Development of a Track-Pattern-Based Medium-Range Tropical Cyclone Forecasting System for the Western North Pacific

Hung Ming Cheung, Chang-Hoi Ho, Minhee Chang, Dasol Kim, Jinwon Kim, and Woosuk Choi

a School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea
b National Institute of Meteorological Sciences, Jeju-do, South Korea
c Division of Data Science, University of Suwon, Gyeonggi-do, South Korea

ABSTRACT: Despite tremendous advancements in dynamical models for weather forecasting, statistical models continue to offer various possibilities for tropical cyclone (TC) track forecasting. Herein, a track-pattern-based approach was developed to predict a TC track for a lead time of 6–8 days over the western North Pacific (WNP), utilizing historical tracks in conjunction with dynamical forecasts. It is composed of four main steps: 1) clustering historical tracks similar to that of an operational 5-day forecast in their early phase into track patterns, and calculating the daily mean environmental fields (500-hPa geopotential height and steering flow) associated with each track; 2) deriving the two environmental variables forecasted by dynamical models; 3) evaluating pattern correlation coefficients between the two environmental fields from step 1 and those from dynamical model for a lead time of 6–8 days; and 4) producing the final track forecast based on relative frequency maps obtained from the historical tracks in step 1 and the pattern correlation coefficients obtained from step 3. TCs that formed in the WNP and lasted for at least 7 days, during the 9-yr period 2011–19 were selected to verify the resulting track-pattern-based forecasts. In addition to the performance comparable to dynamical models under certain conditions, the track-pattern-based model is inexpensive, and can consistently produce forecasts over large latitudinal or longitudinal ranges. Machine learning techniques can be implemented to incorporate nonlinearity in the present model for improving medium-range track forecasts.

KEYWORDS: Hurricanes/typhoons; North Pacific Ocean; Statistical forecasting; Operational forecasting; Clustering

1. Introduction

Among various features of tropical cyclones (TCs), the track forecast is of great importance; preparation and mitigation measures could be ineffective without accurate information on future TC locations. A risk analysis of TCs for the Republic of Korea (ROK) demonstrated that TC-induced damages are more sensitive to TC tracks than their intensity and/or size (Nam et al. 2018). A reduced vulnerability to TCs could be attributed to substantial improvements in track forecasting, as suggested in Park et al. (2015). As TCs seriously threaten the coastal and adjacent regions with intense rainfall, strong winds, storm surges, and flash flooding, TC track forecasts are especially crucial for the western North Pacific (WNP) which has a long coastline and large populations. The WNP basin has 31 TCs per year on an average (Chan 2015). During 1979–2019, about 37% of the WNP TCs survived over 6 days (Fig. 1) while 60% of these long-living TCs reach to the north of 30°N and 20% are directed toward the South China Sea (SCS, south of 25°N, west of 120°E). A majority of these long-lasting TCs eventually impact large populations; hence, accurate forecasting of the passage of TCs for a longer forecast lead time is highly beneficial especially for East Asia.

Currently, TC track forecasts for all lead time rely heavily on dynamical models. In the past, dynamical model forecasts were known to contain systematic errors which can lead to large TC track errors (Carr and Elsberry 2000), thus requiring statistical adjustment techniques (Elsberry et al. 1999) or systematic approaches (Carr et al. 2001). Due to substantial improvements (or reduction in error) in track forecasting (Heming 2016; Landsea and Cangialosi 2018) concurrent with advances in computing technology, the medium-range track forecasting (here defined as beyond 5 days) could be made with dynamical model outputs from various institutes, like the European Centre for Medium-Range Weather Forecasts (ECMWF), the National Centers for Environmental Prediction (NCEP), and the Met Office (UKMO). ECMWF uses their Integrated Forecast System to develop several TC forecast products with a forecast lead time of 10 days (Magnusson et al. 2019). In addition, the Hong Kong Observatory (HKO) provides a 9-day TC track probability forecast based on the output of several ensemble models (Wong and Choy 2018).

Statistical models have received less attention since the rise of dynamical models. According to Neumann (1985), statistical models could be divided into three categories: analog, regression equation, and statistical–dynamical models. Analog models find historical tracks with characteristics similar to the current TC. Regression equation models use regression analysis to fit a number of predictors related to current and historical TC locations and movements. Statistical–dynamical models combine “pure” statistical models and forecast fields of dynamical models to produce a final forecast. Statistical models are advantageous over dynamical models in that they can forecast.
TC tracks by consuming far fewer computational resources than dynamical models (Neumann 1985). On the other hand, any environmental variables at disposal could be chosen (Roy and Kovordányi 2012). The climatology and persistence statistical model (CLIPER5; Aberson 1998), developed in 1998, is still used by the National Hurricane Center (NHC) as a benchmark for evaluating track forecast skills of other models during 5 days. The NHC regards CLIPER5 as a good baseline for forecast skill over a season or longer (Cangialosi 2019). Although there are studies intending to develop statistical models for TC track forecasts (e.g., Bessafi et al. 2002; Xu and Neumann 1985; Fraedrich et al. 2003), we believe that these studies do not focus on the medium range.

