Upgrade of JMA's Typhoon Ensemble Prediction System

Masayuki Kyouda
Numerical Prediction Division, Japan Meteorological Agency
and
Masakazu Higaki
Office of Marine Prediction, Japan Meteorological Agency

1. Introduction

Since February 2008, the Japan Meteorological Agency (JMA) has operated a Typhoon Ensemble Prediction System (TEPS) designed to improve track forecast targeting for tropical cyclones (TCs). The system is run in the Regional Specialized Meteorological Center (RSMC) Tokyo - Typhoon Center's area of responsibility within the framework of the World Meteorological Organization, and employs a low-resolution version of JMA's Global Spectral Model (GSM). A singular vector (SV) method is adopted in TEPS to generate initial perturbations, and dry SVs targeting the mid-latitude area are calculated for the Center's area of responsibility. The system also calculates moist SVs targeting TC surroundings where moist processes are critical. A stochastic physics scheme is used in the system in consideration of model uncertainties associated with physical parameterizations.

A major TEPS upgrade in March 2014 included enhancement for the horizontal resolution of the EPS model, revision of its physical processes and an increased ensemble size.

This report details the performance of TEPS. The system's specifications are outlined in Section 2; the impacts of individual system enhancements on TC forecasting are covered in Section 3; the results of experiments using the current system in comparison with the previous one are presented in Section 4; and Section 5 gives a summary of the report.

2. Typhoon Ensemble Prediction System Specifications

2.1 System Configuration

JMA routinely operates several EPSs to support forecasting work. As well as covering a wide range of prediction periods from early medium-range to seasonal forecasting, the Agency's suite of EPSs also supports the issuance of five-day TC track forecasts (JMA 2013).

The configuration of the current system compared with the previous one is shown in Table 1. Improvements over the previous system include enhancement of the EPS model's horizontal resolution from TL319 to TL479, revision of its physical processes (such as the stratocumulus and radiation schemes) and an ensemble size increase from 11 to 25.

At present, 25 initial conditions are integrated using a low-resolution version of JMA's GSM to produce an ensemble of 132-hour forecasts. Unperturbed analysis is conducted by interpolating the target field in global analysis. The results of sea surface temperature and sea ice analysis are referenced for the lower boundary condition, and the initialized condition is prescribed using the persisted anomaly. Accordingly, anomalies shown based on analysis for the initial time are fixed during the time integration.

Table 1: Configurations of the current and previous systems. Bold red text represents major upgrades applied to JMA's TEPS in March 2014

		Previous system	Current system
Period of operation		7th December 2010 to 10th March 2014	From 11th March 2014
Integration	Ensemble size	11	**25**
	Initial time	00, 06, 12 and 18 UTC	
	Forecast range	132 hours	

EPS model	Model type	GSM1009	**GSM1304** **- Upgraded stratocumulus scheme** **- Upgraded radiation scheme**
	Horizontal resolution	TL319 roughly equivalent to 0.5625° x 0.5625° (55 km) in latitude and longitude	**TL479 roughly equivalent to 0.375° x 0.375° (40 km) in latitude and longitude**
	Vertical resolution (top)	60 levels (0.1 hPa)	
Ensemble setting	Initial perturbation generator	Singular vector method	Singular vector method **with reduced initial perturbation amplitude**
	Initial perturbed area	Extra-tropical northwestern Pacific (20 – 60°N, 100°E – 180°) and the vicinities of up to 3 TCs	
	Model ensemble method	Stochastic physics scheme	

2.2 Frequency

TEPS is operated up to four times a day from base times at 00, 06, 12 and 18 UTC when any of the following conditions is satisfied:

•A TC of tropical storm (TS) intensity or higher is present in the RSMC Tokyo-Typhoon Center's area of responsibility (0 – 60°N, 100°E – 180°).

•A TC is expected to reach TS intensity or higher in the area within the next 24 hours.

•A TC of TS intensity or higher is expected to move into the area within the next 24 hours.

2.3 Approach to Ensemble Initial Conditions

The SV method (Buizza and Palmer 1995) is employed to perturb the initial conditions for the atmosphere. The tangent-linear and adjoint models used for SV computation are lower-resolution versions of those used in the four-dimensional variational data assimilation system (4D-Var) until October 2011. The models involve full dynamical core and physical processes including surface fluxes, vertical diffusion, gravity wave drag, large-scale condensation, long-wave radiation and deep cumulus convection. SVs based on tangent-linear and adjoint models incorporating full physical processes are called moist SVs, while those based on models incorporating simplified physical processes involving surface fluxes and vertical diffusion are called dry SVs.

Two SV calculations are introduced to efficiently capture the uncertainty of TC track forecasting. One produces dry SVs with a spatial target area fixed on the extra-tropical northwestern Pacific (20 – 60°N, 100°E – 180°), and the other produces moist SVs whose spatial target area can be moved within a 750-km radius of a predicted TC's position in one-day forecasting. Up to three movable areas can be configured for different TCs at one initial time. If more than three TCs are present in the Center's area of responsibility, three are selected in the order of concern as prioritized by the RSMC Tokyo-Typhoon Center.

