはめ関数。

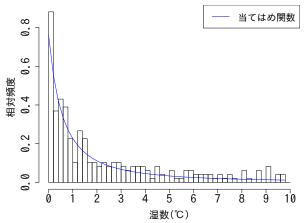


図 4.12.2 図 4.12.1 と同じ。ただし、横軸は GSM で予測された湿数。

まず着氷の実況と、直近の初期値のモデルの気温および湿数の予測値を用いて当てはめ関数を求める。図 4.12.1 は気温と並以上の着氷頻度の関係、図 4.12.2 は湿数と並以上の着氷頻度の関係である。モデルは GSMで、期間は 2013 年 12 月から 2015 年 11 月である。気温が -8 °C 付近と、湿数が 0 °C に近いほど着氷の発生頻度が多くなっていることが分かる。これらの頻度分布から、気温 (T [°C]) に対する頻度を

頻度 = 
$$\frac{-1.6 \times 10^6 \times (T-3)^{-5}}{\exp\{-55.3/(T-3)\} - 1}$$

という関数で、湿数(気温 - 露点温度: $T-T_d$  [ $^{\circ}$ C])に対する頻度を

頻度 = 
$$\frac{0.75}{(T - T_d + 1)^{1.7}}$$

という関数で当てはめる。気温についてはプランク分布を参考に、湿数については湿数の逆数のべき乗で当てはめた。着氷指数はこれら 2 つの式を掛け合わせ、最大値が 100 になるような規格化定数 C を掛けたもの

とする。すなわち、

GSM着氷指数 =

$$C \times \frac{-1.6 \times 10^6 \times (T-3)^{-5}}{\exp\{-55.3/(T-3)\} - 1} \times \frac{0.75}{(T-T_d+1)^{1.7}}$$

とする。ただし、気温が 2 °C より高い場合は着氷が起こることはほとんどないので、気温の予測値が 2 °C より高い場合は着氷指数は 0 とする。気温が 2 °C より高い場合としたのは、気温が 0 °C より少し高い場合でも、機体表面からの蒸発や機体による気流の乱れで気圧が低下し、温度が下がり着氷する場合があるためである。

国内航空悪天 GPV については 2002 年から 2005 年で データを作成し、以下の計算式を用いて作成している。

MSM着氷指数 =

$$C \times \frac{-7 \times 10^8 \times (T-3)^{-5}}{\exp\{-60/(T-3)\} - 1} \times \frac{100}{(T-T_d+1)^{2.5}}$$

なお、LFM 着氷指数も MSM 着氷指数と同じ計算式を用いて作成している。着氷指数を用いて並以上の着氷を予測する場合の閾値は、GSM 着氷指数の場合は10以上、MSM・LFM 着氷指数の場合は15以上とする。これらの値は並以上の着氷が観測された事例を見て、適切な予測となるように決定した。

着氷指数は気温と湿数の当てはめ関数を単純に掛け合わせただけの指数であるが、-8D 法や MSM の雲物理過程で算出された雲水量を用いた方法 (気温が 0 °C 未満で雲水量が大きいほど過冷却水滴が多く、着氷が発生すると考える方法)よりも精度が高いことが確認されている (工藤 2008)。

着氷指数の注意点としては、着氷指数は過冷却水滴 による着氷を予測対象としており、ICI など頻度の低い 着氷は予測対象とはしていないことである。また、着 氷指数は0から100の値を取るが、確率を表すもので はないことにも注意が必要である。これは、MOD 以 上の着氷の相対頻度のみから着氷指数を作成しており、 MOD 未満の実況は含まれていないためである。仮に MOD 未満の実況も含めれば、ロジスティック回帰な どで確率予測を行うことができる。しかし、実況デー タに用いている PIREP 等は、航空路周辺にしか観測 データがなく、自動観測でもないため時間的にも不均 一なデータとなることや、防氷装置によりある程度の 着氷は防ぐことができるため、正確な着氷の確率とは ならない。実際に、ロジスティック回帰による手法も 調査はしたが、着氷指数の予測精度を上回ることはで きなかった。

#### 4.12.3 仕様

モデル毎の着氷指数の仕様を表 4.12.1 に示す。着氷 指数は格子形式で作成されており、北太平洋航空悪天 GPVでは格子間隔 0.5 度の等緯度経度座標で、国内航空

丰	1 -	เก	1	エディ	し伝み	羊ル	七半粉	の仕様
ᅏ	4	١./		+	1,444()	) <del>7 =</del> 7 K	*= **V	(/)1TA=

	北太平洋航空悪天 GPV	国内航空悪天 GPV	LFM 航空悪天 GPV
作成対象モデル	GSM	MSM	LFM
格子間隔	0.5 度	$40~\mathrm{km}$	10 km
格子系	等緯度経度座標	ポーラー	·ステレオ座標
対象領域	$10^{\circ}\text{S}-65^{\circ}\text{N}, 80^{\circ}\text{E}-110^{\circ}\text{W}$	日本及びその周辺	日本周辺
鉛直層	FL010 から FL550 ま	で 2,000 ft 間隔	FL010 から FL450 まで 1,000 ft 間隔
予報対象時間	3 時間毎の毎正時	4	<b></b> 事正時
予報期間	FT=0~36	$FT=0\sim39$	FT=0~9
層別化		なし	
逐次学習の有無		なし	

悪天 GPV では 40 km 格子のポーラーステレオ座標で、 LFM 航空悪天 GPV では 10 km 格子のポーラーステレ 才座標で計算している。鉛直方向には、北太平洋航空 悪天 GPV と国内航空悪天 GPV で FL010 から FL550 まで 2,000 ft 間隔、LFM 航空悪天 GPV で FL010 から FL450 まで 1,000 ft 間隔となっている。予報時間は、 北太平洋航空悪天 GPV で FT=0 から FT=36 までの 3 時間間隔、国内航空悪天 GPV で FT=0 から FT=39 までの 1 時間間隔、LFM 航空悪天 GPV で FT=0 か ら FT=9 までの 1 時間間隔である。過冷却水滴による 着氷は、地域や季節等による特性の変化はないため、 層別化はしていない。また、逐次学習も行っていない。 これは着氷の通報される割合が非常に少なく、全ての 着氷通報を合わせても全通報に対して 0.096%程度しか ないため (工藤 2012)、第 2.2.3 項で述べたように一括 学習の方が適しているためである。なお、MSM 着氷 指数の開発以降、何度か MSM の改良が行われている が、MSM 着氷指数の精度は事例検証では悪化してお らず、2017年2月28日の MSM の改良の際には統計 検証で改善した4ため、作成当初からの計算式を継続 して使っている。

#### 4.12.4 統計検証

着氷指数の統計検証の結果を示す。検証に用いる実況データは、PIREP、ARS、C-PIREPによる着氷の通報を用いる。計算式作成時と同じように、Clearと報じられた通報は除外する。検証期間は 2016 年の 1 年間である。GSM 着氷指数は 2017 年 5 月 25 日に変更された GSM の試験データから作成し、MSM 着氷指数は 2017 年 2 月 28 日に変更された MSM の試験データから作成している。LFM 着氷指数は本期間に予測精度に大きく影響を与えるような LFM の変更は行われなかったため、ルーチンのデータを用いて検証を行った。検証領域は日本周辺  $(20^\circ-45^\circ\text{N}, 120^\circ-150^\circ\text{E})$  で、高度

#### 2016/01/01 - 2016/12/31

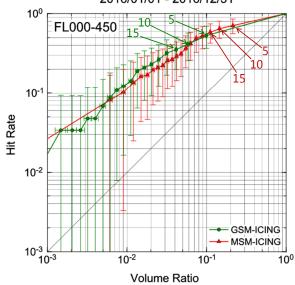


図 4.12.3 2016 年の並以上の着氷に対する着氷指数の閾値別の予測精度。検証領域は日本周辺で、高度は  $FL000\sim450$  である。横軸は体積率、縦軸は捕捉率を示す。 GSM 着氷指数 (GSM-ICING) は緑色、MSM 着氷指数 (MSM-ICING) は赤色で、それぞれの指数の閾値に対応する数字をいくつか図中に示す。それぞれのエラーバーはプートストラップ法による 95%信頼区間を示す。

は  ${
m FL}000{\sim}450$  とした。着氷指数の値は、観測場所から最も近い高度・格子で、観測時刻から最も近い予報時間の予測値を用いた。

図 4.12.3 に、GSM と MSM の着氷指数の閾値を変化させて、それぞれの閾値で並以上の着氷を予測した場合の捕捉率 (Hit Rate) と体積率 (Volume Ratio) を示す (巻末付録 A.3.5, A.3.6)。 GSM は 00, 12UTC 初期値の  $FT=9\sim18$  を用い、MSM は 03, 15UTC 初期値の  $FT=6\sim15$  を用いた。体積率が等しい予測を比べた場合には捕捉率が大きいほど着氷の実況を多く捕らえた良い予測であり、捕捉率が等しい予測を比べた場合には体積率が小さいほど予測領域を限定した良い予測と言えるため、図では線が左上にあるほど予測精度が高い。並以上の着氷の目安となる閾値 (GSM 着氷指数で 10、MSM 着氷指数で 15) 付近で比較すると、精度はほぼ同等となっている。ただし、MSM 着氷指数の

 $<sup>^4</sup>$  通常  $\mathrm{MSM}$  の改良の際には 1 か月程度の試験データしか作成されないため、着氷の観測数が少なく統計検証ができないが、2017 年 2 月 28 日の改良の際には長期間の試験データが作成されたため、統計検証が可能であった。

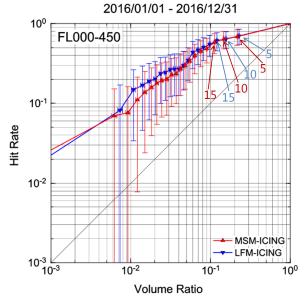


図 4.12.4 図 4.12.3 と同じ。ただし、MSM 着氷指数 (MSM-ICING, 赤色) と LFM 着氷指数 (LFM-ICING, 青色) の 結果。

方が体積率が大きく、そのため捕捉率も大きくなっている。

図 4.12.4 は、MSM 着氷指数と、LFM 着氷指数の検証結果である。ともに 03, 09, 15, 21UTC 初期値で、 $FT=6\sim9$  で検証した。並以上の着氷の目安となる閾値 ( $MSM \cdot LFM$  着氷指数ともに 15) 付近で比較すると、精度はほぼ同等となっている。また、体積率や捕捉率もほぼ同じである。

#### 4.12.5 事例検証

#### (1) 台風と停滞前線による着氷

2017年10月21日から22日にかけて、台風第21号が日本の南を北上し、22日18UTC頃、超大型・強い勢力で静岡県に上陸した。図4.12.5と図4.12.6に22日00UTCの日本域地上天気図と衛星赤外画像を示す。この時間は、日本の南を台風第21号が北上しており、日本付近には停滞前線が解析されている。台風本体の雲域は海上にあるが、停滞前線近傍の雲域により日本付近は広く雲に覆われている。台風からの湿った空気が停滞前線に向かって流れ込んだ影響で、22日は広い範囲で着氷が観測され、この一日で26通の並の着氷が報告された。これは航空機の観測データの保存を開始した2001年6月以降で、1日の観測数としては最多であった。

図 4.12.7 に 22 日 00UTC を対象時刻とする着氷指数の予測を示す。左から GSM 着氷指数、MSM 着氷指数、LFM 着氷指数で、上段から FL330, FL250, FL170 の予測である。図中の四角は対象時刻の前後 90 分以内で高度差 1,000 ft 以内に観測された着氷の実況で、黄色が並の着氷、青色が弱の着氷である。この時刻は、FL170 から FL330 にかけて多くの並の着氷が観測された。FL330 の予測に関しては、九州の南海上の並の

着氷は LFM 着氷指数のみ予測でき、GSM 着氷指数と MSM 着氷指数では予測できなかった。これは LFM の湿数の予測が、GSM や MSM と比べて 0 °C に近かったためである (図略)。この時の気温の予測は概ね-30 °C 程度となっているが、図 4.12.1 で示したように、気温が-30 °C 以下になると着氷の頻度は大きく下がる。このため、着氷の頻度でフィッティングした気温と湿数を掛け合わせる現在の着氷指数の予測手法では、-30 °C 以下で予測値が小さくなる。このような時は、並の閾値を超えるかどうかは湿数の予測に依り、今回の事例のようにモデル間で予測の差が出やすくなる。

FL330 より低い高度に関しては、どのモデルも並の着氷を捕捉できており、領域、高度ともに精度良く予測できていた。着氷が観測された周辺の気温はおおよそ $-30\sim-4$  °C 程度となっており、これらの気温では着氷の発生頻度が高いので、着氷指数で精度良く予測することができる。

#### (2) 冬型の気圧配置での着氷

2017年1月14日12UTCの日本域地上天気図と衛星 赤外画像を図4.12.8と図4.12.9に示す。日本付近は冬 型の気圧配置となっており、日本海側を中心に筋状雲 が流れ込んでいるのがわかる。この筋状雲により、こ の日は強の着氷が2件、並の着氷が1件観測された。

図 4.12.10 に 14 日 12UTC を対象時刻とする着氷指 数の予測(上段)を示す。左から GSM 着氷指数、MSM 着氷指数、LFM 着氷指数で、FL070 の予測である。図 中の四角は対象時刻の前後90分以内で高度差1,000 ft 以内に観測された着氷の実況で、赤色が強の着氷、黄 色が並の着氷である。愛知県上空では強の着氷が観測 されており、これについては全てのモデルでこの着氷 を予測することができていたが、福岡県上空で観測さ れた並の着氷に関しては、全てのモデルで予測できて いなかった。この時の湿数の予測(下段)を見ると、愛 知県上空では概ね湿数が 1.5°C 以下となっていたのに 対し、福岡県上空では湿数が15°C以上となっていた。 同時刻の福岡のラジオゾンデによる高層気象観測では FL070付近は湿数が3°C程度となっていたことから、 モデルの予測は乾燥させすぎており、着氷指数の予測 が外れる要因となった。

冬型の気圧配置での筋状雲の発生や位置を正確に予 測することは難しいため、見逃しとなる場合があるこ とに注意が必要である。

#### (3) 氷晶による着氷

2016 年 7 月 10 日 00UTC の日本域地上天気図と衛星赤外画像を図 4.12.11 と図 4.12.12 に示す。日本付近は停滞前線が解析されており、前線の暖域側で活発な対流雲が見られる。この時間に種子島上空の FL400 という高高度で、並の着氷が 2 件観測された。この航空機が観測した気温は -53  $^{\circ}$ C であった。

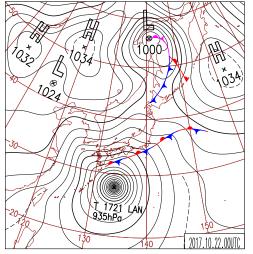


図 4.12.5 2017 年 10 月 22 日 00UTC の日本域地上天気図。

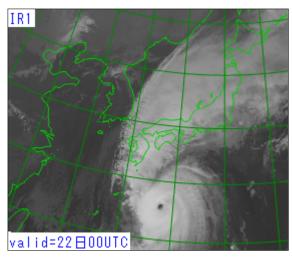


図 4.12.6 2017 年 10 月 22 日 00UTC の衛星赤外画像。

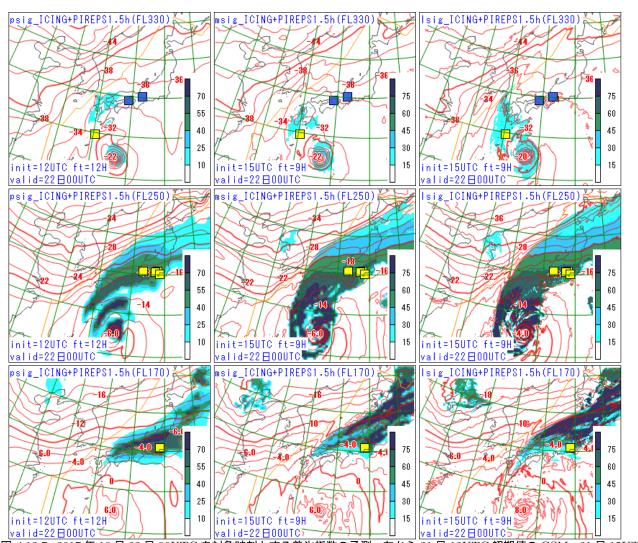


図 4.12.7 2017 年 10 月 22 日 00UTC を対象時刻とする着氷指数の予測。左から 21 日 12UTC 初期値の GSM、21 日 15UTC 初期値の MSM、21 日 15UTC 初期値の LFM の予測で、上段から FL330, FL250, FL170 の予測である。GSM 着氷指数は 10 以上、MSM・LFM 着氷指数は 15 以上から着色している。赤線はそれぞれの気温の予測を示す。図中の四角は対象時刻の前後 90 分以内で高度差 1,000 ft 以内に観測された着氷の実況で、黄色が並の着氷、青色が弱の着氷である。

