Advances in Atmospheric Sciences
Impact Factor: 3.158
E-alert
ISSN 0256-1530
CN 11-1925/O4

A Time Neighborhood Method for the Verification of Landfalling Typhoon Track Forecast

Daosheng XU, Jeremy Cheuk-Hin LEUNG, Banglin ZHANG

Citation: Xu, D. S., J. C.-H. Leung, and B. L. Zhang 2023: A Time Neighborhood Method for the Verification of Landfalling Typhoon Track Forecast, *Adv. Atmos. Sci.*, 40, 273–284. doi: 10.1007/s00376-022-1398-6.

View online: https://doi.org/10.1007/s00376-022-1398-6

Related articles that may interest you

Numerical Study of Boundary Layer Structure and Rainfall after Landfall of Typhoon Fitow (2013): Sensitivity to Planetary Boundary Layer Parameterization

Advances in Atmospheric Sciences. 2019, 36(4), 431 https://doi.org/10.1007/s00376-018-7281-9

Impact of Coastal Radar Observability on the Forecast of the Track and Rainfall of Typhoon Morakot (2009) Using WRF-based Ensemble Kalman Filter Data Assimilation

Advances in Atmospheric Sciences. 2017, 34(1), 66 https://doi.org/10.1007/s00376-016-6028-8

Assimilation of GPM Microwave Imager Radiance for Track Prediction of Typhoon Cases with the WRF Hybrid En3DVAR System

Advances in Atmospheric Sciences. 2021, 38(6), 983 https://doi.org/10.1007/s00376-021-0252-6

Coupling of a Regional Climate Model with a Crop Development Model and Evaluation of the Coupled Model across China

Advances in Atmospheric Sciences. 2019, 36(5), 527 https://doi.org/10.1007/s00376-018-8160-0

Evaluation of NASA GISS Post-CMIP5 Single Column Model Simulated Clouds and Precipitation Using ARM Southern Great Plains Observations

Advances in Atmospheric Sciences. 2017, 34(3), 306 https://doi.org/10.1007/s00376-016-5254-4

Evaluation of Simulated CO2 Concentrations from the CarbonTracker-Asia Model Using In-situ Observations over East Asia for 2009-2013

Advances in Atmospheric Sciences. 2019, 36(6), 603 https://doi.org/10.1007/s00376-019-8150-x

AAS Website

AAS Weibo

AAS WeChat

Follow AAS public account for more information

ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 40, FEBRUARY 2023, 273–284

• Original Paper •

A Time Neighborhood Method for the Verification of Landfalling Typhoon Track Forecast

Daosheng XU, Jeremy Cheuk-Hin LEUNG, and Banglin ZHANG*

Guangzhou Institute of Tropical and Marine Meteorology/
Guangdong Provincial Key Laboratory of Regional Numerical Weather Prediction, *CMA*, *Guangzhou* 510640, *China*

(Received 6 December 2021; revised 10 June 2022; accepted 30 June 2022)

ABSTRACT

Landfalling typhoons can cause disasters over large regions. The government and emergency responders need to take measures to mitigate disasters according to the forecast of landfall position, while slight timing error can be ignored. The reliability of operational model forecasts of typhoon landfall position needs to be evaluated beforehand, according to the forecasts and observation of historical cases. In the evaluation of landfalling typhoon track, the traditional method based on point-to-point matching methods could be influenced by the predicted typhoon translation speed. Consequently, the traditional track evaluation method may result in a large track error even if the predicted landfall position is close to observation.

The purpose of this paper is to address the above issue using a simple evaluation method of landfalling typhoon track forecast based on the time neighborhood approach. In this new method, the timing error was lessened to highlight the importance of the position error during the landfall of typhoon. The properties of the time neighborhood method are compared with the traditional method based on numerical forecast results of 12 landfalling typhoon cases. Results demonstrated that the new method is not sensitive to the sampling frequency, and that the difference between the time neighborhood and traditional method will be more obvious when the moving speed of typhoon is moderate (between 15–30 km h^{-1}). The time neighborhood concept can be easily extended to a broader context when one attempts to examine the position error more than the timing error.

Key words: time neighborhood method, typhoon track, landfalling typhoon, model evaluation

Citation: Xu, D. S., J. C.-H. Leung, and B. L. Zhang, 2023: A time neighborhood method for the verification of landfalling typhoon track forecast. *Adv. Atmos. Sci.*, **40**(2), 273–284, https://doi.org/10.1007/s00376-022-1398-6.

Article Highlights:

- The time neighborhood method provides a meaningful verification metric when the timing error of landfall typhoon forecast can be ignored.
- The new method is insensitive to different sampling frequency in the time neighborhood window.
- The advantage of time neighborhood moved will be more obvious for the typhoon moving with a moderate speed in TRAMS model.

1. Introduction

Landfalling typhoons greatly affect the coastal regions of China and cause grave losses of life and property. For example, about 1126 people were killed when the super typhoon Fred made landfall in China in 1994. Under global warming, typhoons are predicted to be stronger (Kang and Elsner, 2016). A dramatically increasing trend of severe typhoon landfall in China was recorded from 1949–2006 (Xu et al., 2009). Although the forecast skill of typhoon track has been significantly improved in the past several decades, the corresponding verification methods are still very limited, especially for landfalling typhoons.

Traditional point-to-point methods have been widely used in the verification of numerical weather prediction. Although it provides meaningful interpretations of model forecast errors from different perspectives, its limitation in the evaluation of high-resolution model forecasts has been pointed out by Mass et al. (2002). For example, when predictions deviate from observations, the traditional point-to-point methods may result in a "double penalty" in observed-but-not forecasted and forecasted-but-not-observed cases (Ebert, 2009). In recent years, many new methods have

---

* Corresponding author: Banglin ZHANG
Email: zhangbl@gd121.cn

been developed to reduce the impacts of small-scale forecast bias on the verification results (Theis et al., 2005; Mittermaier, 2006; Casati and Wilson, 2007; Jenkner et al., 2008; Ahijevych et al., 2009; Yu et al., 2013; Fang and Kuo, 2015), and one popular method is based on the neighborhood strategy (Roberts and Lean, 2008).

The neighborhood method verifies the forecast within a spatial/temporal neighborhood around each grid point. When the observation is consistent with the forecast in this neighborhood area, the forecast is considered to be a "good" forecast. Compared to the exact grid to grid verification used in traditional methods, the neighborhood method can be seen as a kind of "fuzzy" technique. It can provide guidance on the spatial or temporal scales at which the forecasts should be used to meet certain accuracy requirements. The neighborhood method has been demonstrated to be more meaningful than the traditional point-to-point methods by reducing the impact of displacement error on the calculation of verification score (Clark et al., 2010). The neighborhood method has been widely applied in the evaluation of precipitation forecasts (Mittermaier and Roberts, 2010; Wolff et al., 2014; Zhu et al., 2015; Schwartz, 2017; Stein and Stoop, 2019; Faure et al., 2020). Recently, the neighborhood method has been further extended to the verification of gridded wind forecasts. Skok and Hladnik (2018) showed that the evaluation score based on the neighborhood method can distinguish the forecast results even when the displacement of wind patterns is large.

Although the neighborhood strategy has been widely used for the verification of spatial forecasts (Gilleland et al., 2009), its application to time-varied forecast evaluation (for example, the track forecast of typhoon center) has been infrequent. In the traditional verification method of typhoon deterministic track forecasts, point-to-point matching models are usually adopted to measure the forecast error. For example, in the operational tropical cyclone forecast verification practice in the western North Pacific region, the typhoon landfall position error is calculated as the distance between the actual landfall position and the intersection point of the forecast track and the coast line (Yu et al., 2012). The verifications of typhoon forecast are functions of the forecast lead time, while the forecast errors at different lead times are independent of each other.