Initial perturbations are determined by combining dry and moist SVs linearly. Each SV calculation can produce up to 15 SVs depending on how accurate SV estimates are, which makes the maximum number of SVs 40 (i.e., 10 dry SVs for the fixed area and 30 moist SVs for three movable areas) for each forecast event. Before the binding coefficients are determined, SVs with structures similar to those of others are eliminated. When the value of the inner product of any two SVs is 0.5 or more, one of them is eliminated from the group of SV candidates to be used for initial perturbations. After this process, the coefficients are determined based on variance minimum rotation, which creates a wide spread in the spatial distributions of the perturbations. If no SVs are eliminated, the number of independent initial perturbations is the same as the number of SVs computed. A total of 12 perturbations are randomly selected from the initial perturbations and added to/subtracted from the analysis field to produce 24 perturbed initial conditions in addition to the unperturbed condition. The amplitude of the perturbations is normalized using the moist total energy value.

2.4 Model Ensemble Approach

The stochastic physics scheme (Buizza et al. 1999) is used in consideration of model uncertainties associated with physical parameterizations. This scheme represents random errors associated with parameterized physical processes.

3. Impact of System Enhancement on Typhoon Forecasting

As detailed in Chapter 2, the major differences between the previous and current systems are listed in Table 1. The impacts of each enhancement are shown here.

RSMC Tokyo-Typhoon Center evaluates the accuracy of JMA's operational numerical weather prediction systems as well as its official forecasts, and publishes the results in the form of an annual report. For TC tracks and intensities, the same measures are adopted to compare impacts on TC forecasting in TEPS. The ensemble-mean track is determined by averaging the TC properties of individual ensemble members in TEPS. The track error is defined as the great-circle distance between the TC position as determined from analysis and the corresponding forecast position, meaning that no TC genesis in the forecast range is validated. Intensity is represented by the minimum surface pressure estimate inside a TC. In addition, TC strike probabilities (Lalaurette and van der Grijn 2002), for which data are routinely produced as a form of probabilistic guidance, are introduced into this probabilistic verification. Probability is defined as the chance of a TC center passing within a 120-km radius of a grid point during the next five days. All verification statistics are calculated using best-track data from analysis by the RSMC Tokyo – Typhoon Center within its area of responsibility.

3.1 Increased EPS Model Horizontal Resolution

A preliminary experiment involving the use of TEPS with the TL479-version GSM was conducted to investigate the impact of a higher-horizontal-resolution model on TC forecasting. The experimental system is referred to here as TL479.

The results showed that TL479 supported sharper representation of TCs than the previous TEPS. The improvement was for all tropical depressions rather than being limited to typhoon-category storms. Figure 1 shows scatter plots of TC intensity forecasting derived from unperturbed forecasts against best-track data for the previous and TL479 systems. Essentially poor correspondence between forecasting and best-track distribution is observed, although two differences between the forecasts were clearly present in the areas highlighted with ellipses. When the best-track central pressure was under 950 hPa (shown with black ellipses in both panels), TL479 was superior in deeper low-pressure system expression. Next, when the best-track central pressure was over 1,006 hPa (shown with red ellipses in the panel on the right), TL479 outnumbered the previous one in low-pressure system expression.

The results also showed higher skill in TL479 for TC track forecasting than the previous system. The track TL479 system error for unperturbed forecast in Figure 2 is smaller than that for the previous system. In Figure 3 too, the sum of track errors for unperturbed forecast classified by typhoon number is smaller for TL479 than for the previous system in most cases. The figure also shows large significant error reductions in T1215 and T1217. Figure 4 illustrates a typical case of T1217 tracks derived from three versions of the GSM with different horizontal resolutions initiated at 12 UTC on 26th September 2012. It can be seen that the large reduction is mainly due to a decrease in slow biases after recurvature. The factor may be dependency on the model's horizontal resolution. The results of TC strike probability verification showed Brier (1950) skill scores (BSSs) of 0.35 and 0.32 for TL479 and the previous system, respectively, indicating that the former provided better probabilistic information on TC track forecasting.