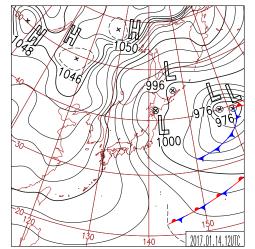


図 4.12.8 2017 年 1 月 14 日 12UTC の日本域地上天気図。

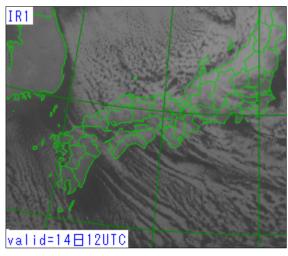


図 4.12.9 2017 年 1 月 14 日 12UTC の衛星赤外画像。

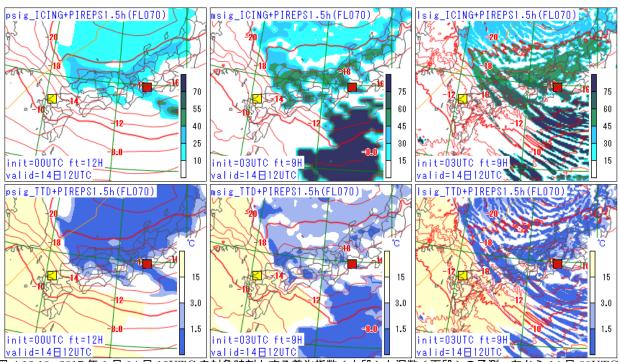


図 4.12.10 2017 年 1 月 14 日 12UTC を対象時刻とする着氷指数(上段)と湿数(下段)の予測。左から 14 日 00UTC 初期値の GSM、14 日 03UTC 初期値の MSM、14 日 03UTC 初期値の LFM の予測で、全て FL070 の予測である。 GSM 着氷指数は 10 以上、MSM・LFM 着氷指数は 15 以上から着色している。赤線はそれぞれの気温の予測を示す。図中の四角は対象時刻の前後 90 分以内で高度差 1,000 ft 以内に観測された着氷の実況で、赤色が強の着氷、黄色が並の着氷である。

図 4.12.13 に 10 日 00UTC を対象時刻とする着氷指数の予測を示す。左から GSM 着氷指数、MSM 着氷指数、LFM 着氷指数で、FL390 の予測である。赤線はそれぞれの気温の予測を示す。 着氷が観測された種子島周辺の気温は-50 °C 程度の予測となっており、航空機が観測した気温と概ね一致している。この気温では過冷却水滴は存在しないことや、周辺に活発な対流雲が観測されていることから、高濃度の氷晶による着氷と推定される。着氷指数は過冷却水滴による着氷を予測対象としているので、当然この着氷は全てのモデルの着氷指数で予測できていなかった。

#### 4.12.6 まとめと利用上の留意点

着氷指数は GSM, MSM, LFM から計算されており、 MSM 着氷指数と LFM 着氷指数は同じ計算式で、GSM 着氷指数は別の計算式から作成している。着氷指数は 0 から 100 の値を取るが、確率を表すものではない。 それぞれの着氷指数の予測精度に大きな差はないが、 GSM 着氷指数は、MSM 着氷指数と比べると体積率が低くなっており、捕捉率も下がるため、見逃しを減らしたい場合は日本周辺では MSM 着氷指数を主に利用するのが良いと思われる。事例検証 (2) で見たように、 冬型の気圧配置での筋状の雲域を正確に予測することは難しく、低気圧や前線近傍での着氷に比べて見逃し

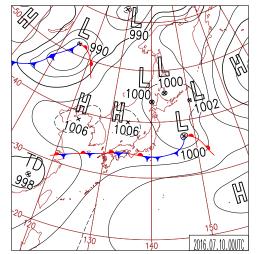


図 4.12.11 2016 年 7 月 10 日 00UTC の日本域地上天気図。

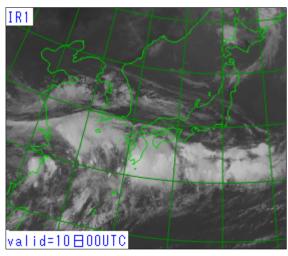


図 4.12.12 2016 年 7 月 10 日 00UTC の衛星赤外画像。

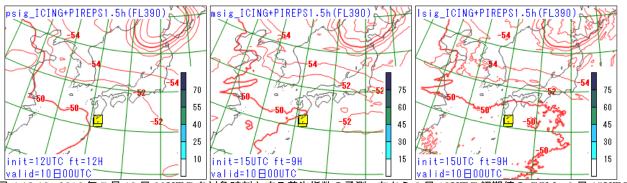


図 4.12.13 2016 年 7 月 10 日 00UTC を対象時刻とする着氷指数の予測。左から 9 日 12UTC 初期値の GSM、9 日 15UTC 初期値の MSM、9 日 15UTC 初期値の LFM の予測で、全て FL390 の予測である。GSM 着氷指数は 10 以上、MSM・LFM 着氷指数は 15 以上から着色している。赤線はそれぞれの気温の予測を示す。図中の四角は対象時刻の前後 90 分以内で高度 差 1,000 ft 以内に観測された着氷の実況で、黄色が並の着氷である。

が多くなることに注意が必要である。

#### 4.12.7 課題と今後の展望

外国の着氷の予測では、近年の数値予報モデルの精度向上により、雲物理過程で計算された雲物理量から直接着氷を予測する手法が用いられるようになって来ている。第 4.12.2 項で、単純に雲水量を用いて着氷を判別する方法より着氷指数の方が精度が高いことを述べたが、Thompson et al. (2017) では、雲水量などから着氷率を求めて、着氷率の大きさに応じて着氷の強度を予測するというより精緻な手法が使われている。今後、着氷指数の改良の参考にしたい。また、第 4.12.1 項の脚注や事例検証 (3) で述べた、-30 °C 以下でも発生する ICI については、過冷却水滴による着氷を予測対象にした現在の着氷指数では予測ができないので、今後 ICI について調査を行い予測手法の確立に努めたい。

#### 参考文献

FAA, 2017: PIREPs Relating to Airframe Icing. Aeronautical Information Manual, Federal Aviation Administration, 7–1–42.

Godske, C.L., T. Bergeron, J. Bjerknes, and R.C.

Bundgaard, 1957: Dynamic Meteorology and Weather Forecasting. *American Meteorological Society*, 769–773.

井藤智史, 2017: GSM 着氷指数の開発と GSM 積乱雲 頂高度予測の改良. 平成 28 年度数値予報研修テキスト, 気象庁予報部, 102–112.

工藤淳, 2006: 着氷について. 航空気象ノート第 65 号, 気象庁航空気象管理官, 17-22.

工藤淳, 2008: 国内航空悪天 GPV. 平成 20 年度数値予報研修テキスト, 気象庁予報部, 92-98.

工藤淳, 2012: C-PIREP 10 年の統計. 航空気象ノート 第 73 号, 気象庁航空気象管理官, 12-19.

Mason, J., 2007: Engine Power Loss in Ice Crystal Conditions. *Boeing Aero fourth quarter*, 12–17.

中山章, 1996: 最新航空気象. 東京堂出版, 73-87.

Thompson, G., M.K. Politovich, and R.M. Rasmussen, 2017: A Numerical Weather Model's Ability to Predict Characteristics of Aircraft Icing Environments. Wea. Forecasting, 32, 207–221.

#### 4.13 積乱雲量・雲頂高度<sup>1</sup>

#### 4.13.1 はじめに

積乱雲は強い乱気流、強い着氷、ひょう、雷などを 伴い、その全てが航空機にとって非常に危険な現象で ある。そのため、できる限り積乱雲を回避することが、 航空機の安全や快適な運航には必要となる。

数値予報課では、積乱雲を予測する指標として、積 乱雲量と積乱雲頂高度を計算している。積乱雲量は積 乱雲の雲量を 1/8 単位で予測した値であり、積乱雲量 が 2/8 以上である領域を積乱雲域としている。また、 航空機が積乱雲を回避するには、水平方向に避けるか 積乱雲の雲頂より高い高度を飛行する方法があり、積 乱雲の雲頂高度は重要な情報となる。

積乱雲頂高度の予測の歴史は古く、1973 年 10 月に 自動作成されるようになった航空用悪天予想図での積 乱雲頂高度の予測に遡る。当時は対流圏下層の安定 度、相対湿度、パーセル法により積乱雲頂高度を求 めていた (菊池 1978)。その後、1988 年 3 月からは、 850, 700, 500, 300 hPa の相対湿度の値および 400 hPa の上昇流の強さに応じて、積乱雲頂高度を予測するよ うになった (予報部数値予報課 1988)。

積乱雲量の予測は、MSM が正式運用となった 2001 年3月から国内航空悪天 GPV で開始し、降水量と SSI の予想に基づく診断的手法 (高田 1997; 工藤 2004) で求 めていた。2006年3月からは、MSM で積雲対流パラメ タリゼーションとして使用されている KF スキーム (山 田 2003) から出力される雲頂高度 (CLDTOP) の予想 に基づく手法に変更した。しかし、CLDTOP はパラメ タリゼーションから出力される値であるため、スキーム の変更等により CLDTOP の予測特性が大きく変わっ てしまう可能性がある。そこで将来的にも安定して積 乱雲の予測を行うために、CLDTOP に相当する雲頂高 度の予想値を、パラメタリゼーションに依らずに算出す る手法を開発した。新しい手法の積乱雲の予測は、KF スキームを参考にしたパーセル法に基づく診断法で求 めており、対流雲の雲頂高度の予測に基づいて積乱雲 量の予測を行うため、積乱雲量と積乱雲頂高度それぞ れについて矛盾のない予測が可能となった。この手法 が現在の積乱雲の予測の基礎となっており、2007年5 月から国内航空悪天 GPV でこの手法による積乱雲量と 積乱雲頂高度の予測を開始した (工藤 2007)。GSM の 予測値から算出する全球航空悪天 GPV・北太平洋航空 悪天 GPV (以下、2 つをまとめて呼ぶ場合は GSM 航 空悪天 GPV とする。) にも 2007 年 11 月からこの手法 による積乱雲頂高度の予測を取り入れた(松下 2007)。 その後も予測手法の改良を行っている (工藤 2009; 井 藤 2017) が、基本的な手法は変わっていない。LFM の 予測値から算出する LFM 航空悪天 GPV では 2012 年 8月から積乱雲の予測を開始している。

1 井藤 智史

空域予報現業では、国内航空悪天予想図 (FBJP) の 作成時に積乱雲量・雲頂高度の予測を利用しているほ か、自動作成される国内悪天12時間予想図 (FBJP112– 412) にも積乱雲量の予測が利用されている。

なお、空域予報に関連する用語については付録 4.11.A にまとめているので、そちらを参照して頂きたい。

#### 4.13.2 仕様

モデル毎の積乱雲量・雲頂高度の予測の仕様を表4.13.1 に示す。全球(北太平洋)航空悪天 GPV では格子間隔1.25 (0.5) 度の等緯度経度座標で積乱雲頂高度を作成している。北太平洋航空悪天 GPV は、福岡FIR の空域予報支援のため高解像度で作成している。国内 (LFM) 航空悪天 GPV では40 (10) km 格子のポーラーステレオ座標で積乱雲量と積乱雲頂高度を作成している。層別化については、GSM 航空悪天 GPVでのみー10°C 面高度で層別化している。国内航空悪天 GPVでも持ち上げ凝結高度 (LCL:Lifted Condensation Level) での気温によって層別化をしていたが、層別化による精度への影響はほとんどなかったので、2009年3月の改良の以降は層別化をせずに積乱雲の判定を行っている(工藤2009)。逐次学習は全ての航空悪天 GPV で行っていない。

#### 4.13.3 計算手法

積乱雲量・雲頂高度の計算手法を以下に示す。 予測手順

- 1. モデルの降水量予測から、対流雲が発生する可能性のある格子点を求める。GSM 航空悪天 GPV では、ショワルター安定指数 (SSI, 925-700 hPa) による対流雲の判定も行う。
- 2. その格子点で、モデル面の最下層から約50 hPa分の平均的な気塊を、LCLまで乾燥断熱的に持ち上げる。
- 3. LCL での気塊の仮温度に摂動項を加えた値が LCL での大気の仮温度よりも高い場合、その格子点で対流雲が発生すると判定する。
- 4. 対流雲が発生すると判定された場合、LCL から気 塊を湿潤断熱的に持ち上げ、ある高さでの大気の 上昇流が負になるか、摂動項を加えた気塊の仮温 度が大気の仮温度よりも低くなる高度を、対流雲 の雲頂高度とする。
- 5. 対流雲の雲頂気温と LCL での気温の差から、その対流雲が積乱雲であるかどうかを判別する。
- 6. 国内 (LFM) 航空悪天 GPV の格子(水平 40 (10) km) 内に含まれる MSM (LFM) の格子(水平 5 (2) km) 点のうち、積乱雲と判別された格子点の割合を、積乱雲量とする。
- 7. 積乱雲量が 2/8 以上である場合、手順 6 で積乱雲 と判別された格子の雲頂高度の平均値を、積乱雲 頂高度とする。

表 4.13.1 モデル毎の積乱雲量・雲頂高度	度の仕様
-------------------------	------

	全球航空悪天 GPV	北太平洋航空悪天 GPV	国内航空悪天 GPV	LFM 航空悪天 GPV
作成対象モデル		GSM	MSM	LFM
格子間隔	1.25 度	0.5 度	40 km	10 km
格子系	等	緯度経度座標	ポーラーステレオ座標	
対象領域	全球	$10^{\circ}\text{S}-65^{\circ}\text{N}, 80^{\circ}\text{E}-110^{\circ}\text{W}$	日本及びその周辺	日本周辺
予報対象時間	6 時間毎の正時	3 時間毎の正時	毎正時	
予報期間	FT=6~36 FT=3~36		FT=1~39	FT=1~9
予報要素	積	乱雲頂高度	積乱雲量・	<b>積乱雲頂高度</b>
層別化	-10 °C 面高度 (4 km	m 以下、4~6 km、6 km 以上)	t <sub>c</sub>	<b>1</b> 1
逐次学習		なし		

表 4.13.2 モデル毎の対流雲・積乱雲判定条件

	GSM 航空悪天 GPV			国内航空悪天 GPV	LFM 航空悪天 GPV	
−10 °C 面高度	4 km 以下	$4\sim6~\mathrm{km}$	6 km 以上	層別的	化なし	
降水量	≥0.8 mm/3h	$\geq$ 1.2 mm/3h	$\geq$ 1.4 mm/3h	≥0.1 mm/h	≥0.1 mm/h	
SSI	≤1.2			使用	]せず	
$\Delta T$	≥26 K	≥28 K	≥30 K	≥28 K	≥25 K	

GSM 航空悪天 GPV では予測手順 5 までを行い、積 乱雲量は求めていない。GSM のモデル値を水平方向には 0.5 度の緯度経度座標(全球領域)に、鉛直方向には 1,000 ft 毎に内挿した値を用いて計算を行う。その後、全球航空悪天 GPV へは 1.25 度の緯度経度座標に内挿した値を出力し、北太平洋航空悪天 GPV へは対象領域の値を切り出して出力する。また、GSM 航空悪天 GPV では -10 °C 面高度で層別化しており、4 km 以下、 $4\sim6$  km、6 km 以上の 3 つに層別化し、積乱雲頂高度を計算している。これは温帯域や熱帯域などの地域差(熱帯域の方が雲頂高度が高くなりやすい。)を考慮するためと、季節による差(夏季の方が雲頂高度が高くなりやすい。)を考慮するためである。国内・LFM 航空悪天 GPV では層別化はしていない。以下、予測手順 1,3,5 について補足説明する。