The track-pattern-based approach is a statistical–dynamical technique that has been used to predict TC activity across a spectrum of time scales, from seasonal to long-term future (Kim et al. 2012; Choi et al. 2016a,b, 2017; Park et al. 2017; Kim and Chan 2018) in different basins. The approach classifies TC tracks into several track patterns (clusters) based on similarities between tracks by various methods such as $k$-means clustering, fuzzy $c$-means clustering, and regression mixture models (Camargo et al. 2007; Kim et al. 2011; Elsner and Liu 2003). Prediction of TC occurrence is made based on a regression between TC occurrence of each track pattern and selected environmental predictors such as sea surface temperature, wind, and vorticity from a dynamical model; predictions of all track patterns are then combined to obtain the final result. Kim et al. (2012) suggested that track-pattern-based models could compete with high-resolution dynamical models in terms of seasonal TC forecasts.

This study aims to explore the possibility of developing a medium-range (6–8 days) TC track forecast system using a track-pattern-based method with an emphasis on forecasting TCs that possibly affect the WNP. The rest of the study is organized as follows: section 2 describes the dataset and methods used, including observation, reanalysis, and model forecast data; section 3 explains the structure of the forecasting system; section 4 presents the verification of the forecasts made by this system, and a comparison with objective deterministic forecasts; Finally, section 5 provides a summary and further discussion.

## 2. Data and methodology

### a. Historical and 5-day forecast TC data

For the historical track selection, best track data during 1979–2019 from the Regional Specialized Meteorological Center (RSMC) Tokyo-Typhoon Center were used. Parameters such as the center position and intensity grade of each track were archived at 6-h intervals. Extratropical cyclones were excluded here. On the other hand, for verification, the actual tracks (latitude and longitude of TC centers at analysis time) from the Korea Meteorological Administration (KMA) were used as a reference for calculating track error.

TCs formed in the WNP were reforecasted using the track-pattern-based forecasting system at multiple initial times on a day for the entire life span of at least 7 or more days. For each of these cases, the genesis position and forecast location over five successive days, and radii of 70% probability circle were collected from the KMA. The forecasting system was validated for the period 2011–19, as the 5-day track forecasts from KMA were available only 2011 onward.

### b. Reanalysis data

As large-scale circulations can modulate the movement of a TC, we utilized the 500-hPa geopotential height (GPH500) and steering flow (STEER) as predictors. These two predictors are highly correlated except for the time when the shear is very high. To compute the environments associated with each historical TC at a particular forecast lead time, reanalysis data from the ERA-Interim of ECMWF (Dee et al. 2011) were analyzed. Data was selected at a horizontal resolution of 0.75° × 0.75° at 6-h intervals.

GPH500 is often used to identify the subtropical high whose expansion or retreat affects the movement of TCs directly (Ho et al. 2004). Winds over a deep layer in the troposphere (STEER, which are calculated as the pressure-weighted and vertically averaged horizontal winds over certain layer of the troposphere) are another indicator of TC movements (Kim et al. 2011). While constant depths are commonly used to calculate the steering flow in climate studies, the use of constant steering layers may be inappropriate for predicting TC motions in weather forecasts. Thus, we utilized the results of Velden (1993), as summarized in Table 1, to approximate the variations in the steering depth and TC intensity. Using a barotropic model, Velden (1993) found that for a specific

![FIG. 1. Distribution of the lifetime of TCs in the western North Pacific during 1979–2019.](image-url)

### Table 1. The tropospheric pressure levels used to calculate steering flow for each tropical cyclone intensity class (Velden 1993).

<table>
<thead>
<tr>
<th>Tropical cyclone intensity class (hPa)</th>
<th>Tropospheric pressure levels (hPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1000</td>
<td>700–850</td>
</tr>
<tr>
<td>990–999</td>
<td>500–850</td>
</tr>
<tr>
<td>980–989</td>
<td>400–850</td>
</tr>
<tr>
<td>970–979</td>
<td>400–850</td>
</tr>
<tr>
<td>960–969</td>
<td>300–850</td>
</tr>
<tr>
<td>950–959</td>
<td>300–850</td>
</tr>
<tr>
<td>940–949</td>
<td>250–850</td>
</tr>
<tr>
<td>&lt;940</td>
<td>200–700</td>
</tr>
</tbody>
</table>
intensity of TC, the track error could be optimized if a specific depth of a steering layer is used. Although his study considers the Atlantic Ocean, the result is similar to that for the Australian region; we adopt this result in our study. As this result did not consider the WNP, the steering flow calculated in this way is not optimal, but can be an improvement from the constant-layer approach. On the other hand, only data at each 50-hPa pressure level was used for efficiency. Because numerical models are known to have large bias in TC intensity (MSLP), we performed a sensitivity test by multiplying the model TC intensity (MSLP) by 0.98, which is equivalent to a low bias of about 20 hPa. The changes in the result due to the change in the model TC are very small (Fig. S7), suggesting that our forecast system is not sensitive to the bias in the model TC intensity.

c. Model forecast data

The global forecast models used in this study include the Climate Forecast System version 2 (CFSv2), Global Ensemble Forecast System (GEFS), and Global Forecast System (GFS), all from the NCEP. See Table 2 for their configurations. CFSv2 and GFS both provide a single simulation, while GEFS provides 21 ensemble members. The mean of all available members in GEFS was computed. The variables used are the same as those listed in section 2b. For the steering flow, when a TC vortex was detected at a certain time step in the model output, the steering depth was selected based on the MSLP of the detected storm vortex. When no vortex was detected, all the layers between 200 and 850 hPa were used. Furthermore, the model data were regridded to the same horizontal grid nest as the reanalysis data.