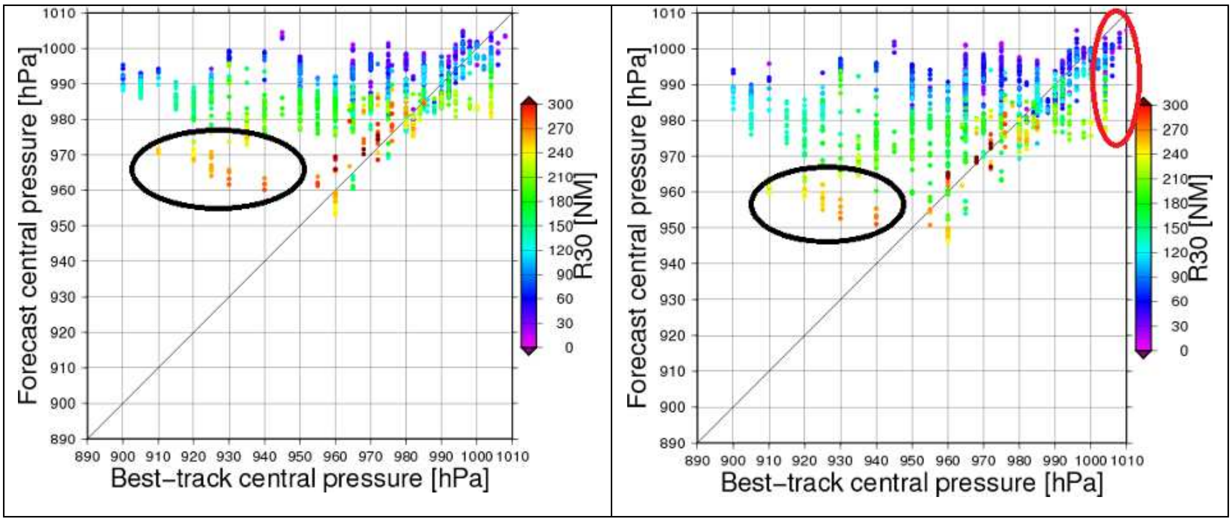

Figure 1: Scatter plots of 72-hour TC central pressure forecasting derived from unperturbed forecast (vertical axis) against best-track central pressure (horizontal axis). The left and right panels represent the results of verification for the previous and TL479 systems, respectively. The sample group verified consists of all TCs over the northwestern Pacific initiated from 00 UTC on 5th June 2012 to 00 UTC on 29th October 2012. Dots are colored according to the 30-knot wind radius in the best-track data (see the color scale in nautical miles). Plots contained in ellipses show typical differences between the two models.

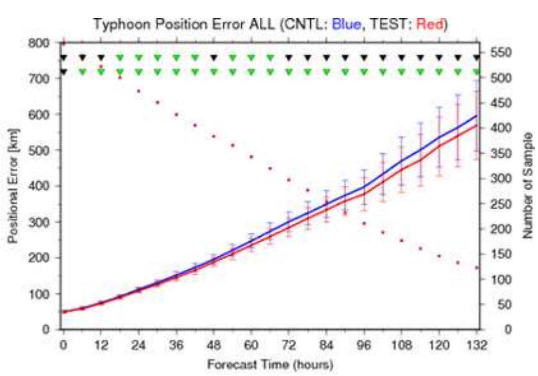

Figure 2: Mean track errors of unperturbed forecast. The sample group verified consists of all TCs over the northwestern Pacific initiated from 00 UTC on 5th June 2012 to 00 UTC on 29th October 2012. The horizontal axis shows the forecast range up to 132 hours ahead, and the blue and red lines represent the results of verification for the previous and TL479 systems, respectively. Dots indicate the numbers of samples verified based on the vertical scale on the right.

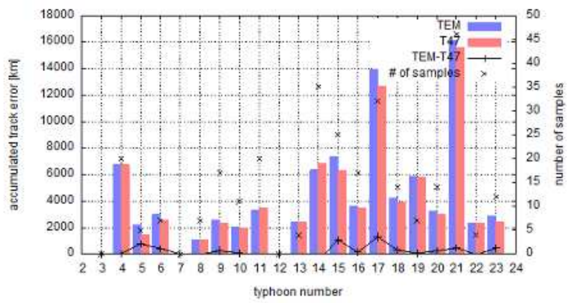

Figure 3: Sums of track errors in unperturbed forecast classified by typhoon number. The sample group verified consists of all TCs over the northwestern Pacific initiated from 00 UTC on 5th June 2012 to 00 UTC on 29th October 2012. The forecast range is 72 hours. The horizontal axis shows the typhoon number, and the blue and red bars represent the results of verification for the previous (P) and TL479 (H) systems, respectively. Plus marks represent differences defined as H – P. Crosses indicate the numbers of samples verified based on the vertical scale on the right.

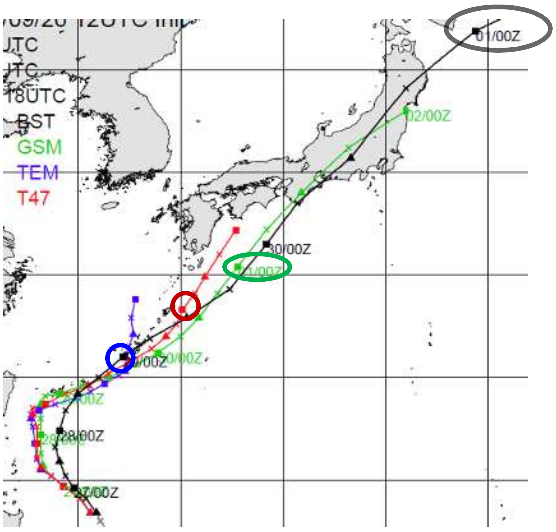

Figure 4: T1217 tracks derived from TL319-version (blue line), TL479-version (red line), and TL959-version (green line) GSMs. The initial time is 12 UTC on 26th September 2012. The black line represents the best track. Marks contained in bold ellipses represent forecast TC positions valid for 00 UTC on 1st October.

3.2 Increased Ensemble Size and Reduced Initial Perturbation Amplitude

To investigate the impact of a larger ensemble size on probabilistic TC track forecasting, an experimental configuration in which the ensemble initial conditions were increased from 11 to 25 was tested. Comparison of BSSs for TC strike probabilities showed higher values from the larger-ensemble TEPS than for the previous one, indicating that an ensemble size increase in the order of a dozen was associated with a higher level of skill. However, the enhancement produced excessive ensemble spread, causing negative impacts on ensemble TC track forecasting such that the initial ensemble spread needed to be reduced.