#### 予測手順 1

モデル毎の対流雲発生の判定条件を表 4.13.2 に示す。 GSM 航空悪天 GPV では、-10 °C 面高度が 4 km 以下では前 3 時間降水量が 0.8 mm 以上、 $4\sim6$  km では 1.2 mm 以上、6 km 以上では 1.4 mm 以上で対流雲が発生すると判定する。また、SSI が 1.2 以下の時に対流雲が発生すると判定する。SSI (850-500 hPa) も試したが精度は変わらなかったことや、SSI (925-700 hPa) の場合 -10 °C 面高度に依らず同じ閾値で判定しても精度が変わらなかったことから、層別化をせずに SSI (925-700 hPa) を用いている。国内航空悪天 GPV では予報対象時刻の前後 30 分間の 1 時間降水量が 0.1 mm以上の場合に、上FM 航空悪天 GPV では予報対象時刻の前 1 時間降水量が 0.1 mm以上の場合に、対流雲が発生すると判定する。

#### 予測手順3

LCL での気塊の仮温度に加える摂動項は、KF スキーム (成田 2008) を参考に、上昇流の強さに応じた摂動項  $(\Delta T_{\rm LCL})$  と、水蒸気に応じた摂動項  $(\Delta T_{\rm RH})$  を与えている。

$$\Delta T_{\rm LCL} = \max \left[ c_1 (100W_{\rm LCL} \frac{\Delta x}{25000} - 100W_{\rm LCL})^{\frac{1}{3}}, 0 \right]$$

$$W_{\rm LCL} = \begin{cases} 0.02 \times \frac{Z_{\rm LCL}}{2000} & (Z_{\rm LCL} \le 2000 \text{ m}) \\ 0.02 & (Z_{\rm LCL} > 2000 \text{ m}) \end{cases}$$

$$\Delta T_{\rm RH} = \begin{cases} 0 & (RH_{\rm LCL} < 75\%) \\ c_2 \frac{0.25(RH_{\rm LCL} - 0.75)q_{\rm mix}}{\partial q_{\rm LCL}/\partial T} \\ & (75\% \le RH_{\rm LCL} \le 95\%) \\ c_2 \frac{(1/RH_{\rm LCL} - 1)q_{\rm mix}}{\partial q_{\rm LCL}/\partial T} & (RH_{\rm LCL} > 95\%) \end{cases}$$

ここで、 $\Delta x$  はモデルの水平格子間隔 [m] 、 $Z_{\rm LCL}$  は持ち上げ凝結高度 [m] 、 $q_{\rm mix}$  は持ち上げた気塊の混合比 [kg/kg] 、T は気温 [K] 、 $W_{\rm LCL}$  ,  $RH_{\rm LCL}$  ,  $q_{\rm LCL}$  はそれぞれ、LCL における大気の鉛直風 [m/s] 、相対湿度、混合比 [kg/kg] である。 $c_1$  と  $c_2$  は最適化係数で、 $c_1$  と  $c_2$  が大きいほど摂動項が大きくなるため、 $c_1$  と  $c_2$  を大きくすると対流雲として判定されやすくなり、雲頂高度も高く予測されやすくなる。モデル毎の最適化係数の値を表 4.13.3 に示す。GSM 航空悪天 GPV では、-10 °C 面高度が 4 km 以下では  $c_1$ =0.7、4~6 km では  $c_1$ =0.8、6 km 以上では  $c_1$ =1.1 としており、国内航空悪天 GPV では  $c_1$ =0.7、LFM 航空悪天 GPV では  $c_1$ =1.0 としている。 $c_2$  は全ての航空悪天 GPV で  $c_1$  の 1.25 倍としており、参考にした KF スキームで用いられている  $c_1$  と  $c_2$  の比率と同じにした。KF スキームで

耒	4.13.3	モデル毎の最適化	係数
ᄯ	4.10.0		IN ZX

	GSM 航空悪天 GPV			国内航空悪天 GPV	LFM 航空悪天 GPV
−10 °C 面高度	4 km 以下	4∼6 km	6 km 以上	層別	化なし
最適化係数 $c_1$	0.700	0.80	1.100	0.700	1.00
最適化係数 $c_2$	0.875	1.0	1.375	0.875	1.25

は $c_1$ =0.2,  $c_2$ =0.25 である。航空悪天 GPV の最適化係数の方が値が大きいのは、モデルでは積雲対流が発生する前の状態の不安定な大気に対して KF が用いられているのに対して、航空悪天 GPV では KF によって不安定が解消された後の大気の状態に対して KF 的な手法を用いているので、モデルと比べて不安定度が小さい分だけ、より大きなパラメータを用いる必要があるからである。

#### 予測手順5

モデル毎の積乱雲判定条件を表 4.13.2 に示す。GSM 航空悪天 GPV では、-10 °C 面高度が 4 km 以下では 対流雲の雲頂気温と LCL での気温の差  $(\Delta T)$  が 26 K 以上、 $4\sim6$  km では 28 K 以上、6 km 以上では 30 K 以上の場合に積乱雲と判定する。国内航空悪天 GPV では  $\Delta T$  が 28 K 以上の場合に、LFM 航空悪天 GPV では  $\Delta T$  が 25 K 以上の場合に積乱雲と判定する。

#### 4.13.4 統計検証

## (1) 国内航空悪天 GPV と LFM 航空悪天 GPV の検証 積乱雲量

積乱雲量の統計検証の結果を示す。検証に用いる実 況データは、本来であれば積乱雲量の実況を用いるべき であるが、実況資料から積乱雲量を定義することは難し い。そこで積乱雲量の検証は、雷監視システム (LIDEN) の発雷実況を用いて行う。国内航空悪天 GPV の格子 内(水平 40 km)で予報対象時刻の前後 30 分以内に 対地雷が1つ以上観測された場合に「実況あり」とし て、積乱雲量の予想値と比較する。検証スコアは分割 表(巻末付録 A.3.1)を作成し、以下で定義する遭遇率 比と巻末付録 A.3.6 で示す体積率を用いる。体積率は、 全事例のうち予測ありの事例の割合であり、予報値だ けで決まる値である。

遭遇率比 = 
$$\frac{FO/(FO+FX)}{(FO+XO)/(FO+FX+XO+XX)}$$

遭遇率比は、全領域での積乱雲の発生率に対する予報領域での積乱雲の発生率であり、遭遇率比が大きいほど予報領域で積乱雲が発生する可能性が高いことを示す。工藤(2007)で述べたとおり、積乱雲域の中である割合で発雷していると仮定した場合、発雷を実況とした場合と積乱雲を実況とした場合とで遭遇率比は等しくなる。遭遇率比と体積率を用いることで、予測対象と異なる発雷の観測を実況とした場合でも、積乱雲の観測を実況とした場合に近い検証結果が得られる。

検証期間は2015年12月~2016年11月の1年間で、

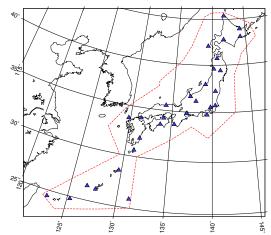


図 4.13.1 積乱雲頂高度と積乱雲量の検証領域。赤点線で囲まれた領域を検証対象とする。 は LIDEN の検知局の位置を示す。

国内航空悪天 GPV は 2017 年 2 月 28 日に変更された MSM の試験データから作成し、LFM 航空悪天 GPV は本期間に予測精度に大きく影響を与えるような LFM の変更は行われなかったため、ルーチンの LFM のデータから作成した。LFM 航空悪天 GPV の積乱雲量の検証は、国内航空悪天 GPV の格子に内挿して行う。ともに 03, 09, 15, 21UTC 初期値で、 $FT=6\sim9$  で検証した。検証領域は図 4.13.1 の赤点線で囲まれた国内の主要航空路を概ね含む領域である。

図 4.13.2 に、積乱雲量を 1/8 から 8/8 まで 1/8 ずつ変化させて積乱雲の予測を行った場合の検証結果を示す。同じ体積率で予報を比較した場合に、遭遇率比が大きいほど積乱雲を予測する指数としての精度が高いと言える。概ねどの季節も国内航空悪天 GPV の方が精度が高くなっている。特に冬季にその差が大きいが、原因は分かっていない。

#### 積乱雲頂高度

積乱雲頂高度の検証はレーダーエコー頂高度を用いて行う。ただし、レーダーエコー頂高度は雲頂高度よりも低くなると考えられるため、ここでは水平  $40~(10)~{\rm km}$ の国内 (LFM) 航空悪天 GPV の格子内に含まれる、水平  $1~{\rm km}$  のレーダーエコー頂高度の最大値を実況とする。レーダーエコー頂高度は  $2~{\rm km}$  単位のレベル値として GPV に格納されているため、積乱雲頂高度の予測値もこれと同じレベル値に変換してから比較する。また、ここでは積乱雲頂高度の予測精度を見るために、積乱雲量 2/8 以上で積乱雲を予想した格子内で、LIDEN で  $1~{\rm cm}$  つ以上対地雷が観測された場合のみ対象とする。

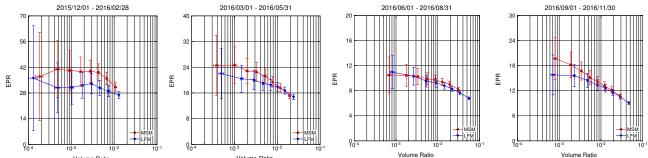


図 4.13.2 閾値別の積乱雲量の検証結果。左から冬季(12 月~2 月)、春季(3 月~5 月)、夏季(6 月~8 月)、秋季(9 月~11 月)の結果。横軸は体積率(Volume Ratio)、縦軸は遭遇率比 (EPR) を示す。赤線が国内航空悪天 GPV、青線が LFM 航空悪天 GPV の検証結果で、右から雲量が 1/8, 2/8, 3/8, 4/8, 5/8, 6/8, 7/8, 8/8 のスコアをプロットしている。エラーバーはプートストラップ法による 95%信頼区間を示す。

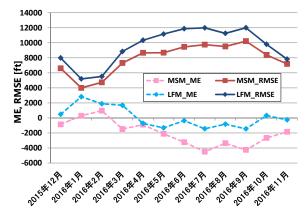


図 4.13.3 月別の積乱雲頂高度の ME(破線)と RMSE(実線)。赤系統の線が国内航空悪天 GPV、青系統の線が LFM 航空悪天 GPV の検証結果。

図 4.13.3 は、月別の平均誤差 (ME) および二乗平均 平方根誤差 (RMSE) である。ME を見ると、国内航空 悪天 GPV は雲頂高度の低くなりやすい冬季にバイアスが適正となっているが、その他の季節は負バイアスとなっており、特に雲頂高度の高くなりやすい夏季に負バイアスが大きくなっている。LFM 航空悪天 GPV は冬季にやや正バイアスとなっており、その他の季節では概ね適正なバイアスとなっている。RMSE は、どちらのモデルも冬季に誤差が小さく夏季を中心に誤差が大きくなっている。また、どの季節も国内航空悪天 GPV の方が誤差が小さくなっている。このため、概ね国内航空悪天 GPV の予測を利用する方が良いが、夏季を中心に負バイアスが大きくなるので、国内航空悪天 GPV の予測値よりも高めに雲頂高度を見積もる必要がある。

## (2) 北太平洋航空悪天 GPV の検証 積乱雲判別と積乱雲頂高度

北太平洋航空悪天 GPV の積乱雲予測の検証結果を示す。北太平洋航空悪天 GPV は主に MSM の領域外(気象レーダーの観測範囲外)で利用されるため、検証には衛星観測から作成される従来型雲量格子点情報(気

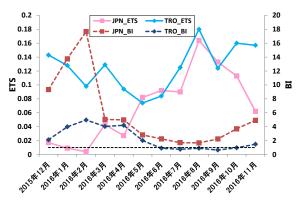


図 4.13.4 北太平洋航空悪天 GPV の月別の積乱雲判別の ETS (左軸、実線)と BI (右軸、破線)。赤系統の線が日本付近、青系統の線が熱帯付近の検証結果。黒破線は BI=1 を示す。

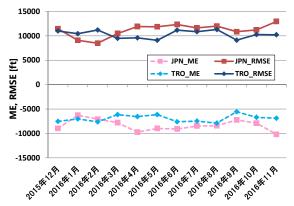


図 4.13.5 北太平洋航空悪天 GPV の月別の積乱雲頂高度 の ME ( 破線 ) と RMSE ( 実線 )。赤系統の線が日本付近  $(20^\circ-45^\circ\mathrm{N},\,114^\circ-180^\circ\mathrm{E})$ 、青系統の線が熱帯付近  $(0^\circ-20^\circ\mathrm{N},\,114^\circ-180^\circ\mathrm{E})$  の検証結果。

象庁観測部 2017) の雲型判別と積乱雲頂高度のデータを用いた。北太平洋航空悪天 GPV は 2017 年 5 月 25 日に変更された GSM の試験データから作成した。検証期間は 2015 年 12 月 ~ 2016 年 11 月である。領域は日本付近  $(20^\circ-45^\circ\text{N},\ 114^\circ-180^\circ\text{E})$  と、熱帯付近  $(0^\circ-20^\circ\text{N},\ 114^\circ-180^\circ\text{E})$ 

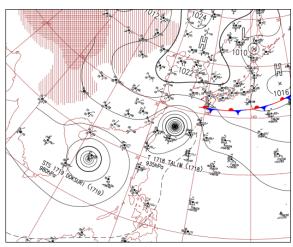


図 4.13.6 2017 年 9 月 14 日 00UTC のアジア太平洋地 上天気図。台風周辺を拡大して表示。

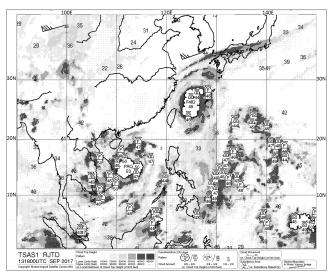


図 4.13.7 2017 年 9 月 13 日 18UTC の TSAS1。台風周辺を拡大して表示。衛星観測により積乱雲として判定された領域が波線または積乱雲を模した記号で示され、その雲量と雲頂高度  $(\times 1,000~{\rm ft})$  が示されている。

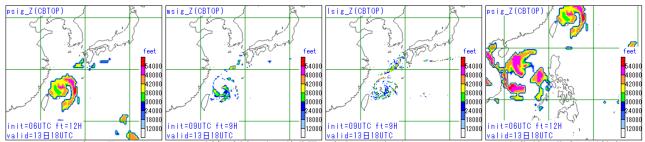


図 4.13.8 2017 年 9 月 13 日 18UTC を対象時刻とする積乱雲頂高度の予測。左から 13 日 06UTC 初期値の北太平洋航空悪天 GPV、13 日 09UTC 初期値の国内航空悪天 GPV、13 日 09UTC 初期値の LFM 航空悪天 GPV の日本付近の予測、13 日 06UTC 初期値の北太平洋航空悪天 GPV のフィリピン付近の予測。

 $114^{\circ}-180^{\circ}E)$  で検証した。月別の積乱雲判別の検証結果を図 4.13.4 に示す。実況は衛星観測の雲型が積乱雲と判定された場合に「積乱雲あり」、それ以外の雲型の場合に「積乱雲なし」とし、予測は積乱雲頂高度が予測されている格子を「積乱雲あり」、それ以外の格子を「積乱雲なし」として、エクイタブルスレットスコア (ETS) とバイアススコア (BI) を求めた。BI は日本付近で 1 より大きく、予測頻度が過大となっており、特に冬季に顕著であることが分かる。熱帯付近は、 $6\sim11$  月は概ね 1 に近く予測頻度は適切となっているが、 $12\sim5$  月まではやや過大となっている。また、ETSについては日本付近で夏季の方が精度が良く、冬季に悪くなっていることが分かる。