The track forecast obtained for various dynamical models can be found in web page of THORPEX (The Observing system Research and Predictability Experiment) Interactive Grand Global Ensemble (https://www.cawcr.gov.au/research/cyclone-exchange/ or http://rda.ucar.edu/datasets/ds330.3/). However, CFSv2 is not included, so we applied a detection and tracking algorithm on the three dynamical models for consistency.

d. Statistical analysis

The one-way analysis of variance (ANOVA) was applied to identify the regions in which the differences in the mean environments between track patterns are statistically significant at the 95% confidence level (section 3c). ANOVA tests for differences in the means of two or more independent groups, in contrast to the Student’s t test that tests for only two groups.

The t test was used to check whether the mean track errors between track-pattern-based forecast was significantly larger than dynamical model forecast track. The null hypothesis was that the mean track error of the track-pattern-based forecast is smaller or equal to that of dynamical model track forecast, while the alternative hypothesis is that the mean track error of track-pattern-based forecast is larger than that of the dynamical model track forecast.

e. Detection and tracking of TCs in dynamical model simulations

To compare the performance of our forecasting system against dynamical forecasts, vortex detection and tracking processes are required to obtain a TC track in dynamical model outputs. The detection and tracking algorithm adopted by UKMO (Heming 2017) was applied with some modifications; it uses relative vorticity at 850 hPa (RV850) to track a TC vortex in a model as it helps give a strong signal for the TC center position even at lower TC intensities. The model TC center is designated at the nearest local minimum MSLP field to the highest value of RV850. The detection procedures are as follows (variables in italics in this subsection represent threshold values for the corresponding parameters; those for the three models are listed in Table 3):

1) Analysis time
   - The center of a TC in the model analysis was determined by initiating a search from the observed position of the TC within a radius of $d_0$ for the grid point which contains the highest value of RV850. The nearest grid point with a local minimum MSLP within a radius of $d_{c2p}$ from the highest RV850 point was identified.

### Table 2. Configuration of dynamical models output.

<table>
<thead>
<tr>
<th>Model resolution (0–192 h)</th>
<th>CFSv2</th>
<th>GEFS</th>
<th>GFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model resolution (192–384 h)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model cycle per day</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Output time step</td>
<td>6-hourly</td>
<td>6-hourly</td>
<td>3-hourly (0–192 h)</td>
</tr>
<tr>
<td></td>
<td>12-hourly (192–384 h)</td>
<td>3-hourly</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2. Configuration of dynamical models output.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model resolution (0–192 h)</td>
</tr>
<tr>
<td>Model resolution (192–384 h)</td>
</tr>
<tr>
<td>Model cycle per day</td>
</tr>
<tr>
<td>Output time step</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

### Table 3. The threshold values for vortex detection and tracking algorithm.

<table>
<thead>
<tr>
<th></th>
<th>CFSv2</th>
<th>GEFS</th>
<th>GFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_0$ (km)</td>
<td>220</td>
<td>220</td>
<td>220</td>
</tr>
<tr>
<td>$d_{c2p}$ (km)</td>
<td>500</td>
<td>300</td>
<td>450</td>
</tr>
<tr>
<td>$d_{close}$  (km)</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>$d_1$ (km)</td>
<td>570</td>
<td>300</td>
<td>450</td>
</tr>
<tr>
<td>$d_1'$ (km)</td>
<td>170</td>
<td>150</td>
<td>100</td>
</tr>
<tr>
<td>$\phi$ (°)</td>
<td>25</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>$\zeta$ ($10^{-5}$ s$^{-1}$)</td>
<td>8</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>$p$ (hPa)</td>
<td>1007</td>
<td>1010</td>
<td>1005</td>
</tr>
<tr>
<td>$\xi_{low}$ ($10^{-5}$ s$^{-1}$)</td>
<td>5</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>$p_{low}$ (hPa)</td>
<td>1015</td>
<td>1015</td>
<td>1015</td>
</tr>
<tr>
<td>$d_{c2c}$ (km)</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>
A closed isobar check was performed at a radius of $d_{\text{close}}$ from the grid point with a local minimum MSLP found above. If a closed isobar is not found, we do not consider this low pressure as a TC vortex.

When a vortex was undetectable, the observed TC position was assigned to be the first model TC position.

A $5^\circ \times 5^\circ$ grid centered on the previously determined grid point at the lowest MSLP was interpolated to a $10^\circ \times 10^\circ$ grid to obtain a more precise TC center.

2) Forecast position after 6 h

- The search for the second TC position was similar to that for the analyzed position. It started from the observed TC position, but with a search radius of $d_1$.

3) Forecast position after 12 h and beyond

- The initial search center (the first estimate) was defined by extrapolating the forecast track for the previous two timesteps. Locating the TC position was the same as step 2, but it started at the first estimated position. The search radius was $d_1$.