Based on the above experimental results, initial perturbation with reduced amplitude was applied to the previous TEPS as a more appropriate design to reduce the excessive ensemble spread. The revision meant that TC tracks predicted by perturbed forecasts would be attracted to the TC track of the unperturbed forecast, especially with a shorter lead time. Accordingly, its impact on the skill of perturbed forecast should be also verified.

The results of an additional experiment on the revision of initial perturbation with reduced amplitude (referred to here as REDAMP) indicated that the revision in the previous system had helped to improve TC track forecasting. Figure 5 compares mean track errors for REDAMP and the previous system, and shows the following:

- The track errors of REDAMP perturbed forecast were smaller than those of the previous system across the whole forecast range.
- The track errors of the REDAMP ensemble mean were smaller than those of the previous system with a shorter lead time.
- The track errors of the REDAMP ensemble mean were almost equal to those of unperturbed forecast with a shorter lead time.
- The track errors of the REDAMP ensemble mean were smaller than those of unperturbed forecast with a longer lead time as with the previous system.

In addition, the resolution of REDAMP was improved over those of the previous system in probabilistic forecasting due to reduced ensemble spread. The results of TC strike probability verification showed that the BSSs for REDAMP and the previous system were 0.382 and 0.374, indicating a positive impact on probabilistic TC track forecast skill.

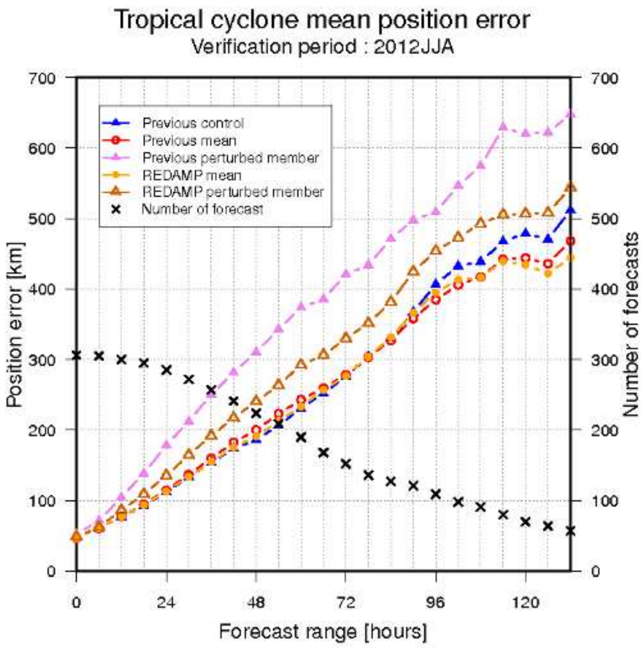

Figure 5: Mean track errors for the REDAMP and previous system. The sample group verified consists of all TCs over the northwestern Pacific initiated from 00 UTC on 5th June 2012 to 06 UTC on 30th August 2012. The horizontal axis shows the forecast range up to 132 hours ahead. The orange line and dark-orange line represent ensemble-mean and perturbed forecasts in REDAMP. The blue line, the red line and the violet line represent unperturbed, ensemble-mean and perturbed forecasts in the previous TEPS. Crosses indicate the numbers of samples verified based on the vertical scale on the right.

A further experiment was conducted to investigate the impact of a larger-ensemble reduced-amplitude system with a design based on REDAMP featuring an ensemble size of 25 (referred to here as REDAMP25). The results indicated that reduced amplitude provided better performance in combination with the increased ensemble size. The comparison in Figure 6 shows that the track errors of the ensemble means for REDAMP25 were slightly smaller than that of the previous system with a shorter lead time.

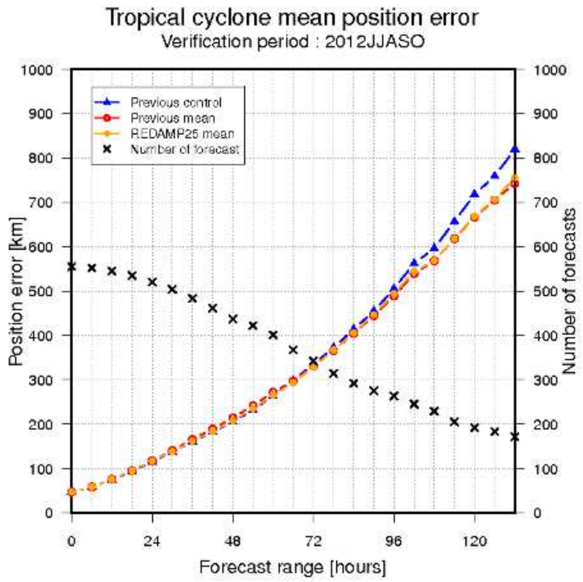

Figure 6: Mean track errors for REDAMP25 and the previous system. The sample group verified consists of all TCs over the northwestern Pacific initiated from 00 UTC on 5th June 2012 to 00 UTC on 29th October 2012. The horizontal axis shows the forecast range up to 132 hours ahead. The orange line represents ensemble-mean forecasts in REDAMP25. The blue line and the red line represent unperturbed and ensemble-mean forecasts in the previous TEPS. Crosses indicate the numbers of samples verified based on the vertical scale on the right.