月別の積乱雲頂高度予測の検証結果を図 4.13.5 に示す。平均誤差 (ME) は日本付近で  $6,000\sim10,000$  ft の負バイアスがあり、熱帯付近で  $6,000\sim8,000$  ft の負バイアスとなっている。二乗平均平方根誤差 (RMSE) は日本、熱帯付近ともに  $9,000\sim12,000$  ft 程度となっている。

#### 4.13.5 事例検証

#### (1) 台風による積乱雲と熱帯域での積乱雲

台風による積乱雲と熱帯域での積乱雲の予測事例を 示す。図 4.13.6 は 2017 年 9 月 14 日 00UTC のアジア 太平洋地上天気図で、図 4.13.7 は 2017 年 9 月 13 日 18UTC の広域雲解析情報図 (TSAS1) である。台風第 18号が東シナ海に、台風第19号が南シナ海にあり、 それぞれ雲頂高度は最大で 49,000 ft と 53,000 ft と解 析されている。また、フィリピンの東海上にも積乱雲 群があり、雲頂高度は最大で 51,000 ft と解析されて いる。図 4.13.8 は、積乱雲頂高度の予測で、左から 13 日 06UTC 初期値の FT=12 における北太平洋航空悪天 GPV、13 日 09UTC 初期値の FT=9 における国内航 空悪天 GPV、13 日 09UTC 初期値の FT=9 における LFM 航空悪天 GPV の日本付近の予測、13 日 06UTC 初期値の FT=12 における北太平洋航空悪天 GPV の フィリピン付近の予測である。台風第18号の予測につ いては、北太平洋航空悪天 GPV の予測領域はやや過 大で、国内航空悪天 GPV と LFM 航空悪天 GPV の予 測領域は過小となっている。この傾向は台風事例だけ

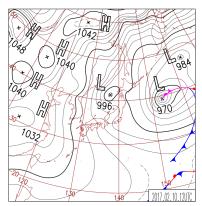


図 4.13.9 2017 年 2 月 10 日 12UTC の日本域地上天気図。

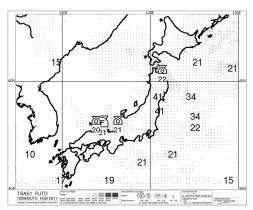


図 4.13.10 2017 年 2 月 10 日 09UTC の TSAS1。日本周辺を拡大して表示。

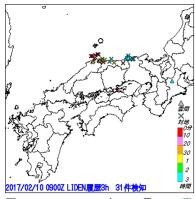


図 4.13.11 2017 年 2 月 10 日 09UTC の前 3 時間発雷履歴。× マークが対地放電、 マークが雲 間放電を示し、色は発雷の時間を 示す。

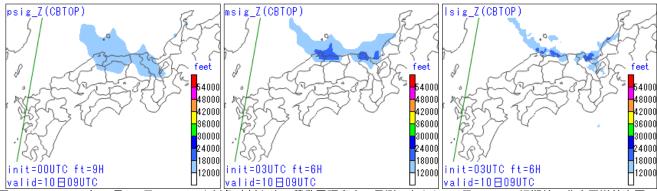


図 4.13.12 2017 年 2 月 10 日 09UTC を対象時刻とする積乱雲頂高度の予測。左から 10 日 00UTC 初期値の北太平洋航空悪 天 GPV、10 日 03UTC 初期値の国内航空悪天 GPV、10 日 03UTC 初期値の LFM 航空悪天 GPV の予測。

ではなく、温帯低気圧の事例でも見られる傾向である。 雲頂高度の予測については、北太平洋航空悪天 GPV と LFM 航空悪天 GPV の最大値は 54,000 ft と高めで、 国内航空悪天 GPV の最大値は 45,000 ft と低めとなっ ている。図 4.13.3 で示したように、国内航空悪天 GPV は低めの予測となる場合が多いので、注意が必要であ る。台風第 19 号の予測については、位置ずれはあるも のの、雲頂高度の予測は 54,000 ft と解析値に近い予測 ができている。

一方で、フィリピンの東海上の積乱雲群については ほとんど予測ができていない。熱帯域での予測は、台 風のようにまとまった擾乱に対してはある程度精度良 く予測できるが、まとまりのない積乱雲については予 測精度は低い場合が多い。

#### (2) 冬型降水時の積乱雲

冬型降水時の積乱雲の予測事例を示す。図 4.13.9 は 2017 年 2 月 10 日 12UTC の日本域地上天気図で、図 4.13.10 は 2017 年 2 月 10 日 09UTC の TSAS1 である。日本の東海上と日本海に低気圧、中国大陸には高気圧があり、日本付近は冬型の気圧配置となっている。近畿から中国地方の日本海側では、日本海寒帯気団収束

帯 (JPCZ) により積乱雲が発達しており、雲頂高度は 21,000 ft と解析されている。また、09UTC の前 3 時 間発雷履歴(図 4.13.11)でも JPCZ 近傍で活発に発雷 しているのが分かる。図 4.13.12 は、積乱雲頂高度の予 測で、左から 10 日 00UTC 初期値の FT=9 における北 太平洋航空悪天 GPV、10 日 03UTC 初期値の FT=6 における国内航空悪天 GPV、10 日 03UTC 初期値の FT=6 における LFM 航空悪天 GPV の予測である。ど のモデルも日本海側の積乱雲を予測しており、北太平 洋航空悪天 GPV の予測領域がやや広いものの、概ね 実況に近い予測をしている。雲頂高度の予測について は、北太平洋航空悪天 GPV が 17,000 ft、国内航空悪 天 GPV が 18,000 ft、LFM 航空悪天 GPV が 19,000 ft とやや低めとなっているが、解析値に近い予測となっ ている。この事例のように冬季は比較的実況に近い予 測となる場合が多く、信頼性が高いと言える。

#### 4.13.6 まとめと今後の展望

積乱雲頂高度は全球・北太平洋・国内・LFM 航空悪 天 GPV で、積乱雲量は国内・LFM 航空悪天 GPV で 計算しており、本節ではそれぞれの計算手法について 示した。統計検証から、北太平洋航空悪天 GPV は日 本付近で積乱雲の予測頻度が過大で、積乱雲頂高度の 予測は日本付近、熱帯付近ともに負バイアスとなって いる。国内航空悪天 GPV と LFM 航空悪天 GPV は、 冬季に予測誤差が小さく夏季に予測誤差が大きくなる。 国内航空悪天 GPV の積乱雲頂高度予測は夏季に負バ イアスが大きいので、予測値よりも高めに雲頂高度を 見積もる必要がある。

今後の展望を以下に述べる。北太平洋航空悪天 GPV の積乱雲の予測頻度は過大となっているが、これは積乱雲量を予測していないことが原因の一つである。国内・LFM 航空悪天 GPV では積乱雲量が 2/8 以上で積乱雲ありとしており、積乱雲の予測頻度を抑えている。今後 GSM 航空悪天 GPV でも積乱雲量の予測を導入して、予測頻度の適正化を図りたい。国内航空悪天 GPV については、夏季に積乱雲頂高度予測の負バイアスが大きくなるので、季節変化を抑えるための手法を検討する必要がある。LFM 航空悪天 GPV については、積乱雲頂高度・積乱雲量ともに国内航空悪天 GPV より予測精度が低くなっている。原因を究明し、予測精度の向上に努めたい。

#### 参考文献

- 井藤智史, 2017: GSM 着氷指数の開発と GSM 積乱雲 頂高度予測の改良. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 102-112.
- 菊池正武, 1978: 航空予報への利用(悪天予想図). 気 象研究ノート, **134**, 385-397.
- 気象庁観測部, 2017: 「ひまわり9号」の待機運用開始と「ひまわり8号・9号」による衛星関連プロダクトについて. 配信資料に関する技術情報第456号, 21-26.
- 工藤淳, 2004: 国内航空悪天 GPV. 平成 16 年度数値予 報研修テキスト, 気象庁予報部, 58-62.
- 工藤淳, 2007: 国内航空用悪天 GPV の積乱雲予測手法の開発. 航空気象ノート第 66 号, 気象庁航空気象管理官, 11–18.
- 工藤淳, 2009: 国内航空悪天 GPV の積乱雲予測手法の 改良. 航空気象ノート第 69・70 号, 気象庁航空気象 管理官, 1-8.
- 松下泰広, 2007: 全球航空悪天 GPV および北太平洋航空悪天 GPV. 平成 19 年度数値予報研修テキスト, 気象庁予報部, 84-86.
- 成田正巳, 2008: Kain-Fritsch スキームの改良とパラ メータの調整. 数値予報課報告・別冊 54 号, 気象庁 予報部, 103-111.
- 高田伸一, 1997: 国内悪天予想資料の CB 予測について. 平成 8 年度航空気象予報技術検討会及び航空気象予報研修, 気象庁予報部, 123-126.
- 山田芳則, 2003: 積雲対流スキーム. 数値予報課報告・ 別冊 49 号, 気象庁予報部, 77-89.

予報部数値予報課, 1988: 国際航空用悪天予想図につい て. 測候時報, **55**, 1–16.

# 第5章 今後の展望

#### 5.1 ガイダンスの今後<sup>1</sup>

第 1 章では、2018 年現在までの気象庁のガイダンスと海外気象機関のガイダンス、第 3 章ではガイダンスの開発と運用について、第 4 章では 2018 年現在のガイダンスについて解説した。これらを踏まえ、今後のガイダンスについて、現時点で計画していること、将来的なガイダンスの方向性を述べる。

#### 5.1.1 ガイダンスの開発と改良

この項では本原稿執筆時点(2018年3月現在)で計画されているガイダンスの開発及び改良について述べる。なお、計画されている内の統合型ガイダンスについては、次項で別途説明する。

#### (1) メソアンサンブル予報を使ったガイダンス

メソモデル (MSM) の予測に対して信頼度・不確実性 の情報を付加することを目的としたメソアンサンブル 予報システム (MEPS) が、本原稿執筆時点で部内試験 運用されている。部内試験運用は1日1回(18 UTC 初 期値)39時間予報を11メンバーで実施しているが、次 期スーパーコンピュータシステム導入(2018年6月予 定)後約1年後までに、1日4回(00,06,12,18 UTC) と高頻度化するとともに、メンバー数が21に増強さ れる予定である (石田 2017)。この MEPS を入力とし たガイダンスを実験的に作成しており、防災情報及び 航空気象予報への利用を考えて、地点形式では、気温、 風、降雪量(地点)、視程及び雲ガイダンス、格子形 式では、降水量(平均・最大) 最大降雪量、発雷確率 ガイダンス、視程分布予想及び航空悪天 GPV がある (2018年3月現在)。現在の手法は MSM ガイダンスと 同様の予測式を各メンバーに適用している簡便なもの であるが、11個の予測を平均することによって、概ね どのガイダンスも MSM ガイダンスより精度が向上す ることを確認している。

また、より適切な予測の信頼度情報を提供するには、MEPS ガイダンスの予測のばらつきから、ガイダンスの予測誤差を見積もる手法を検討する必要がある。調査からは単に MEPS ガイダンスの予測のばらつきのみでは適切な予測誤差を見積もるのは難しいと思われ、第1.4 節で紹介した米国気象局が運用している EKDMOSで行っているような統計的な校正手法 (Veenhuis 2013)を検討する必要がある。

#### (2) 局地モデルを使ったガイダンス

局地モデル (LFM) は、2012 年 8 月から東日本領域で 1 日 8 回 (3 時間ごと)の運用が開始され、2013 年 5 月から日本領域で 1 日 24 回 (1 時間ごと)と高頻度の運用が開始された (永戸ほか 2013)。 LFM ガイダン

スは、精度面での MSM ガイダンス及び LFM の予測値からの優位性、高頻度かつ高分解能の LFM に適したガイダンスの作成方法を検討した後に開発を進めることとしていた。現在、これらの検討を行って降水量や風については予報作業への有効利用が見込めることが確認できつつあり、降水量と風について LFM ガイダンスの開発を進めている。

降水量については、統計的検証及び事例検証からも、 大雨について MSM の予測精度を上回ることが確認さ れており (原ほか 2015)、かつ平成 29年7月九州北部豪 雨等の線状降水帯や局地的大雨の事例について、LFM が MSM の予測を大きく改善できる例も多い。このよ うなことから LFM 降水量ガイダンスの作成が求めら れたため、開発を進め、現在は実験的に運用を行って いる状況である。このガイダンスは今までの降水量ガ イダンスと異なり、平均降水量ガイダンスではカルマ ンフィルタ (KF) 等の統計手法は使わずに、LFM の予 測降水量を LAF (第 2.2 節)を適用し、その後頻度バ イアス補正のみ行う。また、最大降水量ガイダンスで はニューラルネットワーク (NN) や線形重回帰 (MLR) を使わず、高分解能の LFM の降水量予測の分散を利 用した手法とする予定である。なお、現時点での検証 結果ではこの LFM 降水量ガイダンスは、LFM の降水 量予測を改善し、かつ MSM 降水量ガイダンスを上回 る精度が確認されている。

LFM はその高い解像度により、局地風の予測精度が高く、防災情報に有効に利用できることがわかってきた。平成 28 年 12 月の糸魚川の大規模火災は山越えの南風の強風の影響があったが、LFM はこの強風を概ね予測できていた。LFM の予測特性を最大限に活用することで、有効な風ガイダンスとなるよう開発を進める予定である。

#### (3) ガイダンスの改良

現在、最大降水量ガイダンスは 24 時間積算が最大である。近年の特徴として、防災対策のためにより長い時間幅の予測が求められるようになったため、予報現場からは 48,72 時間の最大降水量ガイダンスの要望がある。このニーズに対応するために、新たに 48,72 時間最大降水量ガイダンスの開発を行う予定である。この際には、現在提供している 1,3,24 時間最大降水量ガイダンスの改良も含めて開発を行う予定である。また、近年、モデルの高解像度化、精緻化等の高度化が進んでいるものの、平均降水量ガイダンスや降水確率ガイダンスの説明変数は、海老原 (1999) から 20 年近く変更が行われていない。このため、高度化したモデルに適合した説明変数の調査を行ってゆく必要がある。

その他のガイダンスについても随時改良を行う予定であるが、利用者からの問題点の指摘や要望(ニーズ)

<sup>1</sup> 高田 伸一

開発者の問題意識に基づく改良や乱気流指数や視程分布予想のような新たなガイダンスの提案(シーズ)の 両面から開発を進めてゆく予定である。

#### 5.1.2 統合型ガイダンス

前項では、2018 年現在 MEPS ガイダンス及び LFM ガイダンスを開発していることを述べた。GSM (全球モデル)ガイダンス及び MSM ガイダンスにこれらのガイダンスを加えると、特性の異なる多くの予測資料が溢れ、予報作業が煩雑になる可能性がある。第 1.4 節で紹介したように、ドイツ気象局の ModelMIX、英国気象局の Optimal blended forecasts、米国気象局のNBM など、他国では複数のガイダンス等を結合し、精度の高い統合型ガイダンスを運用している。今後は、気象庁でも統合型ガイダンスを開発を進める予定である。

統合型ガイダンスには、多くのガイダンスをまとめて予報作業を適切に支援する目的以外に、各ガイダンスの予測を平均化することによって、ガイダンスのランダム誤差を軽減し、精度を向上させる目的もある。ここでは簡単な例を用いて、ガイダンスの平均化によって誤差が軽減されることを示す。ある予測事例  $n\ (n=1\sim N)$ の 2 つの異なるガイダンスの予測値を  $f_{1n}, f_{2n}$ 、誤差を  $\epsilon_{1n}, \epsilon_{2n}$ 、平均二乗誤差を  $E_1, E_2$ 、実況を  $y_n$  とすると、平均したガイダンスの平均二乗誤差 E は以下のようになる。

$$E = \frac{1}{N} \sum_{n=1}^{N} \left[ \frac{1}{2} (f_{1n} + f_{2n}) - y_n \right]^2$$

$$= \frac{1}{4} \frac{1}{N} \sum_{n=1}^{N} \left[ (f_{1n} - y_n)^2 + (f_{2n} - y_n)^2 + 2 (f_{1n} - y_n) (f_{2n} - y_n) \right]$$

$$= \frac{1}{4} (E_1 + E_2) + \frac{1}{2} \frac{1}{N} \sum_{n=1}^{N} \epsilon_{1n} \epsilon_{2n}$$
 (5.1.1)