There are instances where the traditional verification result is difficult for users to be able to distinguish between a "near miss" and much poorer forecasts, which may give misleading results that are inconsistent with our general understanding and interpretation of typhoon track forecast error. For example, Figure 1 gives a schematic diagram showing two forecasted tracks (F1 and F2) for an observed typhoon (O). Compared with F2, F1 obviously predicts a typhoon track that is closer to the observation although there is a slight time lag in the typhoon movement. However, one could obtain the same track error of F1 and F2 under the traditional point-to-point verification method, i.e., the distance between F1(t_1) and O(t_1) is similar to the distance between F2(t_1) and O(t_1). In other words, the traditional point-to-point verification method could result in a relatively large track error even when the forecasted typhoon track is very similar to the observation; such cases are referred to as "near miss" forecasts. The above example illustrates that the traditional typhoon track verification method is sensitive to the timing errors of typhoon forecasts, and sometimes it cannot satisfy the needs of users when exact timing is not critical. For example, in operational practice, the government needs to issue typhoon landfall warning and take precaution-

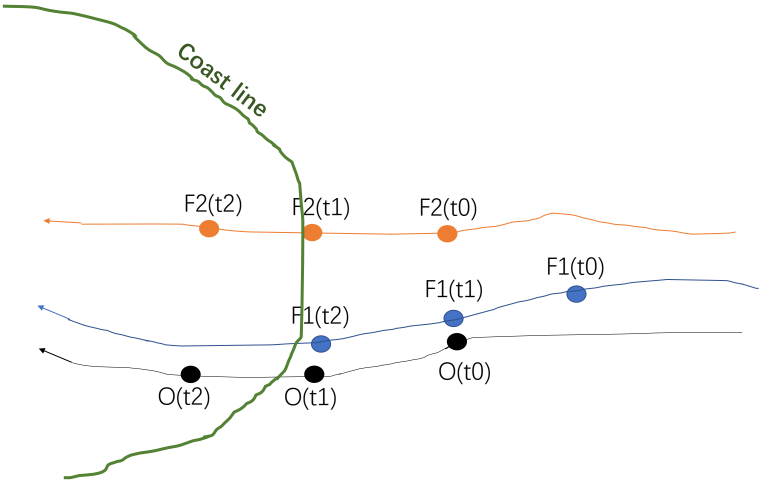

Fig. 1. Schematic diagram illustrating the verification results of "near miss" and much poorer forecasts. The black dots mark the position of observed typhoon center at t_0, t_1 and t_2 [denoted as O(t_0), O(t_1) and O(t_2)], the blue dots mark the "near miss" forecasts of typhoon center at t_0, t_1 and t_2 [denoted as F1(t_0), F1(t_1) and F1(t_2)], the orange dots represent the much poorer forecasts of typhoon center at t_0, t_1 and t_2 [denoted as F2(t_0), F2(t_1) and F2(t_2)]. The bold green line denotes the coastline.

ary measures to avoid typhoon-related disasters about 2–3 days before typhoon landfall. If the precautionary measures are implemented at the right place, the damages can still be avoided successfully even if the timing of landfall is early or late by a few hours. Conversely, the warning and measures will be completely meaningless if the landfall position is far from the actual location. In this sense, the landfall position will be more important than the exact landfall time. Thus, the value of typhoon track forecast could vary according to its application, and the traditional verification method cannot satisfy the needs of all users. Additional approaches should join the current verification toolbox based on "user focused" view.

To alleviate the sensitivity of verification results to the timing error, Colle et al. (2001) used a temporal average precipitation forecast as an evaluated parameter. Mass et al. (2002) also suggested that a temporal or spatial shifting of model fields could be used to emphasize the more realistic structures of weather system simulated by high-resolution models. Duc et al. (2013) extended the concept of spatial neighborhood to the time dimension to alleviate the influence of small-scale variability existing in verification of high-resolution ensemble forecasts. In evaluating medium-range forecast of typhoon tracks, Davis et al. (2016) introduced a type of object-based approach, which defined the forecasts hit observation once their distance between the observation and the forecast was smaller than a pre-selected threshold (refer to their Fig. A1). Although the purpose of Davis et al. (2016) was to identify matching pairs of forecasted and observed tracks, their method can be viewed as a kind of neighborhood method as the influence of small position bias was ignored in verification of typhoon track.

In this paper, the time neighborhood approach is proposed to lessen the effects of timing errors in the verification of landfalling typhoon track prediction. It provides a new metric which better emphasizes the accuracy of forecasts for a given position by excluding the timing error. The metric of the time neighborhood method can better reveal a model's ability to forecast the landfall position prominently under the provision that the timing error is not so important. In the following content, the time neighborhood method will be introduced and applied to the evaluation of typhoon track forecast of 12 landfalling storms occurred near the south China during 2017–18, and its advantages will be illustrated by making comparisons with the traditional point-to-point typhoon track verification method. Section 2 offers a detailed introduction of the time neighborhood method. In section 3, the proposed neighborhood method is applied to the evaluation of typhoon track forecast. Conclusions are presented in section 4.

2. The time neighborhood evaluation method

This section introduces how the time neighborhood method is applied to typhoon track forecast evaluation.

In the traditional evaluation method, the track forecast error at moment t_0 is equal to the distance from the observed typhoon center to the forecasted center at t_0 (the red dashed line in Fig.2). As mentioned above, this traditional point-to-point typhoon track forecast evaluation method is sometimes sensitive to the timing error or typhoon speed error, due to the fact that it judges the predicted typhoon position based on the result of only one forecast time point. Hence, the time neighborhood method is proposed in order to solve the problem. Although the time interval of typhoon track evaluation is often set as six hours, the model forecast of typhoon can be output at a higher time frequency. The higher frequency forecasts may be useful even when they do not

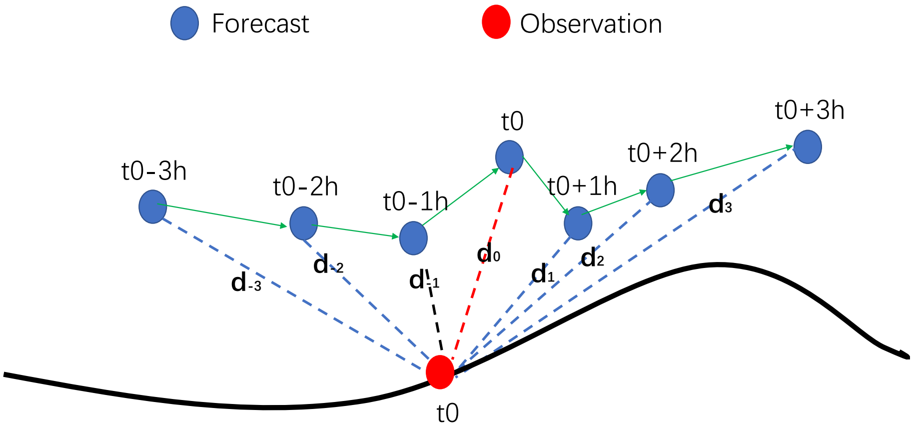

Fig. 2. Schematic diagram of the typhoon track error evaluation based on the time neighborhood method. The black thick solid line denotes the observed typhoon track, and the green arrows indicate the forecasted typhoon track. The blue dots represent forecasted typhoon position at different times (three hours before and after landfalling time t_0). The red dot refers to the observed typhoon center at t_0, and the blue dashed lines represent the typhoon center track deviations (d_{-3}, d_{-2}, …, d_2, d_3) of the forecasts at different times (t_{-3}, t_{-2}, …, t_2, t_3) relative to that at t_0. The track error at t_0 according to the traditional method (d_0) is highlighted with the dashed red line.

match the observations exactly at the same time, since it is possible to avoid the influence of timing error by rewarding the closest forecast with time-expanded samples (Xu et al., 2008; Zhao et al., 2015; Huang and Wang, 2018).

