An Alternative Similar Tropical Cyclone Identification Algorithm for Statistical TC Rainfall Prediction

J. A. Hokson[*], and S. Kanae[*]

[*]Dept. of Civil and Environmental Engineering, Tokyo Institute of Civil Engineering

Corresponding author: Jose Angelo Hokson (hokson.jaa@gmail.com)

Key Points:

- Statistical tropical cyclone rainfall prediction leverages distance-based algorithms to identify and use similar past events.
- Proposed in this study is the use of Sinkhorn Distance as a novel measure of TC similarity in rainfall prediction.
- Incorporating Sinkhorn Distance improves TC rainfall prediction accuracy, offering an alternative similarity measure.

Abstract

There is a need to improve the prediction of Tropical Cyclone (TC) rainfall as climate change has led to increased TC rainfall rates. Enhanced reliability in predicting TC tracks has paved the way for statistical methodologies to utilize them in estimating current TC rainfall, achieved by identifying similar past TC tracks and obtaining their corresponding rainfall data. The widely used Fuzzy C Means (FCM) clustering algorithm, though popular, has limitations stemming from its clustering-centric design, hindering its ability to pinpoint the most appropriate similar TCs. Our study introduces the Sinkhorn Distance as a novel measure of TC similarity in rainfall prediction. Our findings indicate that the incorporation of Sinkhorn Distance significantly enhances the accuracy of TC rainfall predictions across WNP. When compared to the conventional approach using FCM, our Sinkhorn Distance-based methodology consistently yields better results, as demonstrated by metrics like RMSE and correlation coefficients. Collectively, the inclusion of Sinkhorn Distance stands as a valuable addition to our toolkit for discerning similar TC tracks, thus elevating the precision of TC rainfall predictions. With ongoing advancements in statistical and AI techniques, we anticipate even more refined approaches to further enhance our predictive capabilities. This study represents a leap forward in meeting the critical need for more accurate TC rainfall forecasts in the WNP Region.

1 Introduction

In recent years, the world has witnessed an alarming increase in the frequency and intensity of extreme weather events (IPCC 2021), and tropical cyclones (TCs) stand out as one of the most devastating natural phenomena affecting coastal regions (Gori et al., 2022; Lee et al., 2019; Wang et al., 2023). The devastating impact of TC rainfall is well-documented, causing severe flooding, property damage, and loss of lives (Tu et al., 2021). Predicting the rainfall associated with these powerful TCs is of utmost importance for enhancing disaster preparedness, risk mitigation, and timely response strategies.

Numerical weather prediction (NWP) model-based methods for TC rainfall prediction have made significant progress (Luitel et al., 2018; Ren et al., 2018). However, accurately predicting TC rainfall remains challenging due to the complexity and non-linearity of atmospheric processes (Luitel et al., 2018; Ren et al., 2018). Moreover, these NWP-based methodologies are computationally expensive, demanding substantial resources (Hokson & Kanae, in press-a). To address these issues, statistical-based methodologies have been developed as a complementary measure to conventional methods.

Statistical-based methodologies rest on the notion that past weather events have a high likelihood of recurring in the present or future (Bagtasa, 2021). By identifying similar historical TCs, these methods enable the prediction of a TC's rainfall. Leveraging comprehensive historical TC rainfall data, these methodologies provide valuable insights into the fundamental patterns and relationships that govern rainfall behavior during cyclones. A significant advantage of statistical approaches lies in their efficiency, as they often demand fewer computational resources, making them accessible and practical for countries and organizations with limited resources.

Recently, there has been a notable enhancement in the precision and reliability of TC track predictions (Li et al., 2016; Kim et al., 2019). This progress has led to the adoption of various statistical methodologies that leverage TC tracks to establish similarity between current/future TCs and past TCs. This approach is grounded in the concept that TCs exhibiting similar tracks tend to generate akin rainfall patterns. This is attributed to the shared influence of factors such as TC intensity, location relative to landmass, as well as temperature and humidity, as noted by Hokson and Kanae (in press-a). Many studies have capitalized on these principles, including works by Ren et al. (2018), Kim et al. (2019), Kim et al. (2020), Bagtasa (2021, 2022), Hokson and Kanae (in press-a, in press-b), as well as Wang et al. (2023).

In identifying similar TC tracks, and thus similar TCs, for the statistical prediction of rainfall, researchers employ a distance-based similarity measure. Among the methods utilized in previous studies (Kim et al., 2019; Kim et al., 2020; Hokson & Kanae, in press-a, in press-b; Wang et al., 2023) is the Fuzzy C Means (FCM) clustering algorithm. FCM uses a membership coefficient as a similarity index between a target TC and other TC, allowing it to identify various patterns, even those with irregular shapes. Moreover, FCM exhibits computational efficiency, making it a good choice for those with limited resources. However, certain limitations exist, such as the requirement for equal-length data and the dependence on the number and location of cluster centers. If the cluster centers are not adequately optimized (e.g., centers are close to each other), FCM may fail to identify the most similar TCs. In light of these drawbacks, it becomes crucial to explore alternative distance-based similarity measures to enhance the effectiveness and robustness of TC rainfall predictions.