ここで、2 つのガイダンスの平均誤差がゼロであると仮定し(通常ガイダンスの平均誤差はゼロに近い)、その誤差の標本相関係数を  $\rho$  とすると、

$$\rho = \frac{1}{N} \sum_{n=1}^{N} \frac{\epsilon_{1n} \epsilon_{2n}}{\sqrt{E_1 E_2}} \tag{5.1.2}$$

よって、E は以下のように書ける。

$$E = \frac{1}{4} (E_1 + E_2) + \frac{\rho}{2} \sqrt{E_1 E_2}$$
 (5.1.3)

仮に 2 つのガイダンス間の平均二乗誤差がほぼ同じで  $(E_1=E_2)$ 、誤差の相関がない  $(\rho=0)$  という理想的な 状況では、 $E=E_1/2$  となり、平均化することで誤差 が半分になることがわかる。誤差相関が 1 の場合、すなわち、2 つのガイダンスの予測が完全に一致している場合には  $E=E_1=E_2$  となる。このことから、各

#### TGDC RMSE, 2015-04-01~2017-03-31

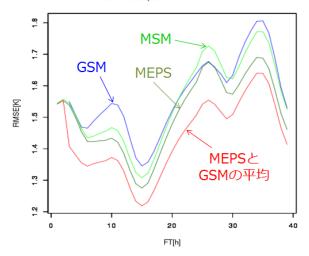


図 5.1.1 MSM, MEPS, GSM の時系列気温ガイダンスの予報時間 (FT) ごとの二乗平均平方根誤差 (RMSE)。 MSM, MEPS, GSM 共に 18UTC 初期値で、統計検証期間は 2015年 4月  $\sim 2017$ 年 3 月。緑線は MSM、深緑線は MEPS、青線は GSM、赤線は MEPS と GSM を平均した予測。

ガイダンスの誤差相関がより小さい場合には、平均化 することで誤差を軽減できることがわかる $^2$ 。

通常は誤差相関が 0 になることはないが、なるべく 誤差相関の低い予測を平均化することで誤差を減じる ことができる。具体的にはアンサンブル予報では、各 メンバーの予測の誤差相関が高いと考えられ、平均化 による誤差軽減の効果は小さいが、異なったモデルで は、誤差相関が低く、誤差軽減の効果が大きい。この ことを気温ガイダンスを例に図 5.1.1 に示す。現在実験 的に運用している MEPS の気温ガイダンスは、アンサ ンブル平均を用いることで、コントロールである MSM の気温ガイダンスの精度を上回っているが、MEPS 気 温ガイダンスと GSM の気温ガイダンスを平均するこ とによって、さらに誤差を大きく軽減できている。図 5.1.2 に MSM 気温ガイダンスと MEPS の各メンバーの 気温ガイダンス及び GSM 気温ガイダンス間の誤差の 標本相関係数を示した。図からわかるように、MSM と MEPS 気温ガイダンスの誤差相関より、MSM と GSM の誤差相関の方が低いことがわかる。MSM と GSM が 別のモデルであるのに対し、MSM と MEPS は基本的 には同じモデルのためである。この例のように、ガイ ダンスの誤差を減らすという観点では、アンサンブル 予報だけでなく、異なったモデルを平均化することも 効果的である。

(5.1.3) からわかるように、ガイダンスの予測を平均 化する際には、なるべく両ガイダンスの平均二乗誤差 に違いがないものを使い、精度が低いガイダンスを使

 $<sup>^2</sup>$  逆相関 (
ho < 0) だと更に誤差 E が小さくなるが、通常、ガイダンス間の誤差相関は負になることはないため、ここでは正で考えている。

わない方が良い。図 5.1.1 には MSM と GSM の気温ガイダンスの RMSE を示したが、その差は  $0.1^{\circ}$  C 未満であり、そのために平均によって精度が大きく向上したと考えられる。よって、統合型ガイダンスの精度を上げるためには、各ガイダンスの誤差及びガイダンス間の誤差相関を調べてから結合する必要がある。

統合型ガイダンスは予報作業の負担を軽減できること、 誤差を減じることができることの他に、以下のメリットも考えられる。

# 顕著現象の捕捉率の精度向上 現在までの調査の結果、MEPSのガイダンスは、 予測誤差に比べてスプレッド (MEPSのメンバー 間の予測値のばらつき)が不足し、顕著現象を捕 捉できない場合がある。他のモデルのガイダンス を加えることによって、顕著現象が起こる可能性 をより捕捉できる可能性がある。

# モデル更新の悪影響の軽減 多くのガイダンスを結合することによって、モデル更新によるガイダンスの精度低下を軽減できる。 これについては、次項で再度説明する。

一方、デメリットとして、全てのガイダンスを結合することによって、1つのモデルとそのガイダンスを一体として行っている予報作業が難しくなる可能性がある。2018年現在の予報作業ではGSM、MSM(一部LFM)のうち1つのモデルをベースにし、気象状況の推移についてのシナリオを組み立てることによって、そのモデルで予測する気象擾乱の位置ずれ、時間ずれ、強度の違いを検討し、それを基にしたガイダンスを修正して予報が作成されているが、この修正作業が行いづらくなる。また、予報官はシナリオに基づく気象解説を行いづらくなるというデメリットもある。これらのデメリットを軽減するには、統合型ガイダンスに最も近いモデルの選択を自動的に行う等、シナリオの組み立てを補助するようなツールを検討する必要がある。

# 5.1.3 モデル更新への対応と逐次学習・固定型ガイ ダンスの今後について

ガイダンスがモデル出力を使った統計処理 (MOS)を継続する限り、モデル更新に対応する根本的な解決はモデル更新前にできるだけ多くの再予報<sup>3</sup>を行い、ガイダンスの予測式を更新することである。実際、季節予報ガイダンスでは大きなモデル更新時には再予報を作成してガイダンスの予測式を更新している。しかし、短期予報用の数値予報モデルにおいては、再予報にはデータ同化を含める必要があることから実施コストが高く、再予報のみで対処することは難しい。他国でも、モデル更新の際にガイダンスのための長期間に再予報を実施しているといった報告はない。このため、第3.3

#### TGDC CC. 2015-04-01~2017-03-31

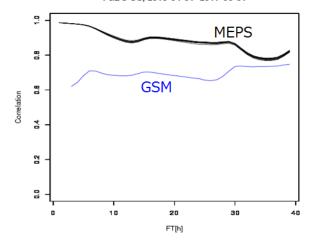


図 5.1.2 MSM 気温ガイダンスと MEPS 気温ガイダンス(各メンバー)、GSM 気温ガイダンスとの予報時間ごとの誤差の標本相関係数。青線は GSM と MSM、黒線は MEPS (各メンバー)と MSM との誤差相関。MEPS の各メンバーと MSM の誤差相関はほぼ同じのため、重なっている。初期時刻、検証期間は図 5.1.1 に同じ。

節で述べたように、モデル更新への対応は各国ともガ イダンスの重要な課題であり、逐次学習型のガイダン スの導入、カナダ気象局の UMOS の導入、ガイダンス のバイアスの逐次補正、モデルの業務化試験 4 データ の利用などが行われている。気象庁では逐次学習型ガ イダンスである KF と NN をほとんどのガイダンスに 導入することでモデル更新への対応を行ってきた。た だし、逐次学習型ガイダンスもすぐにモデル更新に対 応できるわけではないことは第3.3節で述べた通りで ある。また、KF の開発・運用コストは大きいことも 第 3.3 節で述べた。実際に KF, NN, ロジスティック回 帰 (LR)、線形重回帰 (MLR) の各種手法を使ったガイ ダンスを運用してきた経験からも、LR や MLR に比べ て KF と NN の開発・運用コストはかなり高い。今後、 統合型ガイダンスの開発やガイダンスの開発・改良を 効率よく進めてゆく場合には、この高いコストが障害 となる。ドイツ気象局や米国気象局のように多くのモ デルのガイダンス、多くの要素のガイダンスを運用し、 かつ統合型ガイダンスを作成している気象局では、KF を導入せず、全て MLR か LR のみを利用していること も押さえておくべき重要なポイントである。また、警 報・注意報向けのガイダンスの精度向上を行ってゆく には、頻度の少ない現象を対象とするため、第1.2節 で述べたように逐次的な統計処理では後追い的な予測 となってしまう。これらを考えると、気象庁でも降水 量、視程などについて、逐次学習型から一括学習型に 戻すことを検討する時期に来ていると考えられる。

<sup>3</sup> 現業運用(予定含む)と同一の数値予報モデルを用いて過去の多数事例を対象に実施する予測実験であり、ハインドキャストとも呼ばれる。

<sup>4</sup> モデルの更新や新規モデルの運用開始において、業務化を 行うかどうかを判断するために過去期間において実施する試 験

この一括学習型に戻す際に同時に検討すべきポイン トは、前項で述べた統合型ガイダンスである。統合型 ガイダンスでは、1 つのモデルに頼らないため、モデル 更新で精度が落ちるガイダンスがあっても、そのガイ ダンスの結合の重みを下げることによって、精度の低 下を最小限にすることができる。現業化前に実施した 業務化実験等のデータを合わせて、モデル更新後  $1\sim2$ 年程度経過すればガイダンスの予測式が更新できるの で、いったん劣化した新モデルのガイダンスの精度が 向上し、結合の重みを上げれば統合型ガイダンスの精 度の向上が可能である。つまり、モデル更新による悪 影響を最小限に抑えつつ、タイムラグを持ってモデル の改善を統合型ガイダンスの改善に反映させる方法で ある。実際の運用を行う場合には、結合の重みの上げ 下げ、次のモデル更新のタイミング等を検討する必要 があるが、今後のモデル更新に対応できる有効な手法 の一つとなると思われる。

なお、第 3.4 節で述べたように、逐次学習型ガイダンスは、観測所の移設・新設時の対応が比較的スムーズにできるメリットがある。頻繁にあるアメダスの移設・新設において、KF を使った気温ガイダンスは約 2 週間、風ガイダンスは数か月で自動的に移設・新設場所の特性に馴染んでゆく。この点を考えると、アメダスの観測を目的変数としている気温及び風ガイダンスでは KF が有効に働いており、KF の手法の継続が必要と思われる。英国気象局でも気温、湿度及び風ガイダンスには KF を利用している。

#### 5.1.4 新たなガイダンス技術の取り入れ

現在、深層学習と呼ばれるディープニューラルネットワーク (DNN) の技術が発展し、画像認識や音声認識で今までの機械学習技術を大きく超える高い能力が示されている。次節で紹介するように、DNN を気象予測に利用した調査・開発も行われている。ここでは、今後のガイダンスへの新たな技術の利用可能性について簡単に述べる。

まず、今まで説明変数は1点の数値予報格子点値 (GPV) または周辺の GPV を内挿または平均処理などを行って利用していたが、DNN では広範囲の GPV をそのまま説明変数とし、GPV の平面的・立体的な気圧や暖湿流等のパターンから、数値予報のバイアスを補正できる可能性がある。例えば、気温ガイダンスでは、放射冷却が起きて大外れしそうな気圧パターンや鉛直プロファイルのパターン、降水で大外れしそうな降水予測の広がりのパターンから最適な補正ができる可能性がある。つまり、今まで行うことができなかったパターン認識による層別化を導入できる可能性がある。

前述した統合型ガイダンスにおいても、多くのガイダンス・モデルの予測を入力とし、適切な組み合わせを自動的かつ高精度に行うことができる可能性がある。特にどのモデルがどのパターンが得意かを判定しなが

ら結合できれば、大きく精度向上できる可能性もある。 しかしながら、深層学習のガイダンスへの有効性については未知な部分も多く、めったに発生しない顕著現象に対してどの程度の予測精度が可能か、モデル更新に対して安定した予測が出せるか等、DNNを導入するコストに見合った有効性があるかを十分検討しておくことが必要である。

また、深層学習といった最新の機械学習だけではなく、長年にわたって他の分野で使われてきた手法も検討する必要があろう。第 2.10 節では、ランダムフォレスト、ブースティングなどの各種の機械学習技術を紹介した。今後の統合型ガイダンスやガイダンスの改良を進めるにあたって、最新の技術だけに偏らず、長年にわたって有効性が確認されている手法も検討する必要がある。

#### 参考文献

海老原智, 1999: 降水ガイダンスの改良と検証. 平成 11 年度数値予報研修テキスト, 気象庁予報部, 23-33.

永戸久喜, 原旅人, 倉橋永, 2013: 日本域拡張・高頻度 化された局地モデルの特性. 平成 25 年度数値予報研 修テキスト, 気象庁予報部, 18-41.

原旅人, 伊藤享洋, 松林健吾, 2015: asuca が導入された局地数値予報システムの特性. 平成 27 年度数値予報研修テキスト, 気象庁予報部, 9-23.

石田純一, 2017: 数値予報. 平成 29 年度数値予報研修 テキスト, 気象庁予報部, 114-117.

Veenhuis, B. A., 2013: Spread Calibration of Ensemble MOS Forecasts. *Mon. Wea. Rev.*, **141**, 2467–2482.

#### 5.2 ディープニューラルネットワーク1

#### 5.2.1 はじめに

ディープニューラルネットワーク (DNN $^2$ ) とは、多層のニューラルネットワークによる機械学習手法である。第 2.6.1 項でも述べたように、DNN は、2012 年に大規模画像認識のコンペティションである ILSVRC で Hintonらカナダのトロント大学のグループが、後に AlexNet (Krizhevsky et al. 2012) と呼ばれる DNN を用いて圧勝したことをきっかけに大きな注目を集め、2018 年現在もそのブーム $^3$  が続いている。この間、DNN の更なる多層化や新しい手法の開発などが進み、VGG (Simonyan and Zisserman 2014) や GoogLeNet (Szegedy et al. 2015), ResNet (He et al. 2016) と呼ばれるネットワークが次々に発表され、それまでの認識性能の記録を塗り替えてきている。

DNN は、画像認識の分野をはじめとして、音声認識、自然言語処理の分野で大きな成果をあげているが、近年では気象予測の分野でもその適用例が報告されている。そこで本節では、まず第 5.2.2 項で気象庁のガイダンスで用いているニューラルネットワークと DNNの違いを簡単に説明する。次に第 5.2.3 項と第 5.2.4 項で、気象予測の分野に DNN を適用した例を紹介する。最後に、第 5.2.5 項で、これら適用例の気象庁ガイダンスへの応用の可能性を検討し本項のまとめとする。なお、前提となるニューラルネットワークについては、第 2.6 節を参照されたい。

# 5.2.2 気象庁のガイダンスに用いるニューラルネットワークと DNN の違い

現在の気象庁のガイダンスに用いるニューラルネットワークと DNN の違いは、その中間層の数の違い、また、その中間層を多くしたニューラルネットを適切に学習するために必要な技術の有無の違いと言える。第2.6.1 項でも述べたように、現在の気象庁のガイダンスに用いているニューラルネットワークは、長いニューラルネットワークの歴史の中で第2世代に当たる技術を用いた3層の順伝播型ネットワーク(第2.6.4 項参照)である。第2世代ニューラルネットワークは、実差逆伝播法の再発見を機にブームが起き、気象庁のガイダンスでも数値予報モデルの更新に伴う予測特性の変化に追随することが可能になる手法として、カルマンフィルタと並んで多くのガイダンスに導入された。しかし、当時の技術で中間層を多層にしたネットワークの学習を行う場合、その学習時に勾配消失問題(第

2.6.6 項 (2) 参照) が発生してしまうために学習が適切 に行えず、画像を対象とする畳み込みニューラルネッ ト <sup>4</sup>(CNN<sup>5</sup>) といった一部の例外を除いて、ネットワー クを多層にすることが困難であった<sup>6</sup>。ここにブレイ クスルーをもたらしたのが、先に述べた Hinton の研 究である制約ボルツマンマシン (Hinton et al. 2006) と 呼ばれる単層のネットワークや、Bengio らが示した自 己符号化器 <sup>7</sup>(AE<sup>8</sup>, Bengio et al. (2006)) と呼ばれる手 法である。これらを用いることで、多層のネットワー クを入力層に近い側から順番に学習する事前学習 9 を 行い、この事前学習によって得た重みを初期値に使う ことで、DNN であっても適切に学習できることが明ら かになった。その後、前述した画像認識のコンペティ ションで Hinton らのグループが圧勝したのをきっか けに、DNN の有効性が広く知られるようになり、現在 までネットワークの多層化、手法の多様化が進んでい る。しかし、この大きなブレイクスルーをもたらした Hinton や Bengio らが提案した事前学習の手法ですら、 DNN を学習するために欠かせない技術では無くなり つつある。それは、次々と提案される新たな学習手法 やネットワークに与えるパラメータの初期設定値の与 え方によっては、より簡単に勾配消失問題に対応でき ることが分かってきたためであり、DNN の技術は日々 新しくなっていると言える。