In the time neighborhood method, the track forecast error at any given moment in time is defined as the minimum distance between the observed and forecasted typhoon centers within a time window. Technically speaking, the samples within a time expanded neighborhood are compared with the observation at t_0 in order to reduce the impact of timing error on the verification of typhoon track. The scale of the sampling time window (i.e. the size of neighborhood) is designed to be a tunable parameter in this scheme. By varying the length of time window, it is possible to determine the scales for which the forecast has sufficient skill for a particular application. Thus, it is a concept similar to the radius in the spatial neighborhood method for the verification of precipitation, which should be adjusted according to the scales of different weather system. In this study, the sampling time window is defined as half of the time interval of observations for simplicity. Because the typhoon center positions are usually verified every six hours, the radius of the time neighborhood is set here to 3 hours. The distances d_{-3}, d_{-2}, d_{-1}, d_0, d_1, d_2, d_3 (denoted in Fig. 2) between the model forecasts within the sampling time window (t_{-3}, ... , t_3) and the actual observation at t_0 are then calculated for evaluating the typhoon track forecasting error. A simple interpretation of the methodology is that all of the forecasts within the sampling time window are supposed to be equally likely estimates of the true value [in principle, a Gaussian or other kernel could be used to give a greater weighting to the central values, as suggested by Roberts and Lean (2008)]. To reward the closest forecasts in time neighborhood, the minimum value of d_{-3}, d_{-2}, d_{-1}, d_0, d_1, d_2, d_3 (i.e., d_1 in Fig. 2) is defined as the forecast error at time t_0. For the evaluation of the very first and last time point of forecast, only the forecast samples in half of the time window are available.

Note that by definition, the traditional method must have larger track errors than our neighborhood method. On its own, that's not a research finding, so the meaning of this metric needs to be placed into a broader context. Mathematically, the time neighborhood verification depends more on the general direction of the typhoon track and ignores the minor timing error caused by biases in typhoon translation speed. Thus, the time neighborhood method can effectively reduce the sensitivity of verification results to the timing error.

The distinction of the time neighborhood method is somewhat similar to the traditional decomposition of tropical cyclone track errors into along- and cross-track components (Elsberry and Peak, 1986). As illustrated in Fig. 3, the along-track error is defined as the component of absolute error in the direction of the observed typhoon track, and the cross-track error is defined as the component of absolute error in the direction perpendicular to the observed track. Thus, when the typhoon travels along a nearly straight line, the speed error (or timing error) can be measured by the along-track error, while the cross-track error can be viewed as the forecast position error of the typhoon at landfall. In this case, the track error calculated with the time neighborhood method can be viewed as a weighted average of the along- and cross-track errors, with the weighting of along-track error being lessened.

Although the relationship of the above decomposition method and time neighborhood method is very similar for a straight moving typhoon, it should still be stressed that the time neighborhood is not simply equal to a reduction of the impact of along-track error. As can be seen from the schematic shown in Fig. 4, the track error calculated from the time neighborhood method [distance between F(t_2) and O(t_1)] is much smaller than from the traditional method [distance between F(t_1) and O(t_1)]. It is obvious that the reduction of track error by the time neighborhood method is caused by the offset of cross-track error during the deflections of typhoon track. In this case, the landfall position error will be more closely related to the cross-track error, while the metric of neighborhood verification can still describe reasonably well the performance of landfall position forecasts. The neigh-

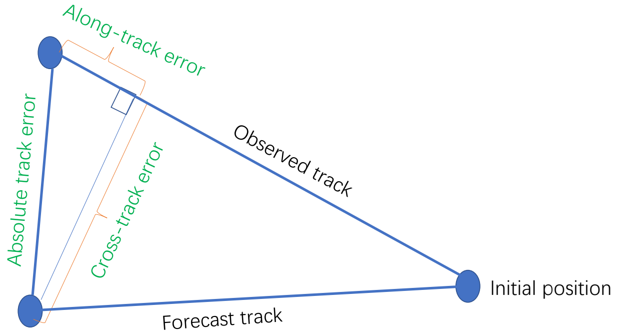

Fig. 3. Schematic diagram of cross- and along- track errors in relation to observed and forecasted typhoon tracks.

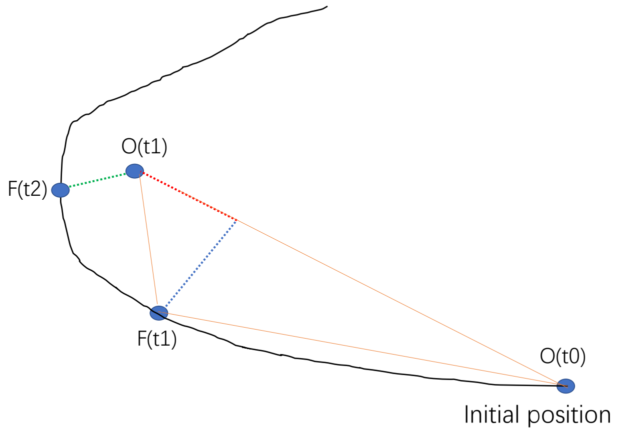

Fig. 4. The decomposition of along-track error (red dotted line) and cross-track error (blue dotted line) during the deflections of typhoon track. The track error using the time neighborhood method (green dotted line) is also displayed. $O(t_0)$ is the initial position and $O(t_1)$ is the observed landfall position at time t_1, $F(t_1)$ and $F(t_2)$ denote forecasted typhoon positions at time t_1 and t_2.

borhood method is applicable to various types of typhoon tracks.

To the best of our knowledge, the metric provided by the time neighborhood method has not been presented in any official annual track verification report (e.g., Chen et al., 2020) or the verification methods for tropical cyclone forecasts proposed by World Meteorological Organization (WMO, 2013). Therefore, the time neighborhood can be seen as a useful supplement of the traditional decomposition of tropical cyclone track errors into along- and cross-track components, at least for the verification of landfall typhoon.

3. Comparison between the time neighborhood approach and traditional point-to-point approach

3.1. *Configurations of the verification*

An intercomparison between the traditional point-to-point and time neighborhood method has been performed by evaluating the track forecasts of 12 landfalling typhoon cases during 2017–18 near South China (Fig. 5). The 72-hour forecast results of the operational Tropical Regional Atmospheric Model System (TRAMS) model with a horizontal grid spacing of 9 km (Xu et al., 2020) are evaluated in this study. The 12 landfalling typhoon cases simulated with the TRAMS model employed initial and lateral boundary conditions from the ECMWF operational global model. The forecasts of typhoon center position were output every one hour, and the observations recorded every six hours. Although there are many different measures for the tracking of typhoon center (Tao et al., 2011; Tory et al., 2013; Biswas et al., 2018), we simply define the location of minimum geopotential height at 850 hPa as the typhoon center, which was adopted in the operational verification system for the TRAMS model (http://www.grapes-trams.org.cn/TyphoonVerification.aspx). Being the first attempt to test the new verification method, the main purpose of this paper is to examine the difference between the time neighborhood and traditional methods. In light of previous studies on the intercomparison of different forecast verification methods (Mittermaier and Roberts, 2010; Ebert, 2009), we believe 12 cases is sufficient to show the advantage of our new method. Of course, when the new method is used to evaluate the performance of different operational models, the number should be extended significantly to include the cases in one or more typhoon seasons so as to be representative.