- When a vortex was undetectable, the search was repeated using the preceding model TC center position as the
starting point and a search radius of \( d_0 \); this was done to prevent missing a vortex when a model vortex accelerated/decelerated abruptly to be considerably far from the first estimated position.

- For GFS, the search distance \( d_1 \) doubled when the forecast interval became 12-hourly after 192 h.

For a lead time beyond 6 h, tracking of a model TC vortex continued if the following criteria were met:

1) RV850 was greater than \( \zeta \) and MSLP was less than \( p \) for vortex detected at higher-latitudes (latitude \( \varphi > \phi \)), and \( \zeta_{\text{low}} \) and \( p_{\text{low}} \) at lower-latitudes (latitude \( \varphi \leq \phi \)). These two thresholds were made to be region-dependent as in certain cases a model TC weakened at an early stage, occurring at lower-latitudes.

2) The model vortex was located south of 45\(^\circ\)N.

3) A decrease in the MSLP of a model TC from a previous time step was \( \leq 20 \text{ hPa} \); this criterion was set up to prevent false identification of a vortex center when two vortices were close to each other in the model.

4) The distance between the model TC center at current and previous timesteps was \( \leq d_{c2c} \) (the distance was double when the forecast interval became 12-hourly after 192 h for GFS).

For GEFS, a TC vortex needed to be detected in at least ten ensemble members to calculate the ensemble mean track. This set of criteria and thresholds could satisfactorily detect and track a model TC in the cases selected here.

3. A new approach to medium-range TC track forecasts

This study aimed to develop a TC track forecast system over the WNP for lead times of 6–8 days. Forecasts were made based on a track-pattern approach; the trajectory of a TC is predicted without specifying the time of arrival at specific locations. Figure 2 shows a conceptual diagram of the system presenting the components for each step and the data flow between blocks or components. The track-pattern-based forecasting system consists of four main blocks: 1) clustering of historical tracks and calculating the daily mean composite environmental fields associated with each track pattern (green block); 2) deriving the daily mean environmental fields in dynamical models (yellow block); 3) evaluating the pattern correlations between composites of environmental fields from observation and dynamical models for lead times of 6–8 days (blue block); and 4) producing the relative frequency maps using the historical tracks and pattern correlation coefficients, and obtaining the final track forecast based on these maps (pink block). Each block is further explained in detail. The TC Kong-rey (2018, Fig. 3a) is used as an example to demonstrate the medium-range TC track forecast made at its formation time.

a. Track patterns and observed composite of environmental fields

1) Selection of historical tracks

Hope and Neumann (1970) developed an analog technique using historical TCs similar to the existing TCs (in terms of geographic location, moving direction and speed) to make forecasts. Akin to this approach, we can utilize the historical tracks with a similar geographic location and moving directions in the earlier stage (first 5 days) of Kong-rey’s track to predict its future stage (beyond 5 days). At the time Kong-rey was analyzed, however, its precise future positions were unknown. We assume that the 5-day forecasts from a forecast institute (chosen from KMA; Fig. 3a) possess reasonable skill and we used it to represent the early trajectory of Kong-rey.

In the following step, an array of six circles were drawn at the current position (at analysis time) as well as at the forecast position for each of the succeeding 5 days. The radii of these six circles...
2019 were included. In total, 72 historical tracks were selected.

The part of a track presented as a dashed line starts from the point closest to the formation position of Kong-rey up to the point of the first occurrence in the sixth circle.

(i) A TC passing through each of the six circles, while it entered into the first circle (500-km radius) and exited from (or dissipated in) the last one (1000-km radius).

(ii) A TC having a lifetime of at least 6 days.

(iii) A TC forming in the period (i.e., month) similar to the test case. The track selection process having a length of three months, centered at the forecast date.

Kong-rey formed on 29 September 2018, so other historical TCs that formed between August 15 and November 13 during 1979–2019 were included. In total, 72 historical tracks were selected.

2) CLUSTERING OF HISTORICAL TRACKS

The collected historical tracks were classified into several track patterns by fuzzy c-means clustering (Bezdek 1981). Fuzzy c-means clustering is similar to k-means clustering, but allows each track to belong to all clusters with varying membership coefficients that indicate how strongly a track belongs to a particular cluster. Each TC track was assigned to a track pattern of which the membership coefficient is the largest, i.e., the hard clusters were created. The membership coefficients were used in calculating pattern correlation coefficient. Kim et al. (2011) further explains the application of this clustering algorithm on TC tracks.

Since our forecast system targets days 6–8 of a TC lifetime, the early stage (before the first occurrence in the 1000 km circle mentioned in step 1) of the historical tracks were not included here. Thus, the classification of a track pattern was based on the part of the track that was close to or which fell within medium range; it was not too sensitive to this procedure because the spread of historical tracks that were enclosed in the first five circles were much smaller than that in or beyond the final circle.

The number of track patterns for each forecast was objectively determined by the Xie–Beni index (Xie and Beni 1991). This index, which could be used for fuzzy clustering, is a ratio of compactness (within a cluster) to separation (between clusters). A small index value suggests that track patterns are well separated and that the track spread within each pattern is small. For each forecast, the index value was calculated for each “number of track pattern” \( n_c \) (where \( 2 \leq n_c \leq \) number of selected historical tracks). The \( n_c \) with the lowest index value was selected for this forecast. The maximum number of allowed track pattern was set to five, considering the limited number of historical tracks that could be selected.