このように、現在の気象庁ガイダンスに用いられるニューラルネットワークと DNN は、どちらもニューラルネットワークという機械学習手法に括られるものの、そのネットワークの構造や学習手法などの技術的な面での多くの違いがある。

# 5.2.3 CNN と LSTM を組み合わせた降水量予測

Shi et al. (2015) は、CNN と LSTM $^{10}$  (第 2.6.4 項 参照 )を組み合わせた畳み込み LSTM(ConvLSTM) を用いた降水量のナウキャスト手法を提案し、既存の予測手法を上回る予測精度が得られることを示した。

Shi らは、香港天文台が開発した最新の降水ナウキャストである ROVER<sup>11</sup> (Woo and Wong 2014) の問題

<sup>1</sup> 白山 洋平

 $<sup>^{2}\,</sup>$  Deep Neural Network

<sup>&</sup>lt;sup>3</sup> Hinton らの研究によって勾配消失問題に対応できるようになったことのほかに、GPU などによって計算機性能が飛躍的に向上したこと、学習に用いる大量の画像や映像データをインターネット上などから集めることが可能になってきたこともブームの要因である。

<sup>&</sup>lt;sup>4</sup> 畳み込み層とプーリング層と呼ばれる 2 種類の層を有する順伝搬型ニューラルネットワークで画像認識に用られる。

<sup>&</sup>lt;sup>5</sup> Convolutional Neural Network

 $<sup>^6</sup>$  画像を対象とした CNN については、80 年代後半に 5 層のネットワークによる予測が成功されている (LeCun et al. 1989)

<sup>&</sup>lt;sup>7</sup> 3 層ニューラルネットワークにおいて、入力層と出力層に 同じデータを用いた学習を行う手法。

<sup>&</sup>lt;sup>8</sup> autoencorder

 $<sup>^9</sup>$  まず入力層、中間層の 1 層目、仮の出力層の 3 層で出力層 と入力層が一致するように学習を行う。次に、この仮の出力 層を取り除き、元のネットワークの 2 層目の中間層を積み上げ、前回の学習した時の中間層を入力層と見なした 3 層でもう 1 度学習する、といった学習を元のネットワークの分だけ繰り返して DNN を学習する手法。

<sup>&</sup>lt;sup>10</sup> Long Short-Term Memory

<sup>&</sup>lt;sup>11</sup> Real-time Optical flow by Variational methods for

点として、大気場の流れの推定処理とレーダーエコー の外挿処理が分離しており、そのことが良い予測精度 を得るためのモデルパラメータの決定を困難にしてい ることを挙げ、この問題は機械学習の観点から解決で きる可能性があると述べている。彼らは、降水のナウ キャストは過去のレーダー画像の時系列データを入力 として、ある固定された期間の将来のレーダー画像の 時系列(予測する時刻数は通常は1より大きい)を予測 する時空間系列データの予測問題であると考え、近年 の DNN の発展、特に時系列データを適切に扱えるリカ レントニューラルネットワーク ( $RNN^{12}$ 、第 2.6.4 項参 照)や通常のRNNよりも長い記憶を扱えるLSTMが この問題に対処するために有用であると考えた。DNN によるアプローチでは、適切なモデルと学習用の十分 な量のデータが必要であるが、降水ナウキャストであ ればデータの要件は満たすため、必要なのは適切なモ デルとなる。Shi らは、モデルとして3つの先行研究 を例に挙げている。1 つは Sutskever et al. (2014) によ る入出力データがともに時系列データである問題に対 する学習のための一般的なフレームワーク、後の2つ は DNN モデルで映像データの先のフレームを予測す る研究 (Ranzato et al. 2014; Srivastava et al. 2015) で ある。ただし Shi らは、Ranzato et al. (2014) は予測 する時刻数が1時刻であったり、CNN の畳み込み計算 の範囲が制限されていて空間方向の相関の考慮が不十 分なことを、Srivastava et al. (2015) は LSTM を用い 複数時刻の予測をするものの、その計算が畳み込み計 算でないため空間方向の相関を考慮していないことを 指摘している。Shi らは、これらの先行研究を念頭に、 未来の複数時刻を予測し、かつ LSTM の計算に畳み込 み計算を導入して空間方向の相関も考慮できるように した ConvLSTM を提案した。

図 5.2.1 及び図 5.2.2 は、Shi らが提案した ConvLSTM による降水量予測の検証結果と実際 の予測例を、HKO のナウキャスト手法である ROVER や Srivastava et al. (2015) が提案した手法と比較した結果である。この ConvLSTM による実験では、モデルとして ConvLSTM ユニットを 2 層を持つ DNN を用いている。また学習用のデータには、2011 から 2013 年の 3 年間のレーダーエコーデータから上位 97 日の降水日を抽出して用いている 13 このデータは、前処理として、グレースケールへの変換、領域の切り出し、6 分毎の 240 枚/日のデータセットへと整理などを行い、訓練・評価・テストの 3 セットに分割して実験に用いられる。図 5.2.1 は、左上から各予測と観測の相関係数、0.5 mm/h のスレットスコア、誤検出率、捕捉率である。いずれも赤線で示された ConvLSTM が

Echoes of Radar

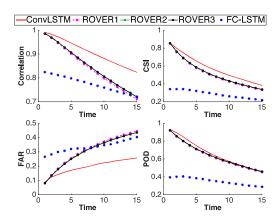


図 5.2.1 Shi et al. (2015) の Figure 7。左上から各予測と観測との相関係数、0.5 mm/3h のスレットスコア(図中は CSI) 誤検出率(図中は FAR) 捕捉率(図中は POD) 赤線は ConvLSTM による予測、青線は Srivastava et al. (2015) が提案した LSTM を用いた手法による予測、残りは既存手法である ROVER による予測で、1~3 は ROVER の設定を変更した 3 通りの実験結果を表している。

最も高精度であることを示している。Shi らは、この ConvLSTM の改善について、ConvLSTM は領域の境 界付近をうまく予測できていること、また ConvLSTM ではデータセット内の複雑な時空間パターンを学習で きることを理由に挙げている。図 5.2.2 は、実際の予測 例で上段から順に、ConvLSTM や ROVER に与える 入力のレーダー画像、入力のレーダー画像に続くレー ダー画像(つまり、予測に対する真値)、ConvLSTM による予測結果、ROVER による予測結果である。こ の予測結果から、ConvLSTM では外挿によって予測 を行う ROVER では予測し得ない領域の外側から流れ てくるエコーを予測していることが分かる。一方で、 ROVER は時間が進んでもシャープな予測をしている が、ConvLSTM はぼやけた予測になっている。Shi ら は、この ROVER のシャープな予測は多くの空振りを 引き起こし、ConvLSTM より精度が低くなる原因に なっていること、また ConvLSTM のぼやけた予測は、 予測の後半で領域全体をシャープに、また精度良く予 測することはほとんど不可能であるという、降水のナ ウキャスト固有の不確実性による誤差を緩和すること ができていると述べている。このようにShiらは、こ れまでは機械学習の恩恵を受けていてなかった降水の ナウキャストに DNN を上手く適用し、既存の予測手 法を上回る予測精度を達成した。

林ほか (2016) は、 $\mathrm{Shi}$  らの手法は過去の降水量のみを入力とした将来の降水量予測であるが、実際の降水は雲の厚さや地形、風向きといった複数の要因に依存して変化することを指摘し、過去の降水量に加えてこれら複数の要因を入力とした予測手法を提案している。林らは、この提案手法の評価のために、降水量データとして  $\mathrm{Shi}$  らと同じ気象レーダーによる観測データを、降水量以外の要因として気象衛星ひまわり 7 号の雲観測データをそれぞれ  $120 \times 120$  のピクセルのグレース

<sup>&</sup>lt;sup>12</sup> Reccurent Neural Network

<sup>13</sup> Shi らは、ナウキャストのターゲットは降水量予測である ため、無降水の日は実験には用いていない。

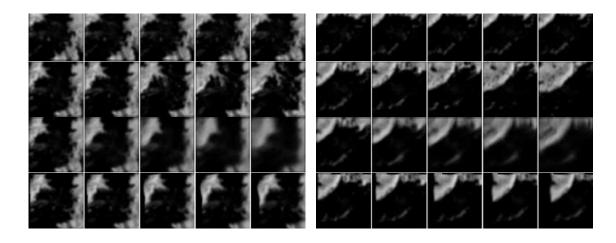


図 5.2.2 Shi et al. (2015) の Figure 8。実際の降水のナウキャスト例(左右で 2 事例分)。2 事例とも上段から順に、ConvLSTM や ROVER に与える入力のレーダー画像、入力のレーダー画像に続くレーダー画像(つまり、予測に対する真値)、ConvLSTM による予測結果、ROVER による予測結果で、それぞれ左から右に時間が進んでいる。

ケールのデータに変換、これらデータを RGB 画像のような多チャンネルのデータとして扱った ConvLSTM を構築し、Shi らの提案した過去の降水量データのみを入力データに用いた予測手法よりも精度が向上することを示している。

#### 5.2.4 気象観測データを用いた翌日の降水量予測

Hernández et al. (2016) は、AE と多層パーセプトロン  $^{14}(MLP^{15})$  を組み合わせた DNN と前日までの気象観測データを用いて翌日の日降水量予測を行い、AE を用いない MLP のみによる予測やいくつかの先行研究が提案しているニューラルネットワークによる予測精度を上回ることを示した。

彼らが構築したネットワークでは、まず用意した前日 までの降水量、気温、気圧などの気象観測データを AE を用いてよりコンパクトな表現に変換する。AEは、入 力層、中間層(入力層よりユニット数を少なくする) 出力層(入力層とユニット数を同じにする)の3層で構 成され、ネットワークの出力値が入力値に近づくよう に学習することで、真ん中の中間層に入力値から非線 形変換されたより次元の小さいコンパクトな特徴量を 抽出する手法である。この隠れ層からの出力値を MLP の入力値とすることで、MLP が過学習しにくくなり、 かつ精度が改善されると Hernandez らは述べている。 なお、これら AE や MLP の学習回数や学習率、それ ぞれの隠れ層のユニット数などのパラメータは、いく つかの候補値の組み合わせを総当たりすることで最適 な設定を決定している。また、Hernandezらはこの構 築したネットワークの実際の予測例を示し、大雨の日 は非常に良い精度で予測できている一方で、日降水量 が20 mm 未満の日は予測が正確ではなく、これら雨量 の少ない日の予測精度の改善が今後の課題であると述べている。

#### 5.2.5 まとめ

本項では、はじめに気象庁のガイダンスに用いているニューラルネットワークと DNN の違いについて説明し、さらに DNN の気象予測の分野への応用例をいくつか紹介した。

第5.2.3 項で初めに紹介した、Shi らの過去の降水量 のみを入力とした ConvLSTM は、そのままでは降水の ナウキャスト手法であるため本報告が対象とするガイ ダンス手法ではない。また、ナウキャスト手法として見 ても、予測時間の後半ではぼやけた降水予測となって おり、このままでは大雨の予測といった防災目的への利 用は難しいと考えられる。 林らが提案した ConvLSTM に複数の要因を取り込む手法はガイダンスへの応用が 期待できる。例えば、入力の多チャンネルデータに数 値予報モデルの降水量や高度場のデータを、教師デー タに降水量ガイダンスの目的変数を与えたネットワー クを学習することで、LSTM によって対象時刻までの 入力データを、CNN によって周囲の格子の気象要素 を考慮した統計的な補正を行うニューラルネットワー クの作成が行えると考えられ、従来にはない新しいア プローチのガイダンス作成手法となりうるだろう。ま た、第5.2.4 項で紹介した、気象観測データのみを用い た翌日の降水量予測は、一般的には数値予報を用いた 場合に比べて予測精度は劣るものと想像されるが、多 様な入力データからデータの特徴を良く表す量(特徴 量)を効率的に抽出しており、この手法はニューラル ネットワークを用いたガイダンスに限らず、多くの既 存のガイダンスにも応用可能な技術であると考えられ る。このように、DNN はガイダンス開発に新たなアプ ローチを与えるとともに、第2.6.9 項でも述べたよう に、現在気象庁が運用しているガイダンスに、DNN に 用いられる新しい技術を取り込んでいくことで、気象

 $<sup>^{14}</sup>$  順伝播型ニューラルネットワークと同じだが、本節では Hernández et al. (2016) にあわせて多層パーセプトロンと 記す。

 $<sup>^{15}\,</sup>$  multilayer perceptron

庁ガイダンスの予測精度の向上にも資すると言えるだろう。

第 5.2.1 項や第 5.2.2 項で述べた通り、DNN の技術は日進月歩であり、新しい技術が次々に発表されている。引き続き最新の DNN の情報収集に努め、既存のガイダンスの精度向上に繋がる技術開発に結びつけていきたい。

#### 参考文献

- Bengio, Y., P. Lamblin, D. Popovici, and H. Larochelle, 2006: Greedy Layer-Wise Training of Deep Networks. Proceedings of the 19th International Conference on Neural Informatin Processing Systems, 153–160.
- 林政行, R. Hadfi, 伊藤孝行, 2016: 深層学習を用いた マルチモーダル学習による降水量予測. 情報処理学 会第 78 回全国大会論文集, 2016(1), 369-370.
- He, K., X. Zhang, S. Ren, and J. Sun, 2016: Deep Residual Learning for Image Recognition. The IEEE Conference on Computer Vision and Pattern Recognition, 770–778.
- Hernández, E., V. Sanchez-Anguix, V. Julian, J. Palanca, and N. Duque, 2016: Rainfall prediction: A Deep Learning approach. Proceedings of the 11th International Conference on Hybrid Artificial Inteligent Systems, 151–162.
- Hinton, G.E., S. Osindero, and Y.W. Teh, 2006: A fast learning algorithm for deep belief nets. *Neural Computation*, 18, 1527–1554.
- Krizhevsky, A., I. Sutskever, and G.H. Hinton, 2012: ImageNet Classification with Deep Convolutional Neural Networks. Proceedings of the 25th International Conference on Neural Informatin Processing Systems, 1097–1105.
- LeCun, Y., B. Boser, J.S. Denker, D. Henderson,
  R.E. Howard, W. Hubbard, and L.D. Jackel, 1989:
  Backpropagation Applied to Handwritten Zip Code
  Recognition. Neural Computation, 1, 541–551.
- Ranzato, M., A. Szlam, J. Bruna, M. Mathieu, R. Collobert, and S. Chopra, 2014: VIDEO(LANGUAGE) MODELING: A BASE-LINE FOR GENERATIVE MODELS OF NATURAL VIDEOS. arXiv preprint arXiv:1412.6604.
- Shi, X., Z. Chen, H. Wang, D. Yeung, W. Wong, and W. Woo, 2015: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. arXiv preprint arXiv:1506.04214.
- Simonyan, K. and A. Zisserman, 2014: VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION. arXiv preprint

- arXiv:1409.1556.
- Srivastava, N., E. Mansimov, and R. Salakhudinov, 2015: Unsupervised Learning of Video Representations using LSTMs. *Proceedings of the 32nd In*ternational Conference on Machine Learning, **37**, 843–852.
- Sutskever, I., O. Vinyals, and Q. V. Le, 2014: Sequence to Sequence Learning with Neural Networks. Proceedings of the 27th International Conference on Neural Informatin Processing Systems, 2, 3104–3112.
- Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, 2015: Going deeper with convolutions. The IEEE Conference on Computer Vision and Pattern Recognition, 1–9.
- Woo, W.C. and W.K. Wong, 2014: Application of Optical Flow Techniques to Rainfall Nowcasting. the 27th Conference on Severe Local Storms.