4. Performance of the Current System in TC Track Forecasting

A big-sample experiment was conducted for the period from 2011 to 2013 to statistically evaluate the performance of the current system before it was put into operation. The major differences between the previous and current systems are listed in Table 1. The verification period included 2,056 TC forecasts in RSMC Tokyo – Typhoon Center's area of responsibility from 1,527 TEPS runs. Verification processing was as detailed in Chapter 3.

4.1 Ensemble Spread

The verification results showed that the size of ensemble track spreads with a forecast time of two days and the corresponding error of ensemble-mean tracks were almost the same size on average in the current system, which was an expected effect from the initial perturbation amplitude reduction explained in Section 3.2. This is seen for each ensemble TC track forecast. Figure 7 shows examples from the previous and current systems initiated at 00 UTC on 12th October 2013. Here, two ensemble-mean TC tracks and the TC tracks of each unperturbed forecast matched well up to three days ahead. It is notable that the ensemble TC track spread of the current system was clearly smaller than that of the previous one during the period. For comparison of up to one day, the reduction of the initial perturbation amplitude also resulted in a decreased wide-direction spread of the ensemble TC track.

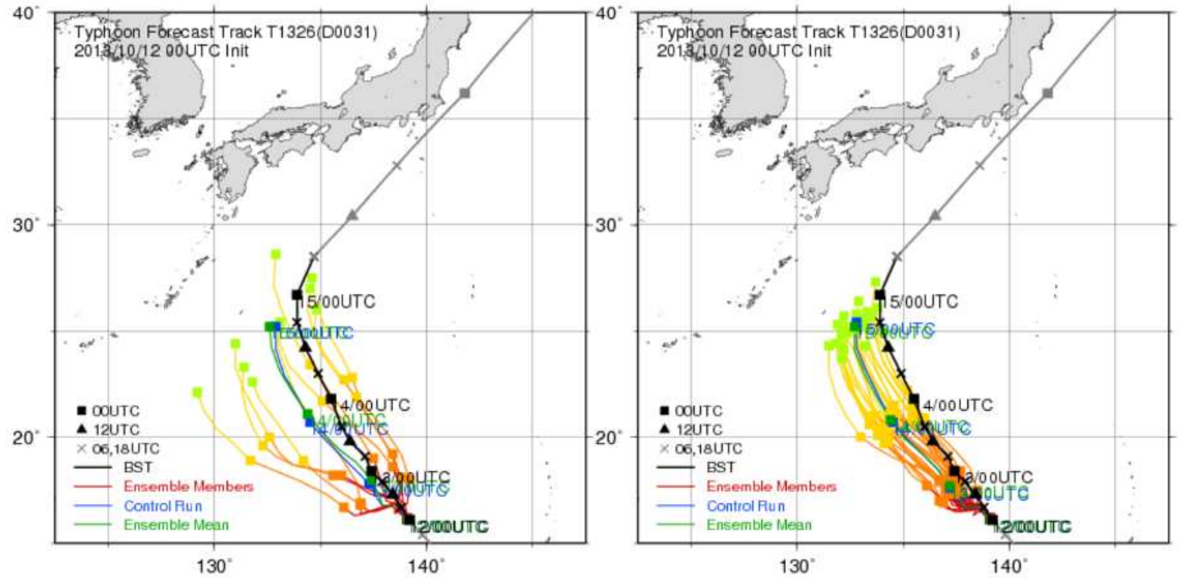

Figure 7: Comparison of ensemble TC tracks up to three days ahead derived from the previous TEPS (left panel) and the current TEPS (right panel). All tracks are for T1326 initiated at 00 UTC on 12th October 2013. Blue lines, warm-color lines and green lines represent TC tracks of the unperturbed forecast, the perturbed forecast and the ensemble mean, respectively. Black lines represent the best track.

4.2 Unperturbed Forecast TC Track

The verification results showed better TC forecast skill in the current system than in the previous one, which was an effect of the higher horizontal resolution in the EPS model as described in Section 3.1. Figure 8 shows that the track errors of unperturbed forecasts in the current system were smaller than in the previous one. In addition, as a similar effect verified in a previous impact experiment, the reduction of slow biases after recurvature and intensity biases was the main factor in enhancement for the horizontal resolution of the EPS model. However, it should be noted that systematic slow bias and shallow-pressure bias remain present.