For TC Kong-rey, \( n_c = 2 \) was selected. Figure 4 displays the two track patterns associated with the historical tracks for Kong-rey. The two track patterns selected for TC Kong-rey are defined here as C1 and C2. They differ mainly in geographical location. C1 contained poleward-moving recurving tracks, while C2 contained those directed toward the SCS. Although an apparent spread exists within a track pattern due to a relatively large number of historical tracks, the two track patterns were clearly separated from each other.

3) COMPOSITE OF OBSERVED ENVIRONMENTAL FIELDS FOR EACH TRACK PATTERN

For each selected historical track, the environmental fields on each of the days 6–8\(^1\) were calculated for the particular day in the TC lifetime. The two predictors of the historical TC having the highest membership coefficient for the two track patterns on day 6 are shown in Figs. 5a, b, d, and e. The western edge of the subtropical high (thick solid line) of C1 at 25°N was close to the 130°E meridian, retreating eastward at higher latitudes (Fig. 5a), while that of C2 was around the 120°E meridian (Fig. 5b). The STEER pattern was largely consistent with GPH500. The strongest flow near the deep low pressure region of C1 in Fig. 5a was southwesterly (Fig. 5d). For C2, the magnitudes of steer wind vectors were quite uniform over the entire domain (Fig. 5e).

---

\(^1\) Day 6 is defined as hour 144–167, day 7 is defined as hour 168–191, and so on.
b. Forecast of two environmental fields

The GPH500 and STEER fields in a dynamical model on each of days 6–8 were similarly calculated. Figures 5c and f show the daily mean fields on day 6 of CFSv2. The western edge of the subtropical high was near the 130°E meridian, and the ridge line was close to the southern coast of Japan (Fig. 5c). The dominant steering flow over the East Sea around the 130°E meridian was southerly (Fig. 5f). These two fields were more similar to those of C1 as compared to C2.

c. Pattern correlation between observed and model composite

If the environments associated with a particular track pattern are similar to the predicted environments in which the TC that we wanted to forecast is embedded, it is likely that this TC will follow a similar track pattern. To quantify this similarity, pattern correlations between each observed environmental fields and those in the forecast models on days 6–8 were calculated. Only the regions lying within 10°–40°N, 105°–150°E, in which significant differences in the mean environmental fields exist between any two track patterns (the differences were identified through one-way ANOVA, stippled areas in Fig. 5) were included in the pattern correlation calculations.

Since the magnitudes and physical units of all predictors were diverse, they were first normalized (normalized zonal and meridional wind speed had values ranging from -1 to 1; on the other hand, the remaining variable had values ranging from 0 to 1) and then combined to form a single array. Consequently, each TC and model had one array formed by three variables (the zonal and meridional component of the steering flow were considered as two independent variables). Each TC-model combination carried a pattern correlation coefficient for one day. If a correlation coefficient was negative, it was set to zero because it implied that the two environmental fields differed. It also helped prevent track patterns from cancelling each other while computing the relative frequency map in the subsequent step. Thereafter, we calculated the average correlation coefficient for each track pattern by adjusting the correlation coefficient of each TC using the corresponding membership coefficient (section 3a). This was done to make a track that was not too close to its track pattern, having minimal influence on the correlation for the track pattern it belonged to. Having two track patterns and three forecast days considering Kong-rey, there were six correlation coefficients in total for one model. Wanting to produce one forecast for each TC case, we took a 3-day average, which would become the weighting factor to be used in the subsequent step. Figure 6 shows a schematic diagram depicting the computation of pattern correlation coefficient and weighting factor; only CFSv2 is shown for simplicity, but the same procedure was applied to other models.

d. Track forecast

1) RELATIVE FREQUENCY MAP

The weighting factors represented a likelihood that a TC would follow a particular track pattern. However, only a few scalars were insufficient to quantify the expectations that a TC can occur at a certain point in space. Hence, we combined these weighting factors with the historical tracks of each track pattern to produce a two-dimensional distribution of TC occurrence; the final forecast was made based on this; Fig. 7 shows its computation. The historical tracks of each track pattern (Fig. 4) were binned into a 1° × 1° grid to obtain TC track frequency. A grid value represented the number of historical tracks passing within a circle of 250-km radius centered at the grid box. The binned track frequency was then multiplied by the corresponding weighting factor calculated in the previous
step; this product was called “weighted track frequency.” Finally, an “overall relative frequency map” was obtained by summing all the weighted track frequencies, which was then divided by the total number of historical tracks (from the historical track selection process). The overall relative frequency map could be regarded as a measure of track frequency in climatology, which could then be adjusted by model forecasts through pattern correlation. On the other hand, a “specific relative frequency map” is similar to the overall relative frequency map, but for a track pattern only (color shading in Fig. 8).