Figure 6 displays the different configurations for the verification with the traditional method and the time neighborhood method. For the traditional method, only the forecast at t_0 was used for verification, which is equal to 6-hour sampling frequency in the time neighborhood method, therefore it is named Test-6hfrq for the convenience of comparison. To examine the influence of different sampling frequencies on the results of the time neighborhood method, 3-hour and 1-hour sampling frequencies are used in the time neighborhood method and are denoted as Test-3hfrq and Test-1hfrq, respectively.

3.2. *Verification results*

The resulting verification for all 12 cases with Test-6hfrq, Test-3hfrq and Test-1hfrq are shown in Fig. 7. Clearly, Test-6hfrq has the largest track error while the time neighborhood method (Test-3hfrq and Test-1hfrq) greatly reduces the track error for some of the cases. However, the difference between Test-3hfrq and Test-1hfrq is very small, which indicates that the forecast time bias of TRAMS model is regularly larger than 3 hours.

As can be seen in Fig. 7, the difference between Test-6hfrq and Test-3hfrq (or Test-1hfrq) vary from case to case. It will be beneficial to distinguish what circumstance lead to a smaller error than, or a similar error to, the traditional method. We first divided the 12 cases into two groups accord-

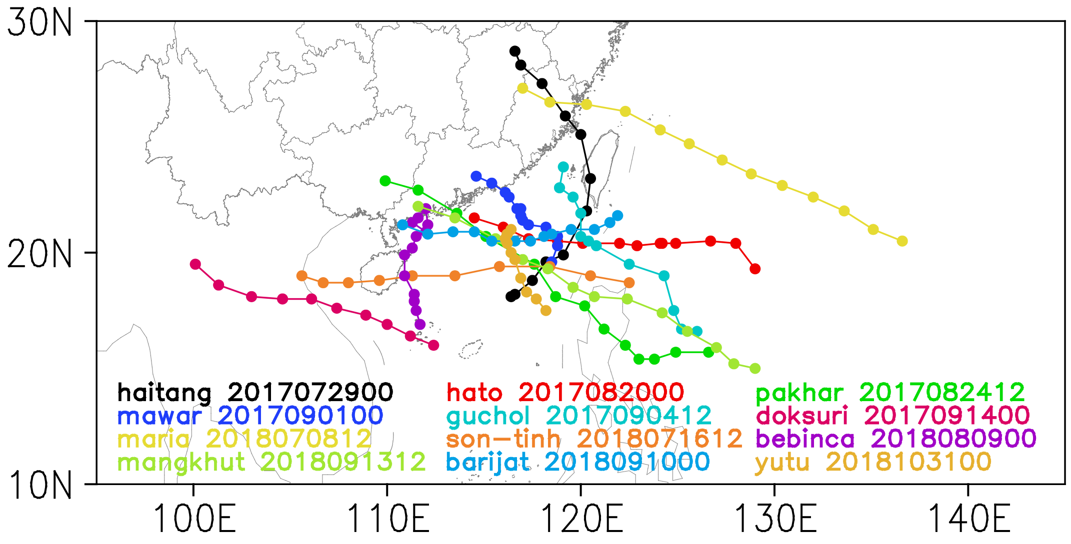

Fig. 5. Observed tracks every 6 h for 12 landfalling typhoon cases used for the verification test in this study. The typhoon names and initial time of forecast are given in the figure.

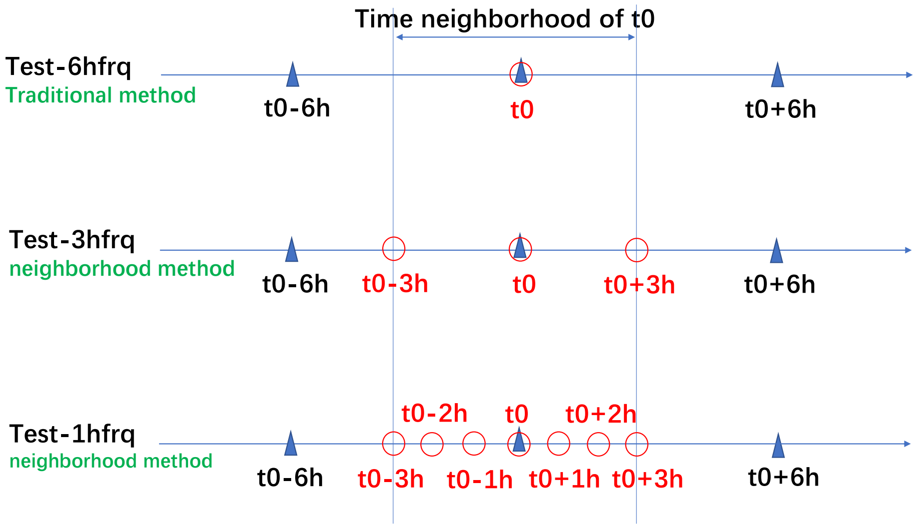

Fig. 6. A schematic diagram of the three sets of forecast evaluation experiments, i.e., Test-6hfrq, Test-3hfrq, and Test-1hfrq. The red circles indicate the forecasts in the neighborhood window (from t_0–3 h to t_0+3 h), and the blue triangles denote the observation. Test-6hfrq is equivalent to the traditional method, while Test-3hfrq and Test-1hfrq belong to the time neighborhood method.

ing to the difference in errors shown in Fig. 7 between the traditional method and the time neighborhood method. The “similar-error” group includes five cases, which are Yutu (2018), Mawar (2017), Bebinca (2018), Maria (2018) and Son-tinh (2018). The remaining seven cases belong to the “smaller-error” group. When the 12 cases are plotted according to their observed and forecasted average moving speed (Fig. 8), the different circumstances for the “similar-error” group and the “smaller-error” group can be easily identified.

The “similar-error” group (denoted with dashed ellipse in Fig. 8) is mainly distributed into two extreme situations. First, when the speed of typhoon motion is very slow (such as Bebinca (2018), Mawar (2017) and Yutu (2018), with speed <15 m s^{-1}), the verification results will be insensitive to the length of time neighborhood window. This can be easily interpreted. Since the sampling time window is defined as half of the time interval of observations in this study, the space distance traveled through the time neighborhood will be reduced when the typhoon moves slowly. Thus, the error will be closer to each other whenever choosing the closest forecast in time neighborhood method or the forecast at t_0 in traditional method. Secondly, when the typhoon translates quickly (such as typhoon Maria (2018) and Son-tinh (2018), whose speed >30 m s^{-1}), the verification result will also be similar. The reason for this is that the translational speed forecasted by the TRAMS model matched the observation quite well. As shown in Fig. 8, the forecast error of averaged translational speed is less than 1km/h both for Maria and Son-tinh.

The translational speed of the other seven cases in the “smaller-error” group are located between 15 and 30 m s^{-1}, which is a moderate speed for the typhoons in the Northwest Pacific Ocean. The TRAMS model underestimated the translational speed for most of the seven cases except for typhoon Haitang being overestimated. In such a circumstance, the time neighborhood will lead to a smaller error than the traditional method. The reduction of error can be attributed to the advantage of the time neighborhood method in identifying the “near miss” cases.