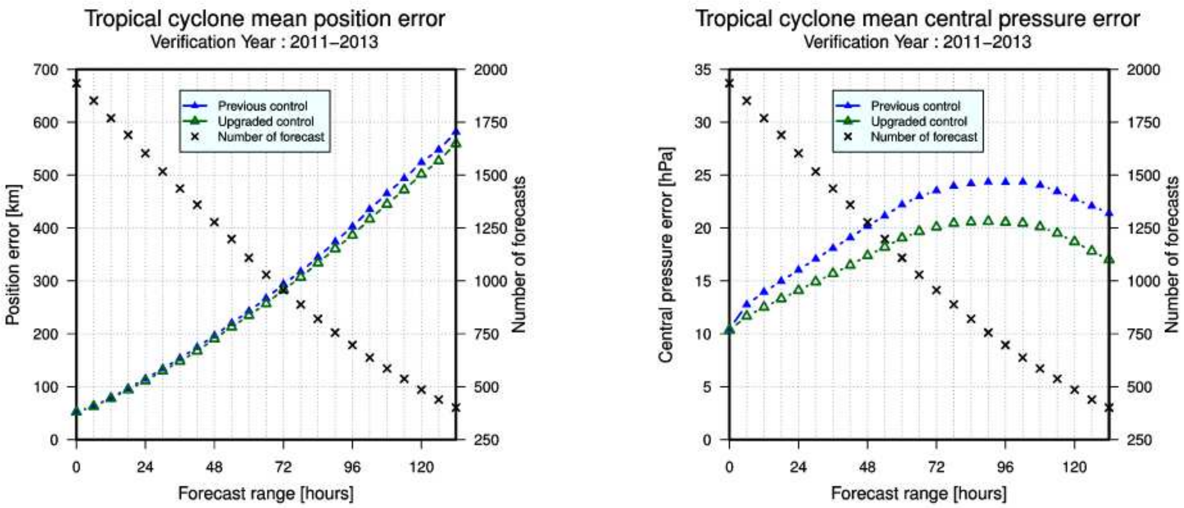

Figure 8: Mean track errors (left panel) and mean biases of central pressure (right panel) for unperturbed forecast. The sample group verified consists of all TCs over the northwestern Pacific from 2011 to 2013. The horizontal axis shows the forecast range up to 132 hours ahead, and the blue and green lines represent the results of verification for the previous and current TEPSs, respectively. Crosses indicate numbers of samples verified based on the vertical scale on the right.

4.3 Ensemble-mean Tracks

The verification results showed that the current system provided skillful ensemble-mean TC track forecasting and that its ensemble TC tracks were better than the TC tracks of the unperturbed forecasts with a longer lead time. Figure 9 shows that the track errors of ensemble means from the current system were smaller than from the previous one. It also shows that the track errors of ensemble means were equal to or smaller than that of the unperturbed forecasts in the current system.

In the cases verified, most track errors with a forecast time of five days for the current system were smaller than for the previous one. However, in the cases verified for T1112, T1204 and T1217, the track errors for the current system were larger than for the previous one. Figure 10 shows related examples for T1217 tracks from the previous and current systems initiated at 12 UTC on 22nd September 2012. In the previous system, two scenarios were considered: the typhoon would either move westward over the South China Sea or approach Japan. As a result, the ensemble-mean TC track supporting the latter scenario became close to the best track with a forecast time of five days. However, no track was close to the best track during this period, and the members representing the latter scenario did not overlap with the best track. On the other hand, the current system produced no track indicating that the typhoon would move westward. As a result, the ensemble-mean TC track veered eastward from the best track, although some TC tracks were close to the best track during the period.

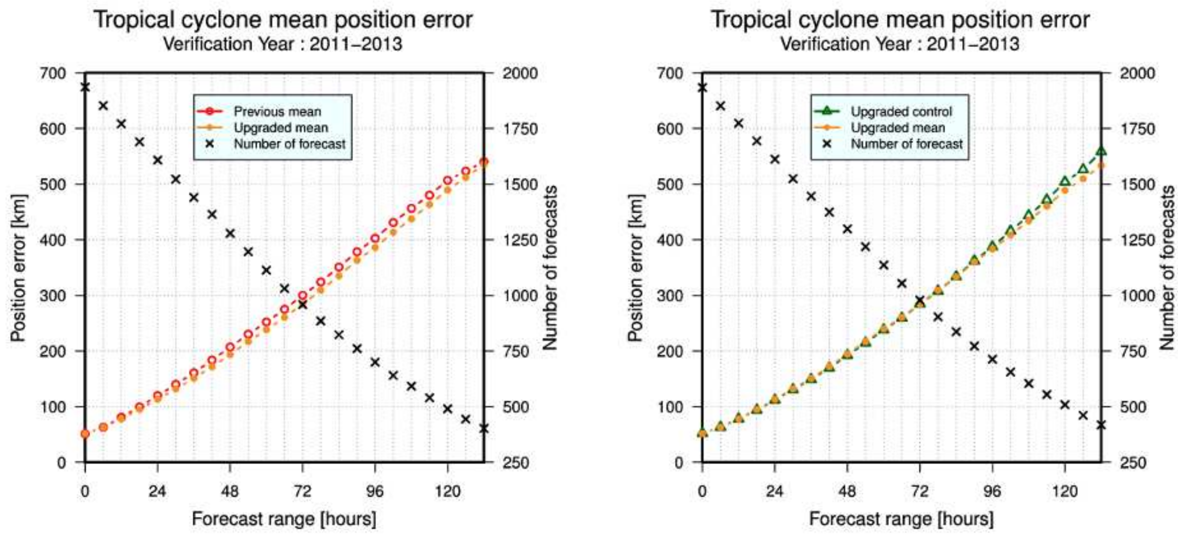

Figure 9: Mean track errors for the ensemble mean from the current TEPS (orange in both panels) in comparison with that from the previous TEPS (red in the left panel) and the unperturbed forecast in the current TEPS (green in the right panel). The sample group verified consists of all TCs over the northwestern Pacific from 2011 to 2013. The horizontal axis shows the forecast range up to 132 hours ahead. Crosses indicate the numbers of samples verified based on the vertical scale on the right.