2) REPRESENTATIVE CURVE

For each track pattern, a representative curve was constructed by the following procedures:

1) First, starting from the location of TC at analysis time, the longitude of the maximum value for each latitude grid box (row), and the latitude of the maximum value for each longitude grid box (column) on the specific relative frequency map were picked. These were the axes of maximum in a northward and westward direction (assuming that a TC could move in only these two directions before dissipation). The collection of these maximum points was terminated when the maximum value of a row or column was <0.5. The longer axis was selected for the next step.

2) Second, a five-point moving average was applied to the axis of maximum of a specific relative frequency map for smoothing. This smoothed line was the representative curve for the corresponding track pattern.

Consequently, the final track forecast by the track-pattern-based system was selected for a higher score $s$, which was calculated using the following formula:

$$s(m,c) = \frac{\sum_{i=0}^{L(m,c)-1} f[m,x(i),y(i)] \times [L(m,c) - i]}{\sum_{i=1}^{L(m,c)} i}.$$  (1)

where, $s(m,c)$ and $L(m,c)$ are the score and length (in grid point) of the representative curve associated with the input model $m$ and track pattern $c$, respectively; and $f[m,x(i),y(i)]$ is

---

**FIG. 6.** Schematic diagram showing the computation of pattern correlation coefficient and weighting factor. CFSv2 is used as an example.
the value of the overall relative frequency map at a grid point \([x(i), y(i)]\), which is passed through by the representative curve of track pattern \(c\) at point \(i\). The formula emphasizes the latter part of a representative curve (the second term in the numerator).

4. Verification

Here, the performance of this forecasting system was evaluated and was compared to those of the global forecast models. For verification, 132 TC cases that affected the WNP for 2011–19 were selected. These TCs survived for at least 7 days and had more than one historical track selected for the historical track selection process. Among the 132 cases, the maximum and minimum number of selected historical tracks were 93 and 3, respectively, while the median was 27. Moreover, about 44% of the cases utilized two track patterns in the forecast, 17% utilized three, and 39% utilized four. On the other hand, there were 17 cases with zero or one selected historical track, and they were coming from 6 TCs (Bopha in 2012, Hagupit in 2014, In-fa in 2015, Lionrock in 2016, Kai-tak and Noru in 2017). These TCs existed in unusual regions or had unusual trajectories, making the selection of historical tracks difficult.

The track error of a forecast is conventionally measured by the great-circle distance between the actual and forecasted tracks at the same time step. Nevertheless, the exact time of occurrence of predicted TCs were not specified in the track-pattern-based forecast. Alternatively, the cross-track error (the shortest distance at a particular time step between the actual track and the representative curve) was used as track error. The track error of a track detected from dynamical model output using a vortex tracking algorithm was similarly

![FIG. 7. Schematic diagram showing the computation of a relative frequency map.](image)
calculated. Due to a large gap between two 6-hourly data points on a detected track, the track was interpolated to a higher-resolution line with 100 points; an example using TC Kong-rey is shown in Fig. 9. Powell and Aberson (2001) also use interpolation to forecast track to increase the accuracy of landfall location. The black markers are the TC positions at the first time step on day 6 (asterisk) and 7 (triangle) of the actual track. The red (track-pattern-based forecast) and blue markers (dynamical model forecast) are the forecasted TC positions nearest to the corresponding black markers. As the weighting factors between the forecasts using different dynamical models are similar, the three track-pattern-based forecasts produced are also quite similar.

The daily mean track errors between the forecasts made by the track-pattern-based system were first compared (Fig. 10). We used $E[A]$ to represent the daily mean track errors of a type of forecast $A$ (e.g., $E[\text{TP-GEFS}]$ indicates the errors of a track-pattern-based forecast using GEFS as an input model; while $E[\text{VT}]$ indicates the errors of all dynamical models). For a homogeneous sample comparison, only cases existing in all three groups were considered. Note that the pattern correlation calculations only focused on the days 6–8; hence, the forecast was accurate for this period, and a discussion on the relatively large errors within the first 5 days are beyond the scope of this study. For either day, the distributions of track error between the three model inputs were quite similar. The $E[\text{TP-GFS}]$ had slightly lower distributions on all lead days.

One of the track-pattern-based forecasts was compared with the track error of dynamical models ($E[\text{VT}]$). The $E[\text{TP-GFS}]$ was selected as it performed slightly better than the others despite small differences among the performances. Figure 11 shows that $E[\text{TP-GFS}]$ was larger than $E[\text{VT-GEFS}]$ and $E[\text{VT-GFS}]$ on day 6, in terms of mean, median, and interquartile range (IQR), but was comparable to $E[\text{VT-CFSv2}]$. On days 7

---

**Fig. 9.** The track-pattern-based forecast (red dashed line) using (a) CFSv2, (b) GEFS, and (c) GFS for input data and the respective model track (blue dashed line) for TC Kong-rey. The black line indicates the actual track. The black star and triangle represent the location of the TC on days 6 and 7 of its lifetime, while the red and blue markers represent the locations nearest from the black to red and blue line, respectively. The 5 gray circles represent the 70% probability circle from the KMA 5-day track forecast as in Fig. 3a.