#### 5.3 ガイダンスグループ1

2018年現在、気象庁内のさまざまな課室や施設等機関で統計や機械学習を利用したプロダクトの開発や調査が行われている。また、ディープニューラルネットワークなどの新しい技術の登場により、気象庁でも新たな手法を利用したガイダンスの開発が求められている。そこで、各課室が持つツールやノウハウを共有することによってプロダクトの開発と改良を促進することにに、新しい技術を互いに習得することによってプロダクトの開発を進めることを目的として、平成29年度より気象庁技術開発推進本部の実況監視・予測技術開発部会の下にガイダンスグループが立ち上げられた。過去にも、予報部内に平成9~16年にガイダンスワーキンググループが立ち上げられ、ガイダンスの改良・開発の促進に寄与したが、ガイダンスグループは庁内を横断したグループである点が異なる。

平成 29 年度の主な参加メンバーは、表 5.3.1 となっている。ガイダンスグループへの付託事項は、

- ガイダンス等統計的プロダクトの作成手法とノウ ハウを共有し各課プロダクトの高度化を推進する こと
- 機械学習や人工知能などの最新技術の動向を調査 し業務への利用可能性を検討すること

となっており、前者に関しては、

- 1. 各メンバーの担当プロダクトで使っている統計・ 機械学習技術の共有
- 2. 統計・機械学習技術の基礎知識の習得
- 3. 関連ツールの紹介
- 4. 外国の論文やプロダクトの紹介
- 5. 担当プロダクトの改良調査の報告

などを行っている。特にメソアンサンブルや全球アン サンブルのアンサンブル予報を使った応用技術の高度

表 5.3.1 ガイダンスグループ構成メンバー(平成 29 年度)。 以下のメンバーの他に情報通信課及び気象大学校からも 参加。

4445 4 5 119	BB)++D \L\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
参加メンバー	関連担当業務
数値予報課・アプリケー ション班	天気予報ガイダンス、航空 気象予報ガイダンス、航空 悪天 GPV、週間予報ガイ ダンス
予報課・週間予報係	週間予報ガイダンス
気候情報課・季節予報係	季節予報ガイダンス
予報課・アジア太平洋気象 防災センター	台風予報ガイダンス
予報課・開発班	雷ナウキャスト
航空予報室	航空気象予報プロダクト
観測システム運用室	推計気象分布
海洋気象情報室	海氷解析
気象衛星センター	積乱雲情報プロダクト

<sup>1</sup> 高田 伸一

化に力点を置いている。

後者に関しては、ディープニューラルネットワークを 使った調査結果の共有及びツールの共有を行っており、 実際に業務への利用調査も実施している。また、若手 を中心に基礎的な理論と技術の習得を目的とした活動 も行っている。今後は、部外の機械学習の専門家との 共同研究などの連携を行ってゆくことを検討している。

# 付録 A 数値予報課報告・別冊で用いた表記と統計的検証に用いる代表的な指標\*

本報告で使用した表記と統計的検証に用いる代表的な指標などについて以下に説明する。

#### A.1 数値予報課報告・別冊で用いた表記

#### A.1.1 時刻の表記について

本報告では、時刻を表記する際に、通常国内で用いられている日本標準時 (JST: Japan Standard Time) のほかに、協定世界時 (UTC: Coordinated Universal Time) を用いている。数値予報では国際的な観測データの交換やプロダクトの利用等の利便を考慮して、時刻は UTC で表記されることが多い。JST は UTC に対して 9 時間進んでいる。

#### A.1.2 予測時間の表記について

数値予報では、統計的な検証や事例検証の結果を示す際に、予報対象時刻のほかに、初期時刻からの経過時間を予報時間  $(FT: Forecast\ Time^1)$  として表記している。

#### 本報告では、予報時間を

「予報時間」=「予報対象時刻」-「初期時刻」で定義し、例えば、6時間予報の場合、FT=6と表記しており、時間の単位[h]を省略している。

#### A.1.3 アンサンブル予報の表記について

アンサンブル予報では、複数の予測の集合(アンサンブル)を統計的に処理し、確率予測等の資料を作成する。本報告では、予測の集合の平均を「アンサンブル平均」、個々の予測を「メンバー」と呼ぶ。また、摂動を加えているメンバーを「摂動ラン」、摂動を加えていないメンバーを「コントロールラン」と呼ぶ。

#### A.1.4 緯度、経度の表記について

本報告では、緯度、経度について、アルファベットを用いて例えば「北緯 40 度、東経 130 度」を「40°N, 130°E」、「南緯 40 度、西経 130 度」を「40°S, 130°W」などと略記する。

## A.2 統計的検証に用いる代表的な指標

# A.2.1 平均誤差、二乗平均平方根誤差、誤差の標準 偏差

予測誤差を表す基本的な指標として、平均誤差 (ME: Mean Error、バイアスと表記する場合もある)と二乗平均平方根誤差 (RMSE: Root Mean Square Error) がある。これらは次式で定義される。

$$ME \equiv \frac{1}{N} \sum_{i=1}^{N} (x_i - a_i)$$
 (A.2.1)

RMSE 
$$\equiv \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - a_i)^2}$$
 (A.2.2)

ここで、N は標本数、 $x_i$  は予測値、 $a_i$  は実況値である。 ME は予測値の実況値からの偏りの平均であり、0 に近いほど実況からのずれが小さいことを示す。 RMSE は最小値の 0 に近いほど予測が実況に近いことを示す。

RMSE は ME の寄与とそれ以外を分離して、

$$RMSE^2 = ME^2 + \sigma_e^2 \tag{A.2.3}$$

$$\sigma_e^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - a_i - ME)^2$$
 (A.2.4)

と表すことができる。 $\sigma_e$  は誤差の標準偏差である。

#### A.3 カテゴリー検証で用いる指標

カテゴリー検証では、まず、対象となる現象の有無を予測と実況それぞれについて判定し、その結果により標本を分類する。そして、それぞれのカテゴリーに分類された事例数を基に、予測の特性を検証するという手順を踏む。

#### A.3.1 分割表

分割表は、カテゴリー検証においてそれぞれのカテゴリーに分類された事例数を示す表(表 A.3.1)である。付録 A.3.2 から A.3.12 に示す各スコアは、表 A.3.1 に示される各区分の事例数を用いて定義される。また、以下では全事例数を N=FO+FX+XO+XX、実況「現象あり」の事例数を M=FO+XO、実況「現象なし」の事例数を X=FX+XX と表す。

表 A.3.1 カテゴリー検証で用いる分割表。FO, FX, XO, XX はそれぞれの事例数を示す。

10. C1. C1. C1. C1. C1. C1. C1. C1. C1. C1					
		実	実況		
		あり	なし	計	
予測	あり	適中 (FO)	空振り (FX)	FO+FX	
]. \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	なし	見逃し (XO)	適中 (XX)	XO+XX	
計		M	X	N	

#### A.3.2 適中率

適中率は、予測が適中した割合であり、次式で定義される。

適中率 
$$\equiv \frac{\text{FO} + \text{XX}}{N} \quad (0 \leq$$
 適中率  $\leq 1) \quad (\text{A.3.1})$ 

最大値の1に近いほど予測の精度が高いことを示す。

#### A.3.3 空振り率

空振り率は、予測「現象あり」の事例数に対する空振り(予測「現象あり」かつ実況「現象なし」)の割合

<sup>\*</sup> 小林 大輝

<sup>&</sup>lt;sup>1</sup> Forecast Range などと記述されることも多い。

であり、次式で定義される。

空振り率 
$$\equiv \frac{\mathrm{FX}}{\mathrm{FO} + \mathrm{FX}} \quad (0 \leq 空振り率 \leq 1) \ (\mathrm{A.3.2})$$

最小値の0に近いほど空振りが少ないことを示す。 本報告では分母をFO+FXとしているが、代わりにNとして定義する場合もある。

#### A.3.4 見逃し率

見逃し率は、実況「現象あり」の事例数に対する見逃し(実況「現象あり」かつ予測「現象なし」)の割合であり、次式で定義される。

見逃し率 
$$\equiv \frac{\mathrm{XO}}{M} \quad (0 \leq 見逃し率 \leq 1) \qquad (A.3.3)$$

最小値の0に近いほど見逃しが少ないことを示す。 本報告では分母をMとしているが、代わりにNとして定義する場合もある。

#### A.3.5 捕捉率

捕捉率  $(H_r: \text{Hit Rate})$  は、実況「現象あり」のときに予測が適中した割合であり、次式で定義される。

$$H_r \equiv \frac{\text{FO}}{M} \quad (0 \le H_r \le 1) \tag{A.3.4}$$

最大値 1 に近いほど見逃しが少ないことを示す。捕捉率は、ROC 曲線(付録 A.4.5)のプロットに用いられる。

#### A.3.6 体積率

体積率 ( $V_r$ : Volume Ratio) は、全事例のうち予測の「現象あり」の事例の割合を示す。

$$V_r \equiv \frac{\text{FO} + \text{FX}}{N} \tag{A.3.5}$$

複数の予測の捕捉率が等しい場合、体積率が小さい予 測ほど空振りが少ないよい予測と言える。

#### A.3.7 誤検出率

誤検出率  $(F_r$ : False Alarm Rate) は、実況「現象なし」のときに予測が外れた割合である。空振り率 (A.3.3)とは分母が異なり、次式で定義される。

$$F_r \equiv \frac{\text{FX}}{X} \quad (0 \le F_r \le 1) \tag{A.3.6}$$

最小値の0に近いほど、空振りが少なく予測の精度が高いことを示す。誤検出率は捕捉率(付録 A.3.5)とともにROC 曲線(付録 A.4.5)のプロットに用いられる。

#### A.3.8 バイアススコア

バイアススコア (BI: Bias Score) は、実況「現象あり」の事例数に対する予測「現象あり」の事例数の比であり、次式で定義される。

$$\mathrm{BI} \equiv \frac{\mathrm{FO} + \mathrm{FX}}{M} \quad (0 \leq \mathrm{BI}) \tag{A.3.7}$$

予測と実況で「現象あり」の事例数が一致する場合に1となる。1より大きいほど予測の「現象あり」の頻度が過大、1より小さいほど予測の「現象あり」の頻度が過小であることを示す。

#### A.3.9 気候学的出現率

現象の気候学的出現率  $P_c$  は、標本から見積もられる現象の平均的な出現確率であり、次式で定義される。

$$P_c \equiv \frac{M}{N} \quad (0 \le P_c \le 1) \tag{A.3.8}$$

この量は実況のみから決まり、予測の精度にはよらない。予測の精度を評価する際の基準値の設定にしば しば用いられる。

#### A.3.10 スレットスコア

スレットスコア (TS: Threat Score) は、予測または 実況で「現象あり」の場合の予測適中事例数に着目し て予測精度を評価する指標であり、次式で定義される。

$$TS \equiv \frac{FO}{FO + FX + XO} \quad (0 \le TS \le 1) \qquad (A.3.9)$$

出現頻度の低い現象( $N\gg M$ 、したがって、 $XX\gg FO$ , FX, XO となって、予測「現象なし」による寄与だけで適中率が1 に近い現象)について XX の影響を除いて検証するのに有効である。本スコアは最大値の1 に近いほど予測の精度が高いことを示す。なお、スレットスコアは現象の気候学的出現率の影響を受けやすく、異なる標本や出現率の異なる現象に対する予測の精度を比較するのには適さない。この問題を緩和するため、次項のエクイタブルスレットスコアなどが考案されている。

#### A.3.11 エクイタブルスレットスコア

エクイタブルスレットスコア (ETS: Equitable Threat Score) は、前項のスレットスコアが現象の気候学的出現率の影響を受けやすいため、気候学的な確率で「現象あり」が適中した頻度を除いて求めたスレットスコアであり、次式で定義される (Schaefer 1990)。

$$ETS \equiv \frac{FO - S_f}{FO + FX + XO - S_f} \quad \left(-\frac{1}{3} \le ETS \le 1\right)$$
(A.3.10)

ただし、

$$S_f = P_c(FO + FX) \tag{A.3.11}$$

である。ここで、 $S_f$  は「現象あり」をランダムに  ${
m FO+FX}$  回予測した場合(ランダム予測)の「現象あり」の適中事例数である。本スコアは、最大値の1 に近いほど予測の精度が高いことを示す。また、ランダム予測で0となり、 ${
m FO=XX=0}$ , ${
m FX=XO=}N/2$  の場合に最小値 -1/3 をとる。

#### A.3.12 スキルスコア

スキルスコア (Skill Score) は気候学的確率などによる予測の難易を取り除いて、予測の技術力を評価する指数であり、一般に次式のように定義される。

スキルスコア 
$$\equiv \frac{S_{fcst} - S_{ref}}{S_{pfct} - S_{ref}}$$
 (A.3.12)

ここで、 $S_{fcst}$ ,  $S_{pfct}$ ,  $S_{ref}$  は、評価対象の予測・完全予測・比較の基準となる予測(気候学的確率など)の各スコア(適中率)である。本スコアは、最大値の1 に近いほど予測の精度が高いことを示し、比較の基準となる予測よりも精度が劣る場合、負の値となる。

代表的なスキルスコアは Heidke のスキルスコア (HSS: Heidke Skill Score) で、気候学的な確率で「現象あり」および「現象なし」が適中した頻度を除いて求める適中率であり、次式で定義される。

$${\rm HSS} \equiv \frac{{\rm FO} + {\rm XX} - S}{N-S} \quad (-1 \leq {\rm HSS} \leq 1) \ ({\rm A.3.13})$$

ただし、

$$S = P_c(\text{FO} + \text{FX}) + P_x(\text{XO} + \text{XX}),$$
 
$$P_c = \frac{M}{N}, \quad P_x = \frac{X}{N}$$
 (A.3.14)

である。ここで、 $P_c$  は「現象あり」、 $P_x$  は「現象なし」の気候学的出現率、S は「現象あり」を FO+FX回(すなわち、「現象なし」を残りの XO+XX回)ランダムに予測した場合(ランダム予測)の適中事例数である。HSS は、最大値 1 に近づくほど精度が高く、ランダム予測で 0 となり、FO=XX=0,FX=XO=N/2の場合に最小値 -1 をとる。前項のエクイタブルスレットスコアもスキルスコアの一つで、Gilbert Skill Score とも呼ばれている。

#### A.4 確率予測に関する指標など

# A.4.1 ブライアスコア

プライアスコア (BS: Brier Score) は、確率予測の統計検証の基本的指標である。ある現象の出現確率を対象とする予測について、次式で定義される。

BS 
$$\equiv \frac{1}{N} \sum_{i=1}^{N} (p_i - a_i)^2 \quad (0 \le BS \le 1) \quad (A.4.1)$$

ここで、 $p_i$  は確率予測値 (0 から 1 )、 $a_i$  は実況値 (現象ありで 1、なしで 0 )、N は標本数である。BS は完全に適中する決定論的な  $(p_i=0$  または 1 の ) 予測 (完全予測と呼ばれる)で最小値の 0 をとり、0 に近いほど予測の精度が高いことを示す。また、現象の気候学的出現率  $P_c$  を常に確率予測値とする予測(気候値予測と呼ばれる)のブライアスコア  $BS_c$  は

$$BS_c \equiv P_c(1 - P_c) \tag{A.4.2}$$

となる。ブライアスコアは、現象の気候学的出現率の影響を受けるため、異なる標本や出現率の異なる現象に対する予測の精度を比較するのには適さない。例えば上の  $BS_c$  は  $P_c$  依存性を持ち、同じ予測手法(ここでは気候値予測)に対しても  $P_c$  の値に応じて異なる値をとる (Stanski et al. 1989)。この問題を緩和するため、次項のブライアスキルスコアが考案されている。

#### A.4.2 ブライアスキルスコア

ブライアスキルスコア (BSS: Brier Skill Score) は、ブライアスコアに基づくスキルスコアであり、通常気候値予測を基準とした予測の改善の度合いを示す。本スコアは、ブライアスコア BS、気候値予測によるブライアスコア  $\mathrm{BS}_{\mathrm{c}}$  を用いて

$$BSS \equiv \frac{BS_c - BS}{BS_c} \quad (BSS \le 1)$$
 (A.4.3)