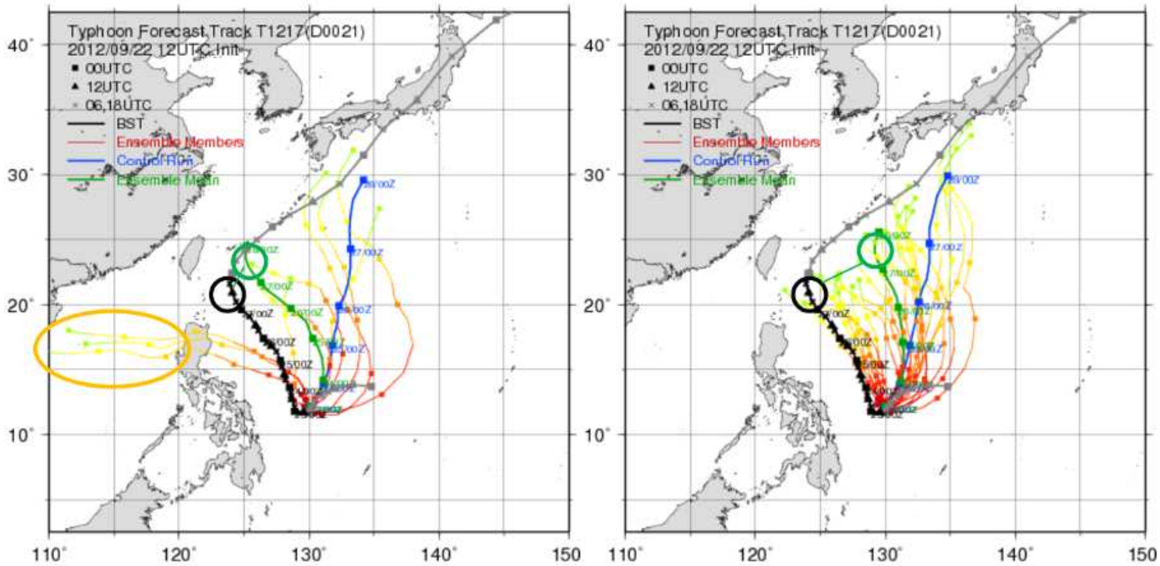

Figure 10: Comparison of ensemble TC tracks up to five days ahead derived from the previous TEPS (left panel) and the current TEPS (right panel). All tracks are for T1217 initiated at 12 UTC on 22nd September 2012. Blue lines, warm-color lines and green lines represent TC tracks of the unperturbed forecast, the perturbed forecast and the ensemble mean, respectively. Black lines represent the best track.

4.4 Strike Probability

The results of TC strike probability verification showed BSS values of 0.356 and 0.338 for the current and previous systems, indicating that the current system comprehensively provided more skillful probabilistic information on TC tracks. The diagram in Figure 11 showing the resolution and reliability of TC strike probabilities indicates that the numbers of samples with forecast probabilities of 100% and high probability from the current system were larger than from the previous one, indicating that the current system provided higher-resolution probabilities and that the improvement of its BSS came from the higher-resolution component of the BSS. It also shows that the curves for the current system exhibited higher departure from the perfect-reliability line (especially in areas of low probability where the numbers of samples were quite large), indicating that the current system provided probabilities with larger biases.

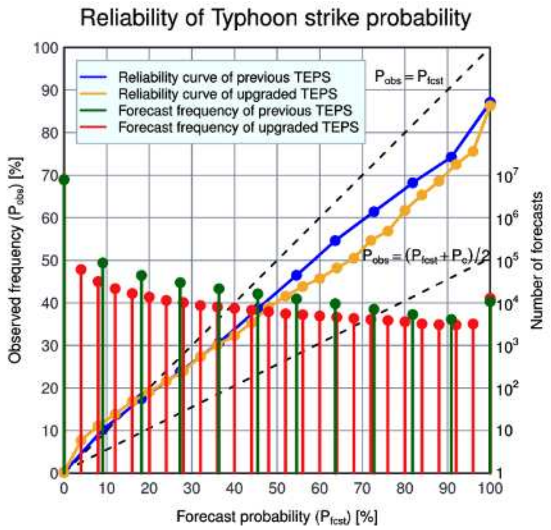

Figure 11: Diagram showing the resolution and reliability of TC strike probabilities derived from the previous and current TEPSs. The sample group verified consists of all TCs over the northwestern Pacific from 2011 to 2013. The horizontal and vertical axes show the forecast probability (P_{fcst}) and the best-track frequency (P_{obs}), respectively, and the blue and orange lines represent the results of verification for the previous and current TEPSs, respectively. The green and red lines indicate the numbers of forecasts based on the vertical scale on the right for the previous and current TEPSs. The bins are classified by ensemble member, and the red line with a forecast probability of 0% overlaps with the green line. The broken line with a slope of 1 and an intercept at 0 represents the curve of a forecast with no bias (called the "perfect-reliability line"). The other broken line with a slope of $P_{fcst}/2$ and an intercept at $P_c/2$ represents the curve of a no-skill forecast where P_c is around 1%.

4.5 Verification of Best and Worst Members

As explained in Section 3.2, a reduced initial perturbation amplitude has been adopted in the current TEPS with the expectation that the system will give optimal performance with an increased ensemble size. To further determine whether this expectation is achieved with such a configuration, additional verification was conducted as described below.