**Fig. 10.** Homogeneous comparison of the daily mean track error (km) of the representative curve in the pattern-based forecast using various dynamical models as input. A cross inside a box represents the mean value, while a ring represents an outlier (defined as the data points that fall above the 75th percentile plus 1.5 interquartile range). The numbers at the top denote the number of samples included on each day.
and 8, an increasing trend of error distribution was observed across all groups. The mean of $E_{[\text{TP-GFS}]}$ was significantly larger than that of both $E_{[\text{VT-GEFS}]}$ and $E_{[\text{VT-GFS}]}$ at a 99% confidence level on days 6–8 ($E_{[\text{VT-GEFS}]}$ at a 95% confidence level on day 8) (Table 4).

We looked for the situations in which the track-pattern-based forecast has performance comparable to dynamical models. Here, we considered coexisting TCs having their centers west of 140°E at the analysis time. Coexisting TCs are defined as two or more TCs existing simultaneously, and when the distance between their TC centers at the time they first became coexisting is less than 20°. When two TCs are in close proximity to each other, the movement of one TC may be influenced by the circulation of the other; hence, the large-scale steering flow cannot fully explain the TC movement. From Fig. 12, it can be seen that $E_{[\text{TP-GFS}]}$ was largely reduced on days 6–8. The distribution of $E_{[\text{TP-GFS}]}$ varied minimally from day 6 to day 8. On the other hand, $E_{[\text{VT}]}$ increased gradually from day 6 to day 8 (except for $E_{[\text{VT-CFSv2}]}$ on day 8). Consequently, only $E_{[\text{TP-GFS}]}$ on day 6 was significantly larger than $E_{[\text{VT-GEFS}]}$ and $E_{[\text{VT-GFS}]}$ (Table 5). The result for all coexisting cases were also checked; similar results were obtained but $E_{[\text{TP-GFS}]}$ was not significantly larger than any $E_{[\text{VT}]}$ (Table S1, Fig. S4 in the online supplemental material). For cases not coexisting, $E_{[\text{TP-GFS}]}$ was significantly larger than $E_{[\text{VT-GEFS}]}$ and $E_{[\text{VT-GFS}]}$ at a 99% confidence level on all three days (Table S2, Fig. S5).

In addition, thermodynamic variables, which are usually considered to be more relevant to variations in TC intensity, could also play a role in TC motion due to the relationship between TC intensity and steering layer as stated earlier. We tried to include such a factor by including relative humidity at 300, 500, and 850 hPa on top of GPH500 and STEER. Although the pattern correlation coefficients differed from the original predictor set, it can be seen that the overall performances of the forecast using these two predictor sets were similar (Fig. S6).

### Table 4. The p values of a one-tailed Welch’s t test for comparing the track errors of the track-pattern-based forecast ($E_{[\text{TP-GFS}]}$) and each of the dynamical model ($E_{[\text{VT}]}$).
The asterisk and pound signs indicate that $E_{[\text{TP-GFS}]}$ is significantly larger than a particular $E_{[\text{VT}]}$ at the 99% and 95% confidence level, respectively.

<table>
<thead>
<tr>
<th>Lead time</th>
<th>VT-CFSv2</th>
<th>VT-GEFS</th>
<th>VT-GFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 6</td>
<td>0.40</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
<tr>
<td>Day 7</td>
<td>0.34</td>
<td>0.01*</td>
<td>0.00*</td>
</tr>
<tr>
<td>Day 8</td>
<td>0.42</td>
<td>0.04*</td>
<td>0.00*</td>
</tr>
</tbody>
</table>

### 5. Summary and discussion

This study presents a medium-range (6–8 days) TC track forecasting system for the WNP using a track-pattern-based approach. A TC track is forecasted based on historical TC track patterns and the similarity between observed and dynamically forecasted fields associated with individual track patterns. Although this new forecast approach does not correct TC track forecast errors as compared to dynamical models for a lead time of 6–8 days, it performs compatibly to the coexisting TCs (observed at analysis time) to the west of 140°E on days 7 and 8.

An advantage of the track-pattern-based system is its ability to consistently make forecasts over a large latitudinal and longitudinal ranges. Sometimes, a TC track detected in dynamical model outputs does not dissipate at similar locations to the actual TC. In such a case, the detected track could be significantly shorter than the actual one (e.g., the track detected in CFSv2 for TC Bolaven in 2012, Fig. 13). This could arise from the incapability of a vortex tracking algorithm in tracking the TC, or dynamical model errors in simulating a TC as it moves to midlatitudes. This problem, however, does not occur on a representative track because the latitudinal/longitudinal range of the selected historical tracks could be broad enough to cover the length of the actual track. Such an issue with dynamical models also implicitly contributes to their errors as in Figs. 11 and 12. When a detected track is significantly shorter than the actual one, the shortest distance is always calculated between varying points on the actual track to the final point on the detected track. Thus, the error could grow with time to become significantly large as the actual TC moves to higher latitudes, e.g., in Fig. 13, the point on the model track nearest to TC Bolaven on days 6–8 (indicated by the three black markers) are all at the tail of the model track (indicated by the blue square). The track-pattern-based forecasting is also fast and computationally efficient: with a decent computer [CPU: Intel(R) Xeon(R) Gold 5118; RAM: Hynix DDR4 16 GB 2666 MHz × 8], the average running time for the forecasting system is 242 s (~4 min) per case, ranging from 84 s (~1.4 min) to 762 s (~12.7 min).