It is helpful to examine selected examples to better understand the behavior of the time neighborhood method. Two interesting examples (Typhoon “Mangkhut” (2018), denoted as example A; Typhoon “Maria” (2018), denoted as example B) are selected from the above 12 typhoon cases and analyzed in detail below. The verification results of the two cases based on the traditional method have obvious differences. The track error of example A in Test-6hfrq was about 80km between 30-60 hours lead time (Fig. 7i), while the track error of example B was about 20 km in Test-6hfrq (Fig. 7g). With the time neighborhood method, the objective verification results of example A and B are very close to each other. The detailed reasons for the differences between traditional and time neighborhood methods will be discussed below.

Example A: Typhoon “Manghkut”(2018)

Typhoon “Manghkut”(2018) translated to the northwest and finally landed southwest of Guangdong province. The observed center positions (every six hours) and corresponding forecasts (every one hour) are displayed in Fig. 9a. The forecasted direction of typhoon motion was consistent with observation, except its speed was slower than the observations by about 3 hours (comparing the green and black dots in Fig. 9a). Compared to the traditional method, an obvious reduction of track error is estimated by the time neighborhood method (Fig. 7g). For the lead times of 30–66 h, the errors are reduced roughly from 80 km to about 20 km.

It is helpful to understand the difference between the two verification results from the perspective of decomposing

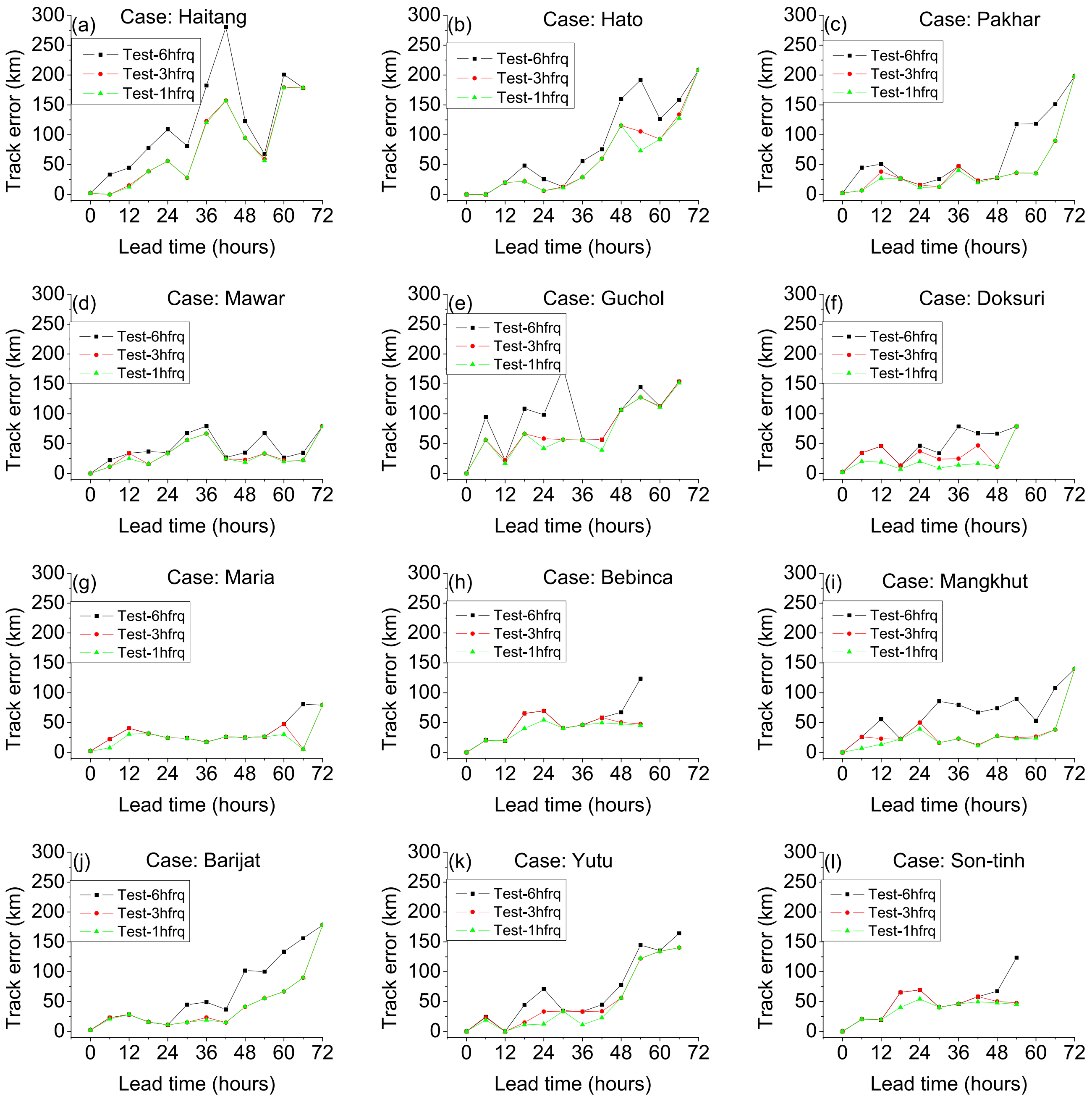

Fig. 7. The track error of 12 typhoon cases, verified with Test-1hfrq (green), Test-3hfrq (red) and Test-6hfrq (black), respectively. The names for the typhoon cases are listed at the top of each panel.

the track error into cross-track and along-track components. The absolute cross-track-error and along-track error of typhoon "Manghkut" (2018) are plotted every six hours in Fig. 10. Results show that the cross-track error (the bias of moving direction) of forecasts is very small for typhoon "Manghkut" (2018), but the along-track error (the bias of moving speed) increases with forecast time. This implies that the absolute error is mainly due to the bias of forecasted typhoon motion. The large along track error causes the large track errors according to the traditional point-to-point approach (Fig. 7i), even if the forecasted typhoon position is close to the observation (Fig. 9a). As the cross-track error (the bias of moving direction) of forecasts is very small for typhoon "Manghkut" (2018), and the along-track error (the bias in translational speed) can be effectively offset by the inclusion of time-expanded samples, the absolute errors are then significantly reduced in time neighborhood method.

The difference between Test-3hfrq and Test-1hfrq is not apparent for "Manghkut" (2018) (Fig. 7i). Verification of track error for 36 h forecast was taken as an example to interpret this phenomenon. As shown in Fig. 11a, the track error of Test-1hfrq can be defined as the distance between the observation [denoted as O(36)] and the nearest forecast [denoted as F(39)]. For Test-3hfrq (Fig. 11b), the forecast closest to O(36) is also F(39), which is coincident with Test-1hfrq. Thus, the results of Test-1hfrq and Test-3hfrq are equivalent at 36 h. The similarity of verification results at other time points can also be attributed to the same reason.