In the case where ensemble-mean TC tracks were well matched up to three days ahead (shown in Figure 7 in Section 4.1), verification shows track errors of 211 and 205 km, respectively, for the ensemble means with a forecast time of three days from the previous and current systems. These were also smaller than the mean value of about 300 km (plotted in the panel on the left of Figure 9). It is noteworthy that the largest track error among the ensemble TC tracks (referred to here as the worst error) from the current system is 380 km, which was significantly smaller than the worst error (699 km) from the previous one.

In the case where the track error of the ensemble mean with a forecast time of five days from the current system (627 km) was larger than that from the previous one (264 km; shown in Figure 10 in Section 4.3), verification also shows that the smallest track error among the ensemble TC tracks (referred to here as the best error) from the current system was 103 km, which was significantly smaller than the best error (249 km) from the previous one.

The results from verification of the best and worst errors are shown in Figure 12. Both errors of the current system were smaller than those of the previous one up to 132 hours ahead. The proportions of TC tracks with a track error of less than 120 km to the ensemble TC tracks from the current system were also higher than from the previous one except for forecast ranges of less than one day. This indicates that the current system provides more skillful ensemble TC track prediction than the previous one as a result of reduced ensemble spread and increased ensemble size.

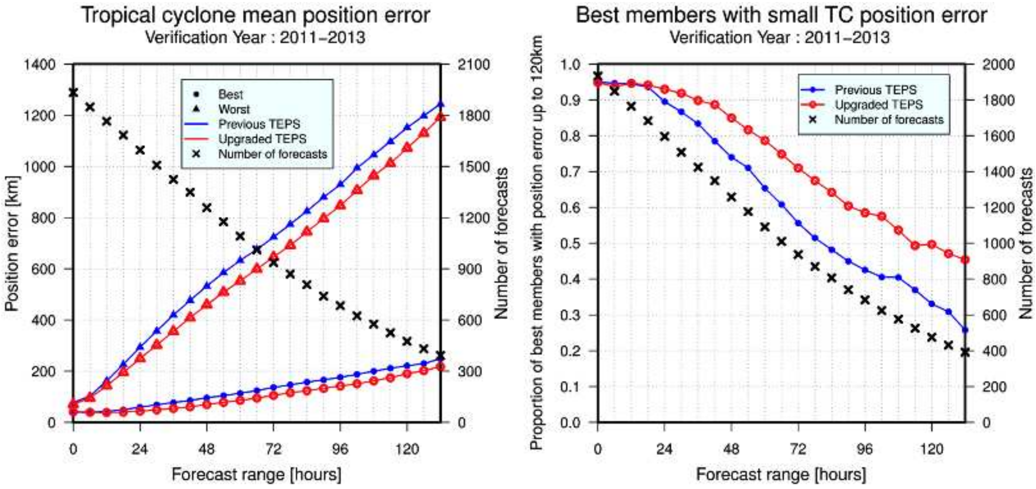

Figure 12: Comparison of best position, worst position and smaller position in ensemble TC tracks from the previous TEPS (blue) and the current TEPS (red). The sample group verified consists of all TCs over the northwestern Pacific from 2011 to 2013. The horizontal axis shows the forecast range up to 132 hours ahead. The circles and triangles in the panel on the left represent values averaged from the smallest and largest track errors in each ensemble TC track. The circles in the panel on the right represent mean proportions of TC tracks with a track error of less than 120 km to the ensemble TC tracks. Crosses indicate the numbers of samples verified based on the vertical scale on the right.

5. Summary

The Typhoon Ensemble Prediction System (TEPS) is a Japan Meteorological Agency operational numerical weather prediction system designed to improve track forecast targeting for tropical cyclones (TCs). Improvements in a major TEPS upgrade implemented in March 2014 included enhancement for the horizontal resolution of the EPS model from TL319 to TL479, revision of its physical processes, an ensemble size increase from 11 to 25, and reduction of the initial perturbation amplitude.

Preliminary experiments to investigate the impacts of each enhancement were conducted, and a big-sample experiment was conducted for the period from 2011 to 2013 to statistically evaluate the performance of the current system before it was put into operation. The performance of the current TEPS in TC track forecasting can be summarized as follows:

- The current TEPS supports sharper representation of TCs than the previous one not only for typhoon-category storms but for all tropical depressions.
- The current TEPS has better skill in TC track forecasting than the previous one, especially for TC tracks with slow bias after recurvature.
- The current TEPS provides more skillfully predicted ensemble-mean TC tracks than those derived from unperturbed forecasts with a longer lead time.
- The current TEPS provides better probabilistic information in TC track forecasting than the previous one.

REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1 – 3.

Buizza, R. and T. N. Palmer, 1995: The singular vector structure of the atmospheric global circulation. J. Atmos. Sci., 52, 1434 – 1456.

Buizza, R., M. Miller and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. Quart. J. Roy. Meteor. Soc., 125, 2887 – 2908.

Japan Meteorological Agency, 2013: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research, JMA, p. 188.

Lalaurette, F. and G. van der Grijn, 2002: Ensemble forecasts: Can they provide useful early warnings? ECMWF Newsletter, 96, 10 − 18.