This forecasting system has some limitation due to its reliance on historical data which may inevitably induce the track error. The grid values of the specific relative frequency map

**Fig. 11.** As in Fig. 10, but for dynamical models (vortex tracking, in blue) and representative curve in the pattern-based forecast of a selected model as input (orange).
primarily rely on the spatial distribution of track frequency divided by the number of historical tracks in each track pattern; if the track frequency over an area is too small while the total number of historical tracks is large, the specific relative frequency cannot be high in the region regardless of the weighting factors. In other words, if a forecasted TC moves into a region where such occurrences are historically rare, a representative curve in proximity would not be selected as a final track forecast due to a low score as calculated by Eq. (1), even if the weighting factor of the track pattern associated with this representative curve is high. Simultaneously, the specific relative frequency over a region is high where historical tracks are dense as long as the associated weighting factors are not significantly small. A number of cases exist in which a forecasted TC moved to a region with a rare occurrence of TCs; if the final track forecast closely followed the axis of maximum weighted track frequency, track error could be noticeable. In addition, if historical track is lacking in a forecast, it implies that the forecasted TC is a rare event. A track forecast using climatology based on a small sample size would be highly uncertain.

In addition to the track error discussed above, other possible sources of errors exist: 1) an inaccurate 5-day track forecast; it affects the selected historical TC tracks to be selected. If the track direction of the 5-day forecast deviates significantly from that of the actual track, the historical tracks that are similar to the actual track could possibly be excluded (incidentally, among the 132 cases analyzed here, all of the actual tracks pass through the circle centered at the fifth forecasted TC locations); 2) an inaccurate model prediction of environments in the medium range; it could lead to erroneous pattern coefficients and weighting factors, affecting the magnitude of the weighted track frequency of any track patterns. An error could eventually arise in the final track forecast.

An issue about environmental fields needs to be addressed. The raw GPH500 and wind fields from dynamical models and reanalysis were used, so signatures of TC remained in these fields. The current results are still valid without removal of TC vortices because the similarity between the forecast and reanalysis fields are still correctly evaluated. However, the location and size of TCs in the forecast and reanalysis fields have to be aligned for a large pattern correlation coefficient; this can be difficult at forecast lead time. Moreover, the influence of the background environment on TC movements is not fully utilized in such cases. As a result, it is more desirable to perform removal of TC vortices from all forecast and reanalysis fields before carrying out the pattern correlation calculations.

The track-pattern-based model can be systematically biased toward slow-moving TCs. During the historical track selection, only TCs that last at least 6 days are retained for further computations. They are likely slow-moving ones because fast-moving storms can move to the higher latitudes and undergo extratropical transition or dissipate early in their lifetime. Hence, TC predictions based on this track-pattern-based model are expected to be based on historical TCs of slower moving speed. This issue should be taken into consideration in TC forecasts using the track-pattern-based model.

From the verification results, it is hardly arguable that this track-pattern-based forecasting system could offer any breakthrough in the medium-range TC track forecasting with the current settings. Machine learning techniques will be considered for further improvements in track forecast: as a subset of artificial intelligence, machine learning could help computer systems acquire an ability to obtain information from given dataset and perform tasks independently; artificial neural network, which is a kind of supervised learning algorithm, is a well-known example. A significant strength of a neural network is its ability to model a non-linear relationship or system, e.g., the complex atmospheric physical processes. Over the last decade, a number of studies have utilized various neural networks for short-term TC track forecast (e.g., Alemany et al. 2019; Kim et al. 2018). For our case, since all components of the track-pattern-based model are linear, replacing some of them with nonlinear machine learning method can possibly improve the model. For example, a convolutional neural network, which is capable of extracting spatial patterns on an image, can be employed to produce weighting factors for each track pattern.


C H E U N G E T A L . 1517

Acknowledgments. This work was funded by the Korea Meteorological Administration Research and Development Program under Grant KMI2020-00610. The work by Jinwon Kim was supported by the NIMS/KMA (NIMS-2016-3100). We thank the National Typhoon Center of KMA for providing 5-day TC forecasts. The actual track data of KMA are obtained at https://www.kma.go.kr/data/typhoonData/typInfoTYList.do?pgmNo=689 (in Korean only), except for that of 2011, obtainable through the National Typhoon Center of KMA. The work by Jinwon Kim was supported by the NIMS/KMA (NIMS-2016-3100). We thank the National Typhoon Center of KMA for providing 5-day TC forecasts. The ERA-Interim data of ECMWF are available at https://www.ecmwf.int/data sets/data/interim-full eg/besttrack.html. The CFSv2 model data are available at https://www.ncei.noaa.gov/data/climate-forecast-system/access/operational-9-month-forecast/6-hourly-by-pressure. The GEFS model data are available at https://www.ncdc.noaa.gov/has/HAS.FileAppRouter?datasetname=GFS&subqueryby=STATION&appname=&outdest=FILE. The GFS model data are available at https://www.ncdc.noaa.gov/has/HAS.FileAppRouter?datasetname=GFS3&subqueryby=STATION&appname=&outdest=FILE.

REFERENCE