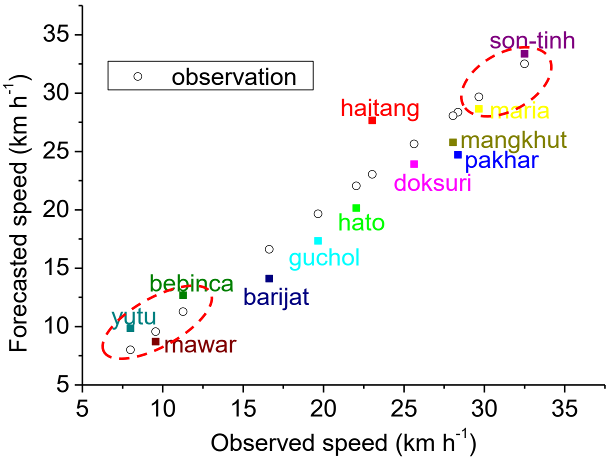

Fig. 8. Comparison of the observed and forecasted translation speed for 12 typhoon cases. The translation speed is averaged over the 72 h period for each case. The observation is denoted by empty circles, and the forecast is denoted with solid squares with different colors. The name of each case is also labelled in the same color with corresponding solid squares. Five cases with similar error are circled with dashed ellipse.

Example B: *Typhoon "Maria"(2018)*

The track of "Maria"(2018) is very similar to that of "Manghkut"(2018), except for landing at Fujian province (Fig. 9b). However, the difference among track errors from different verification methods is very small during 0–66 h, and the results for 18–54 h are exactly the same (Fig. 7g). Comparing the forecasted typhoon center positions and observation at the same time (the green dots in Fig. 9a), it can be seen that the along track error of "Maria"(2018) is very small, so the closest forecast in time neighborhood is actually the same as the forecasts used in the traditional method. This is why the verification result is not sensitive to different methods.

As illustrated in the analyses of the above two examples in which the typhoon advanced nearly in a straight line, the along track error is a very important issue to determine whether a typhoon track evaluation is sensitive to the timing error. The translational speeds of forecasted typhoon centers were then calculated and compared with observations. As shown in Fig. 12a, the motion of typhoon center in example A is slower than observation during the lead time of 12–36 hours, while the forecasted speed during 36–60 hours is close to observation. The track error caused by slower movement at the early time was retained until the end of the forecast

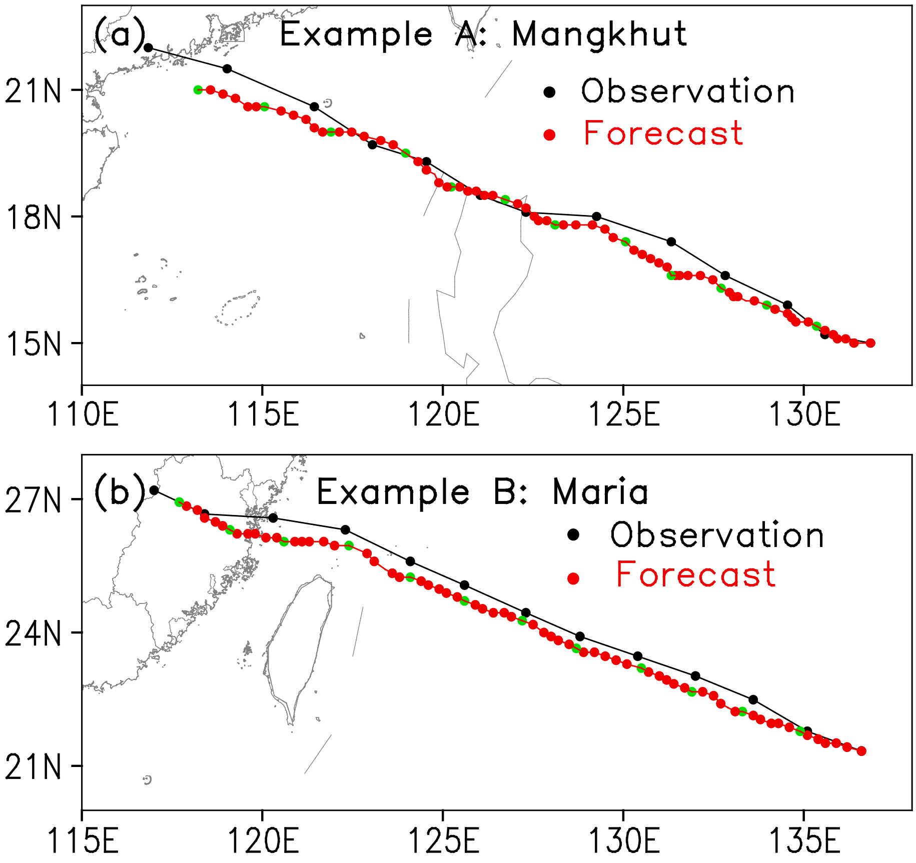

Fig. 9. (a) The observed (black dots) and forecasted (red dots) track of typhoon "Manghkut" (Example A), where the observations are displayed every six hours, and the forecasts are displayed every one hour. The forecasts are marked with green color every 6 hours to compare with observation; (b) same as (a), but for typhoon "Maria" (Example B).

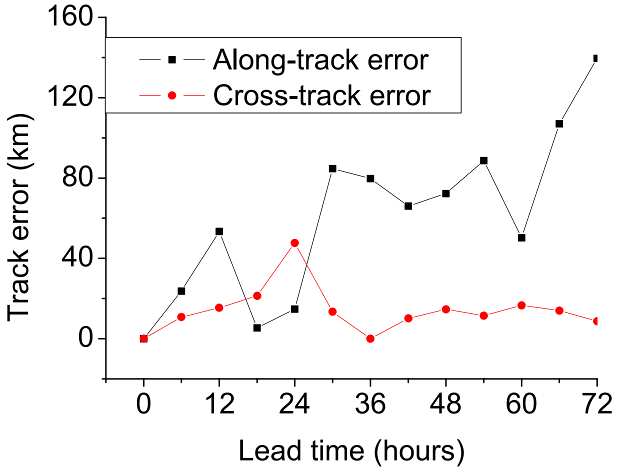

Fig. 10. The absolute cross-track error and along-track error of typhoon "Manghkut".

in the traditional verification method (refer to Test-6hfrq in Fig. 7i), while in the time neighborhood method, this kind of temporary speed bias does not affect the track error (refer to Test-3hfrq and Test-1hfrq in Fig. 7i), as the influence of speed bias is compensated by choosing the "closest" forecast to calculate track error. For example B, the forecasted moving speed is consistent well with observation in general during the lead time of 0–54 hours (Fig. 12b), so the samples selected for verification in time neighborhood method is just the same as traditional method, and then the difference of verified track error from the two methods is not obvious (refer to Fig. 7g).

4. Conclusions and discussion

User needs for typhoon verification information are

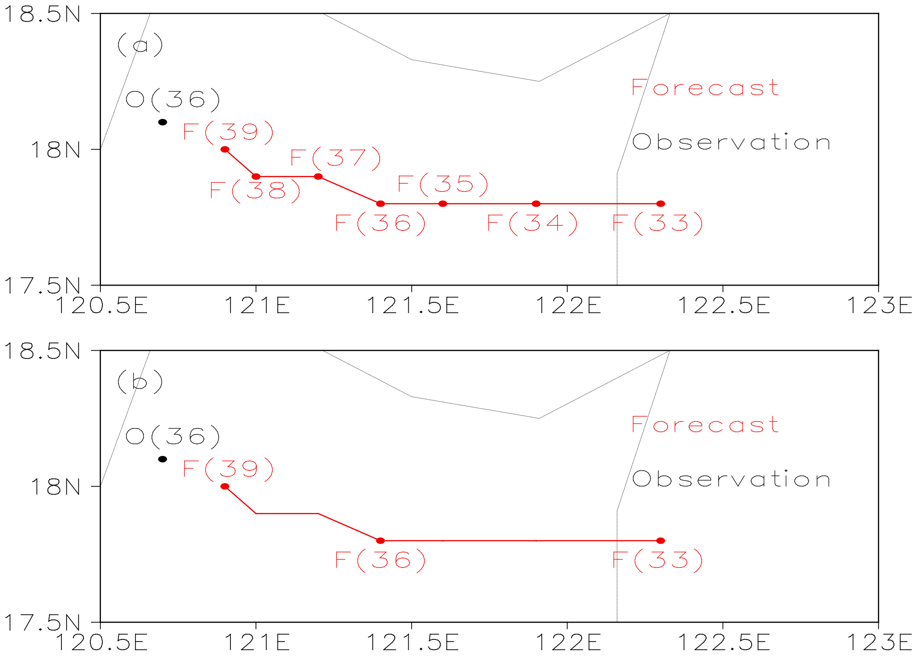

Fig. 11. The observed center of typhoon "Manghkut" at 36 h [black solid dot denoted as O(36)] and the forecasted typhoon track during 33–36 h (red line). The typhoon center is shown with a red solid dot every 1 h for Test-1hfrq (a) and every 3 h for Test-3hfrq (b).