で定義され、完全予測で1、気候値予測で0、気候値予 測より誤差が大きいと負となる。

#### A.4.3 Murphy の分解

Murphy (1973) は、ブライアスコアと予測の特性との関連を理解しやすくするため、ブライアスコアを信頼度 (Reliability)、分離度 (Resolution)、不確実性 (Uncertainty) の 3 つの頃に分解した。これを Murphy の分解と呼ぶ(高野 2002 などに詳しい)。

確率予測において、確率予測値を L 個の区間に分け、標本を確率予測値の属する区間に応じて分類することを考える。確率予測値が l 番目の区間に属する標本数を  $N_l$   $(N=\sum_{l=1}^L N_l)$ 、このうち実況が「現象あり」であった事例数を  $M_l$   $(M=\sum_{l=1}^L M_l)$ 、確率予測値の l 番目の区間の区間代表値を  $p_l$  とすると、Murphy の分解によりブライアスコアは以下のように表される。

$$BS =$$
信頼度  $-$  分離度  $+$  不確実性  $(A.4.4)$ 

信頼度 = 
$$\sum_{l=1}^{L} \left( p_l - \frac{M_l}{N_l} \right)^2 \frac{N_l}{N}$$
 (A.4.5)

分離度 = 
$$\sum_{l=1}^{L} \left( \frac{M}{N} - \frac{M_l}{N_l} \right)^2 \frac{N_l}{N}$$
 (A.4.6)

不確実性 = 
$$\frac{M}{N} \left( 1 - \frac{M}{N} \right)$$
 (A.4.7)

信頼度は、確率予測値  $(p_l)$  と実況での現象の出現相対頻度  $(M_l/N_l)$  が一致すれば最小値の 0 となる。分離度は、確率予測値に対応する実況での現象の出現相対頻度  $(M_l/N_l)$  が気候学的出現率  $(P_c=M/N)$  から離れているほど大きい値をとる。不確実性は、現象の気候学的出現率のみによって決まり、予測の手法にはよらない。例えば、 $P_c=0.5$  の場合に不確実性は最大値の 0.25 をとる。また、不確実性=BS $_c$  が成り立つ。これらを用いて、ブライアスキルスコアを次のように書

くことができる。

$$BSS = \frac{\text{$\beta$ mig - finite}}{\text{$\gamma$ mig the finite of the$$

#### A.4.4 確率值別出現率図

確率値別出現率図(Reliability Diagram, Attributes Diagram とも呼ばれる)は、予測された現象出現確率  $P_{\rm fest}$  を横軸に、実況で現象が出現した相対頻度  $P_{\rm obs}$  を縦軸にとり、確率予測の特性を示した図である(図 A.4.1 参照、Wilks 2011 などに詳しい)。一般に、確率予測の特性は確率値別出現率図上で曲線として表される。この曲線を信頼度曲線 (Reliability curve) と呼ぶ。

信頼度曲線の特性は、Murphy の分解(付録 A.4.3)の信頼度、分離度と関連付けることができる。横軸  $P_{\rm fcst}$  の各値について、信頼度(あるいは分離度)への寄与は、信頼度曲線上の点から対角線  $P_{\rm obs}=P_{\rm fcst}$  上の点(あるいは直線  $P_{\rm fcst}=P_c$  上の点)までの距離の二乗として表現される。 $P_{\rm fcst}$  の各値でのこれらの寄与を、標本数に比例する重みで平均して信頼度(あるいは分離度)が得られる。例えば、no-skill line(直線  $P_{\rm obs}=(P_{\rm fcst}+P_c)/2$ )上の点では、信頼度と分離度への寄与は等しい大きさを持ち、ブライアスキルスコアへの寄与が 0 となる。また no-skill line と直線  $P_{\rm fcst}=P_c$  との間の領域(分離度への寄与 > 信頼度への寄与、図 A.4.1 灰色の領域)内に位置する点は、ブライアスキルスコアに正の寄与を持つ。

特別な場合として、気候値予測(付録 A.4.1)では 1 点  $(P_{\rm fcst},P_{\rm obs})=(P_c,P_c)$  が信頼度曲線に対応する。 また、次の 2 つの特性を示す確率予測は精度が高い。

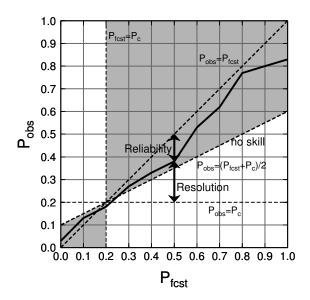


図 A.4.1 確率値別出現率図の模式図。横軸は予測現象出現確率、縦軸は実況現象出現相対頻度、実線が信頼度曲線である。対角線、直線  $P_{\rm obs}=P_c$  との差の二乗がそれぞれ信頼度 (Reliability)、分離度 (Resolution) への寄与に対応している。灰色の領域内の点はブライアスキルスコアに正の寄与を持つ。

- 信頼度曲線が対角線に(信頼度への寄与が最小値の0に)近い。
- 信頼度曲線上の大きい標本数に対応する点が点  $(P_{\mathrm{fcst}},P_{\mathrm{obs}})=(P_c,P_c)$  (気候値予測)から離れた位置(確率値別出現率図の左下または右上寄り)に分布する(分離度が大きい)。

#### A.4.5 ROC 曲線、ROC 面積、ROC 面積スキルス コア

現象の予測出現確率にある閾値を設定し、これを予 測の「現象あり」「現象なし」を判定する基準とするこ とが可能である。様々な閾値それぞれについて作成し た分割表を基に、閾値が変化したときの $F_r$ - $H_r$ 平面上 の軌跡をプロットしたものが ROC 曲線 (ROC curve: Relative Operating Characteristic curve、相対作用特 性曲線)である(図 A.4.2 参照、高野 2002 などに詳し い)。 平面内の左上方の領域では  $H_r > F_r$  であり、平面 の左上側に膨らんだ ROC 曲線特性を持つ確率予測ほど 精度が高いものと見なせる。 したがって、ROC 曲線から 下の領域(図 A.4.2 灰色の領域)の面積(ROCA: ROC area、ROC 面積)は、情報価値の高い確率予測ほど大 きくなる。ROC 面積スキルスコア (ROCASS: ROC Area Skill Score) は、情報価値のない予測  $(H_r = F_r)$ を基準として ROC 面積を評価するものであり、次式 で定義される。

$$ROCASS \equiv 2(ROCA - 0.5) \quad (-1 \le ROCASS \le 1)$$

$$(A.4.9)$$

本スコアは、完全予測で最大値の1 をとる。また、情報価値のない予測(例えば、区間[0,1] から一様ランダムに抽出した値を確率予測値とする予測など)では

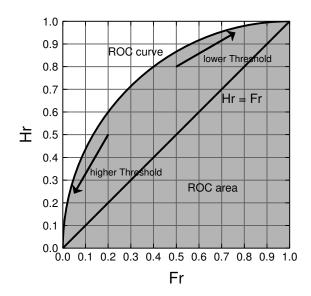


図 A.4.2 ROC 曲線の模式図。横軸は  $F_r$ 、縦軸は  $H_r$  である。灰色の領域の面積が ROC 面積である。

0となる。

#### 参考文献

- Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600.
- Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. Wea. Forecasting, 5, 570–575.
- Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of common verification methods in meteorology. Research Rep., 89-5, Forecast Research Division, Atmospheric Envirnment Service, Environment Canada, 114 pp.
- 高野清治, 2002: アンサンブル予報の利用技術. 気象研究ノート, **201**, 73-103.
- Wilks, D. S., 2011: Statistical Methods in the Atmospheric Sciences, International Geophysics, Vol. 100. Academic Press, 334-340 pp.

付録 B 電子計算室報告、同別冊、数値予報課報告·別冊 発行履歴

 発行年月	発行号	表題
2018年(平成30年)3月	数値予報課報告・別冊第 64 号	ガイダンスの解説
2017年(平成29年)3月	数値予報課報告・別冊第 63 号	数値予報モデル開発のための基盤整備および開発 管理
2016年(平成28年)3月	数値予報課報告・別冊第 62 号	確率的な気象予測のためのアンサンブル予報の課 題と展望
2015年(平成27年)3月	数値予報課報告・別冊第 61 号	観測データ利用の現状と課題
2014年(平成26年)3月	数値予報課報告・別冊第 60 号	次世代非静力学モデル asuca
2013年(平成25年)3月	数値予報課報告・別冊第 59 号	物理過程の改善に向けて (II)
2012年(平成24年)3月	数値予報課報告・別冊第 58 号	物理過程の改善に向けて (I)
2011年(平成23年)3月	数値予報課報告・別冊第 57 号	データ同化の改善に向けて
2010年(平成22年)3月	数値予報課報告・別冊第 56 号	非静力学メソ4次元変分法
2009年(平成21年)3月	数値予報課報告・別冊第 55 号	全球モデルの課題と展望
2008年(平成20年)3月	数値予報課報告・別冊第 54 号	気象庁非静力学モデル II
		──現業利用の開始とその後の発展──
2007年(平成19年)3月	数値予報課報告・別冊第 53 号	数値予報と衛星データ ――同化の現状と課題――
2006年(平成18年)3月	数値予報課報告・別冊第 52 号	アンサンブル技術の短期・中期予報への利用
		──激しい気象現象の予測向上を目指して──
2005年(平成17年)3月	数値予報課報告・別冊第 51 号	全球モデル開発プロジェクト (II)
2004年(平成16年)3月	数値予報課報告・別冊第 50 号	全球モデル開発プロジェクト (I)
2003年(平成15年)3月	数値予報課報告・別冊第 49 号	気象庁非静力学モデル
2002年(平成14年)3月	数値予報課報告・別冊第 48 号	変分法データ同化システムの現業化
2000年(平成12年)10月	数値予報課報告・別冊第 47 号	新しい数値解析予報システム(数値予報解説資料 (33) 平成 12 年度数値予報研修テキスト合併)
2000年(平成12年)3月	数値予報課報告・別冊第 46 号	全球モデル開発の現状と展望 ――気象業務の基幹 モデルとして――
1999年(平成11年)3月	数値予報課報告・別冊第 45 号	数値予報のための衛星データ同化
1998年(平成10年)3月	数値予報課報告・別冊第 44 号	メソ数値予報の実用化に向けて
1997年(平成 9年)3月	数値予報課報告・別冊第 43 号	データ同化の現状と展望
1996年(平成 8年)3月	数値予報課報告・別冊第 42 号	ーヶ月予報に向けた全球モデルの開発 ──バイア スの小さな予報モデルを目指して──
1994年(平成 6年)9月	数値予報課報告・別冊第 41 号	数値予報の実際(数値予報解説資料(27)平成6年 度数値予報研修テキスト合併)
1994年(平成 6年)3月	数値予報課報告・別冊第 40 号	気候監視のための海洋データ同化システム ――大 気海洋結合モデルによる季節予報に向けて――
1993年(平成 5年)3月	数値予報課報告・別冊第 39 号	数値予報とリモートセンシング
1992年(平成 4年)3月	数値予報課報告・別冊第 38 号	力学的 1ヶ月予報の課題と展望
1991年(平成 3年)3月	数値予報課報告・別冊第 37 号	狭領域モデルの課題と展望
1990年(平成 2年)3月	数値予報課報告・別冊第 36 号	気象データと客観解析
1989年(平成 元年)3月	数値予報課報告・別冊第 35 号	力学的長期予報をめざして
1988年(昭和63年)3月	数値予報課報告・別冊第 34 号	数値予報モデルの物理過程
1987年(昭和62年)3月	数値予報課報告・別冊第 33 号	低緯度の数値予報
1986年(昭和61年)3月	数値予報課報告・別冊第 32 号	メソスケール現象と数値予報
1985年(昭和60年)3月	電子計算室報告・別冊第 31 号	延長予報に関する最近の話題
1984年(昭和59年)3月	電子計算室報告・別冊第 30 号	ノーマル・モード・イニシャリゼーション

発行年月	発行号	表題
1983年(昭和58年)3月	電子計算室報告・別冊第 29 号	北半球およびファインメッシュ予報モデル(8L NHM および 10L FLM)と解析システム
1982年(昭和57年)3月	電子計算室報告・別冊第 28 号	スペクトル法による数値予報(その原理と実際)
1981年(昭和56年)3月	電子計算室報告・別冊第 27 号	数値予報モデルの時間差分スキームと物理過程
1980年(昭和55年)3月	電子計算室報告・別冊第 26 号	気象衛星資料と数値予報
1979年(昭和54年)3月	電子計算室報告・別冊第 25 号	4 層北半球プリミティブ・モデルの改良について
1978年(昭和53年)3月	電子計算室報告・別冊第 24 号	数値予報による延長予報
1977年(昭和52年)3月	電子計算室報告・別冊第 23 号	数値予報と天気予報
1976年(昭和51年)3月	電子計算室報告・別冊第 22 号	客観解析
1975年(昭和50年)3月	電子計算室報告・別冊第 21 号	4 層北半球プリミティブ・モデルについて
1974年(昭和49年)3月	電子計算室報告・別冊第 20 号	数値予報特別研修のまとめ
1973年(昭和48年)10月	電子計算室報告・別冊第 19 号	プリミティブ・モデルについて (数値予報解説資料 (6) 合併 )
1973年(昭和48年)3月	電子計算室報告・別冊第 18 号	プリミティブ・モデルをめぐって
1972年(昭和47年)10月	電子計算室報告別冊第 17 号	新しく予報を担当される方のための電計資料の見方(数値予報解説資料(5)合併)
1971年(昭和46年)10月	電子計算室報告別冊第 16 号	じょう乱の構造について (数値予報解説資料 (4) 合 併 )
1971年(昭和46年)9月	電子計算室報告別冊第 15 号	中間規模じょう乱をめぐって
1970年(昭和45年)11月	電子計算室報告別冊第 14 号	北半球 3 層非地衡風バランス・モデル (数値予報解説資料 (3) 合併)
1969年(昭和 44年)10月	電子計算室報告別冊第 13 号	北半球 3 層非地衡風バランス・モデル (数値予報解説資料 (2) 合併)
1969年(昭和44年)9月	電子計算室報告別冊第 12 号	数値予報のはじめ(数値予報解説資料(1)合併)
1968年(昭和43年)10月	電子計算室報告別冊第 11 号	予報技術改善の方向
1968年(昭和43年)3月	電子計算室報告別冊第 10 号	数值予報
1966年(昭和41年)10月	電子計算室報告別冊第9号	北半球4層傾圧予報について
1965年(昭和40年)11月	電子計算室報告別冊第8号	IUGG 大気科学委員会第一回活動概要報告
1964年(昭和39年)2月	電子計算室報告 VIII	
1963年(昭和38年)6月	電子計算室報告別冊第7号	バロクリニツク大気の性質
1963年(昭和38年)6月	電子計算室報告別冊第6号	アジア地区のバロクリニツク予報
1962年(昭和37年)7月	電子計算室報告別冊第5号	北半球バロトロピツク予報
1962年(昭和 37年) 6月	電子計算室報告別冊第4号	気象庁電子計算室におけるルーチン傾圧モデルの 概要
1961年(昭和36年)5月	電子計算室報告 別冊 No.3	500MB 面渦度及びその予報図の利用法
1961年(昭和36年)3月	電子計算室報告 VI VII	
1960年(昭和35年)8月	電子計算室報告 別冊 No.2	機械でつくる天気図について
1960年(昭和35年)7月	電子計算室報告 V	
1960年(昭和35年)5月	電子計算室報告 別冊 No.1	渦度分布図の利用法並びに高層天気図の予報えの 応用について、1パラメーターモデルによる上昇速 度とその利用法について
1960年(昭和 35年) 4月	電子計算室報告 IV	ECCOMMINACOVIC
1960年(昭和35年)1月	電子計算室報告 III	
1959年(昭和34年)10月	電子計算室報告 II	
	· · · · · · — · · · · · —	

# ガイダンスの解説

数値予報課報告・別冊第 64 号 平成 30 年 3 月 16 日発行

編 集 気象庁予報部 数値予報課

〒 100-8122 東京都千代田区大手町 1-3-4

発 行 気象庁予報部

Copyright © 気象庁予報部 2018 Printed in Japan 著作権法で定める範囲を超えて、無断で転載または複写することを禁止します。