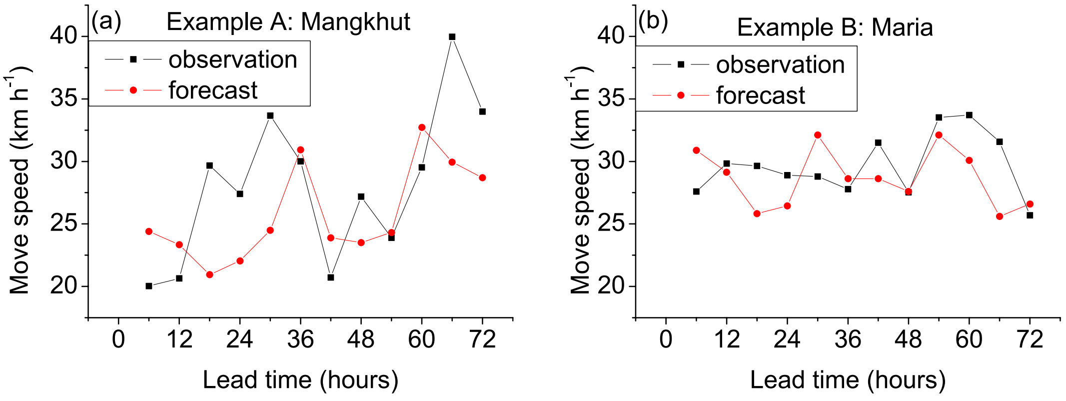

Fig. 12. The observed and forecasted translational speed of typhoon center for (a) "Manghkut" and (b) "Maria".

rather diverse, ranging from the modeler's need for information on the accuracy of the detailed three-dimensional structures of incipient storms at sea to the disaster planner's need for information on the accuracy of forecasts of landfall timing, location and intensity. Emergency managers and government agencies would be expected to be interested in verification information about the typhoon landfall position which directly impact their decision-making processes.

To emphasize the potential value in the accurate forecast of landfall position, a new verification method, the time neighborhood method, was proposed to evaluate typhoon track forecasts in this paper. By including the time-expanded samples, the closest forecast near the time of observation is rewarded by defining the minimum error as the track error. After taking into account the higher time frequency forecasts as supplementary information for the traditional method, the time neighborhood method can effectively identify the "near miss" cases (the cases have tracks similar to the observed ones but with timing difference), which will be helpful to evaluate the performance of model forecast in typhoon landfalling position more effectively.

Analyses of 12 landfalling typhoon cases in the South China Sea are performed to identify some properties of the time neighborhood method. When the typhoon moved slowly (with speed less than 15km/h), the verification results are similar for the time neighborhood method and the traditional method, no matter whether the forecasted speed error is large or small. When the typhoon translated with a high speed (larger than 30m/s), the TRAMS model predicted the speed quite well, which also in turn led to the similar error between two different verification methods. The difference in general error statistics between neighborhood and traditional methods mainly exists in the typhoon cases with a moderate translational speed (between 15 and 30 km/h), caused by the obvious underestimate bias of speed forecasted by TRAMS model. Two typical typhoon cases (Mangkhut (2018) and Maria (2018)) were analyzed in greater detail to further illustrate the influence of forecasted speed error on the metric of time neighborhood method.

Despite the essential difference as noted in section 2, the properties of the time neighborhood method can still be explained conveniently based on the view of decomposing the track error into cross-track and along-track components. As shown by the verification of two selected examples in this paper, the contribution of along-track error on absolute track error is reduced in the time neighborhood method under the situation that the cross-track error is small. This is easy to understand because the time-expanded samples will reach the observation center position sooner or later if the moving direction of forecast is consistent with observation (that is, with small cross-track error), and then the along-track error will be effectively compensated.

Under the special scenarios that the small timing error can be ignored, the time neighborhood method proposed in this paper is a supplement of the traditional method, in which the higher time resolution forecast information is taken into account to determine whether a forecast is "good" or "bad". Meanwhile, as the differences between competing model forecasts are lessened in time neighborhood method, the new metric cannot be simply seen as a replacement of the track error from traditional method. The new metric emphasizes more the accuracy of forecasts from different models for a given position.

The new metric can easily be extended to other situations. Mass et al. (2002) presented an interesting example showing less importance of timing than position when a farmer makes a decision about irrigation. Thus, for this farmer, the metric of time neighborhood will be a more satisfactory tool to assess the value of different forecast models. This also indicated the possibility to extend our method to the precipitation forecast, which will provide a more reliable reference of the performance of different models for those users who do not care about the exact time error very much. We suggest the neighborhood method can be applied for model intercomparison studies within a broader context, such as typhoon intensity, temperature at 2m and wind at 10m height, etc.

Some details are simplified for easy implementation in this paper and they still need to be further studied in the future work. For example, the optimal sampling time window is still difficult to be defined objectively. It should be determined according to the specific applications. This may introduce some uncertainty into the verification results. Thus, how to construct a reasonable formula for computing the optimal sampling time window quantitatively based on related factors for landfall typhoon is left as an open question.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant Nos. U1811464 and U2142213), Guangdong Basic and Applied Basic Research Foundation (Grants Nos. 2020A1515110275, 2020A1515110040, 2022A1515011870) and the Special program for innovation and development of China Meteorological Administration (CXFZ2021Z006 and CXFZ2022P026). We thank the two anonymous reviewers and the editors for their comments, which greatly improved this paper.

REFERENCES

Ahijevych, D., E. Gilleland, B. G. Browm, and E. E. Ebert, 2009: Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts. *Wea. Forecasting*, **24**, 1485–1497, https://doi.org/10.1175/2009WAF2222298.1.

Biswas, M. K., D. Stark, and L. Carson, 2018: GFDL vertex tracker users guide version 3.9 a. Developmental Testbed Center, 1–35. [Available online from http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRFv3.9a_ScientificDoc.pdf]

Casati, B., and L. J. Wilson, 2007: A new spatial-scale decomposition of the Brier score: Application to the verification of lightning probability forecasts. *Mon. Wea. Rev.*, **135**, 3052–3069, https://doi.org/10.1175/MWR3442.1.

Chen, G. M., M. Q. Yang, X. P. Zhang, and R. J. Wan, 2020: Verification of tropical cyclone operational forecast in 2019. Technical Report for 52st Session of ESCAP/WMO

Typhoon Committee. Available from http://www.typhooncommittee.org/52nd/6-TC52Technical.html.

Clark, A. J., W. A. Jr. Gallus, and M. L. Weisman, 2010: Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF model simulations and the operational NAM. *Wea. Forecasting*, **25**, 1495–1509, https://doi.org/10.1175/2010WAF2222404.1.

Colle, B. A., C. F. Mass, and D. Ovens, 2001, Evaluation of the timing and strength of MM5 and Eta surface trough passages over the eastern Pacific. *Wea. Forecasting*, **16**,553-572, https://doi.org/10.1175/1520-0434(2001)016<0553:EOTTAS>2.0.CO;2.

Davis, C. A., D. A. Ahijevych, W. Wang, and W. C. Skamarock, 2016: Evaluating medium-range tropical cyclone forecasts in uniform- and variable-resolution global models. *Mon. Wea. Rev.*, **144**, 4141–4160, https://doi.org/10.1175/MWR-D-16-0021.1.

Duc, L., K. Saito, and H. Seko, 2013: Spatial-temporal fractions verification for high-resolution ensemble forecasts. *Tellus A*, **65**, 18171, https://doi.org/10.3402/tellusa.v65i0.18171.

Ebert, E. E., 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, https://doi.org/10.1175/2009WAF2222251.1.

Elsberry, R. L., and J. E. Peak, 1986: An evaluation of tropical cyclone forecast aids based on cross-track and along-track components. *Mon. Wea. Rev.*, **114**, 147–155, https://doi.org/10.1175/1520-0493(1986)114<0147:AEOTCF>2.0.CO;2.

Fang, X. Q., and Y.-H. Kuo, 2015: A new generic method for quantifying the scale predictability of the fractal atmosphere: Applications to model verification. *J. Atmos. Sci.*, **72**, 1667–1688, https://doi.org/10.1175/JAS-D-14-0112.1.

Faure, G., P. Chambon, and P. Brousseau, 2020: Operational implementation of the AROME model in the tropics: Multiscale validation of rainfall forecasts. *Wea. Forecasting*, **35**, 691–710, https://doi.org/10.1175/WAF-D-19-0204.1.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, https://doi.org/10.1175/2009WAF2222269.1.

Huang, B., and X. G. Wang, 2018: On the use of cost-effective valid-Time-Shifting (VTS) method to increase ensemble size in the GFS hybrid 4DEnVar system. *Mon. Wea. Rev.*, **146**, 2973–2998, https://doi.org/10.1175/MWR-D-18-0009.1.

Jenkner, J., C. Frei, and C. Schwierz, 2008: Quantile-based short-range QPF evaluation over Switzerland. *Meteor. Z.*, **17**, 827–848, https://doi.org/10.1127/0941-2948/2008/0344.

Kang, N.-Y., and J. B. Elsner, 2016: Climate mechanism for stronger typhoons in a warmer world. *J. Climate*, **29**, 1051–1057, https://doi.org/10.1175/JCLI-D-15-0585.1.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? The results of two years of real-time numerical weather prediction over the pacific northwest *Bull. Amer. Meteor. Soc.*, **83**, 407–430, https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2.

Mittermaier, M., and N. Roberts, 2010: Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Wea. Forecasting*, **25**, 343–354, https://doi.org/10.1175/2009WAF2222260.1.

Mittermaier, M. P., 2006: Using an intensity-scale technique to assess the added benefit of high-resolution model precipitation forecasts. *Atmospheric Science Letters*, **7**, 36–42, https://doi.org/10.1002/asl.127.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Schwartz, C. S., 2017: A comparison of methods used to populate neighborhood-based contingency tables for high-resolution forecast verification. *Wea. Forecasting*, **32**, 733–741, https://doi.org/10.1175/WAF-D-16-0187.1.

Skok, G., and V. Hladnik, 2018: Verification of gridded wind forecasts in complex alpine terrain: A new wind verification methodology based on the neighborhood approach. *Mon. Wea. Rev.*, **146**, 63–75, https://doi.org/10.1175/MWR-D-16-0471.1.

Stein, J., and F. Stoop, 2019: Neighborhood-based contingency tables including errors compensation. *Mon. Wea. Rev.*, **147**, 329–344, https://doi.org/10.1175/MWR-D-17-0288.1.

Tao, W.-K., and Coauthors, 2011: High-resolution numerical simulation of the extreme rainfall associated with Typhoon Morakot. Part I: Comparing the impact of microphysics and PBL parameterizations with observations. *Terrestrial, Atmospheric and Oceanic Sciences*, **22**, 673–696, https://doi.org/10.3319/TAO.2011.08.26.01(TM.

Theis, S. E., A. Hense, and U. Damrath, 2005: Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteorological Applications*, **12**, 257–268, https://doi.org/10.1017/S1350482705001763.

Tory, K. J., S. S. Chand, R. A. Dare, and J. L. McBride, 2013: The development and assessment of a model-, grid-, and basin-independent tropical cyclone detection scheme. *J. Climate*, **26**, 5493–5507, https://doi.org/10.1175/JCLI-D-12-00510.1.

WMO, 2013: Verification methods for tropical cyclone forecasts. WWEP 2013-7, 89 pp. [Available online from (http://www.wmo.int/pages/prog/arep/wwrp/new/documents/WWRP_2013_7_TC_verification_15_Nov_en.pdf].

Wolff, J. K., M. Harrold, T. Fowler, J. H. Gotway, L. Nance, and B. G. Brown, 2014: Beyond the basics: Evaluating model-based precipitation forecasts using traditional, spatial, and object-based methods. *Wea. Forecasting*, **29**, 1451–1472, https://doi.org/10.1175/WAF-D-13-00135.1.

Xu, D. S., and Coauthors, 2020: Updates in TRAMS 3.0 model version and its verification on typhoon forecast. *Meteorological Monthly*, **46**, 1474–1484, https://doi.org/10.7519/j.issn.1000-0526.2020.11.008. (in Chinese with English abstract)

Xu, Q., L. Wei, H. J. Lu, C. J. Qiu, and Q. Y. Zhao, 2008: Time-expanded sampling for ensemble-based filters: Assimilation experiments with a shallow-water equation model. *J. Geophys. Res.*, **113**, D02114, https://doi.org/10.1029/2007JD008624.

Xu, X. C., Y. B. Yu, and D. J. Zhao, 2009: Variational characteristics of tropical cyclones making landfall in China with different intensity. *Journal of Tropical Meteorology*, **25**, 667–674, https://doi.org/10.3969/j.issn.1004-4965.2009.06.004. (in Chinese with English abstract)

Yu, H., and Coauthors, 2012: Operational tropical cyclone forecast verification practice in the western North Pacific region. *Tropical Cyclone Research and Review*, **1**, 361–372, https://doi.org/10.6057/2012TCRR03.06.

Yu, H., G. M. Chen, and B. Brown, 2013: A new verification measure for tropical cyclone track forecasts and its experimental application. *Tropical Cyclone Research and Review*, **2**(4), 185–195, https://doi.org/10.6057/2013TCRR04.01.

Zhao, Q. Y., Q. Xu, Y. Jin, J. McLay, and C. Reynolds, 2015: Time-expanded sampling for ensemble-based data assimilation applied to conventional and satellite observations. *Wea. Forecasting*, **30**, 855–872, https://doi.org/10.1175/WAF-D-14-00108.1.

Zhu, K. F., Y. Yang, and M. Xue, 2015: Percentile-based neighborhood precipitation verification and its application to a landfalling tropical storm case with radar data assimilation. *Adv. Atmos. Sci.*, **32**, 1449–1459, https://doi.org/10.1007/s00376-015-5023-9.