The Sinkhorn Distance (Cuturi, 2013) is one possible distance-based similarity measure we can use in identifying similar TCs for the statistical prediction of TC rainfall. It compares probability distributions and handles large-scale datasets with complexity and uncertainty, making it popular in AI and machine learning research. Unlike FCM, it doesn't need equal-length data and allows direct similarity checks between two TCs without involving other real or arbitrary TCs. In this study, we explore the potential of the Sinkhorn Distance as a similarity measure for identifying similar TCs in the statistical prediction of TC rainfall. To assess its accuracy in rainfall prediction, we employ various statistical measures. We establish the methodology utilizing FCM as the reference approach and compare the results obtained using the Sinkhorn Distance. By doing so, we aim to determine whether the Sinkhorn Distance could serve as a promising alternative to FCM in improving the accuracy of TC rainfall predictions.

This study is part of an ongoing effort to enhance our methodologies for the statistical prediction of TC rainfall. In a previous study (Hokson & Kanae, in press-b), we investigated the utilization of additional along-track variables, alongside the TC track, to identify similar TCs. We discovered that these additional variables had minimal impact on TC rainfall prediction accuracies. In another previous study (Hokson & Kanae, in press-a), we proposed a novel constraint involving the TC central pressure at selected locations, which yielded significant improvements in our TC rainfall prediction accuracy. These findings highlight the importance of exploring innovative approaches to optimize our predictions.

This rest of this paper is organized as follows. Section 2 describes the data and the study area. Section 3 discusses the methodology, including the proposed use of Sinkhorn Distance Algorithm. Section 4 presents and discusses the result. Section 5 offers a summary and conclusion.

2 Data and study area

2.1 Data

In this study, we utilized twodatasets for the statistical prediction of TC rainfall. The first dataset, RSMC Best Track Dataset (https://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/trackarchives.html), provided 6-hourly TC tracks. We downloaded the second dataset, APHRODITE Monsoon Asia Precipitation data V1101 and V1101EX_R1 (Yatagai et al., 2012), to obtain 0.25° daily rainfall data. These datasets played a crucial role in our investigation of TC rainfall prediction.

2.2 Study area

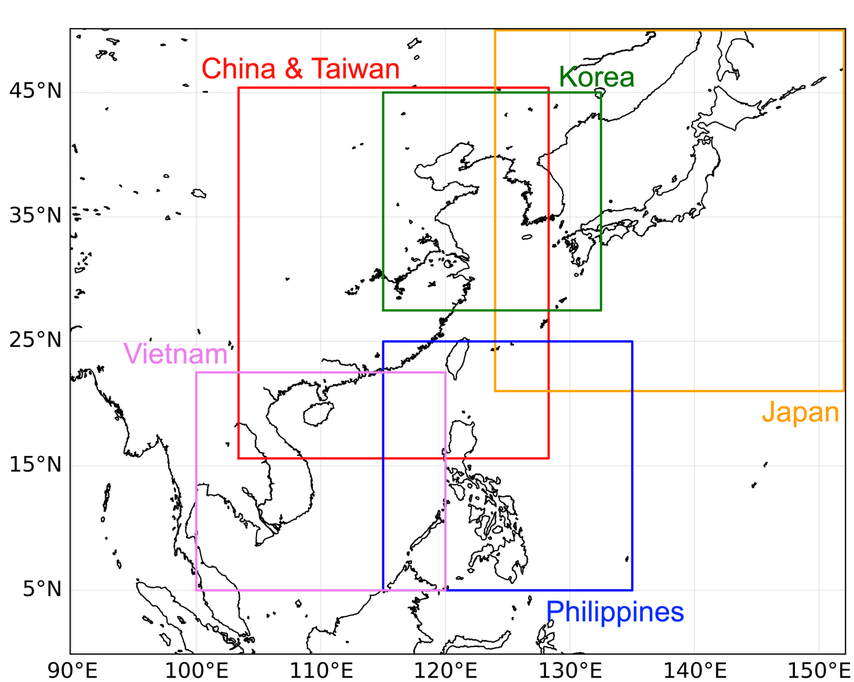

Figure 1. The five areas of simulations. Adapted – with modifications – from Magee et al. (2021).

The Western North Pacific (WNP) Region is recognized as the world's most active tropical cyclone basin, witnessing an average of 26 typhoons annually (Lee et al., 2020). These cyclones impact hundreds of thousands to millions of people every year in the region, underscoring the critical importance of accurately predicting TC rainfall. In contrast to regional methods, our study adopts a country-specific approach, focusing on individual predictions for each country (Figure 1) in WNP. The countries under examination include China and Taiwan (15.6° - 45.4° N, 128.3° - 103.4° E), Japan (21.0° - 50.0° N, 124.0° - 152.0° E), Korea (27.5° - 45.0° N, 115.0° - 132.5° E), Philippines (5.0° - 25.0° E, 115.0° - 135.0° E), and Vietnam (5.0° - 22.5° N, 120.0° - 100.0° E). This decision was driven by the need for more localized and precise forecasts, a crucial aspect in enhancing disaster preparedness.

3 Statistical prediction of TC rainfall

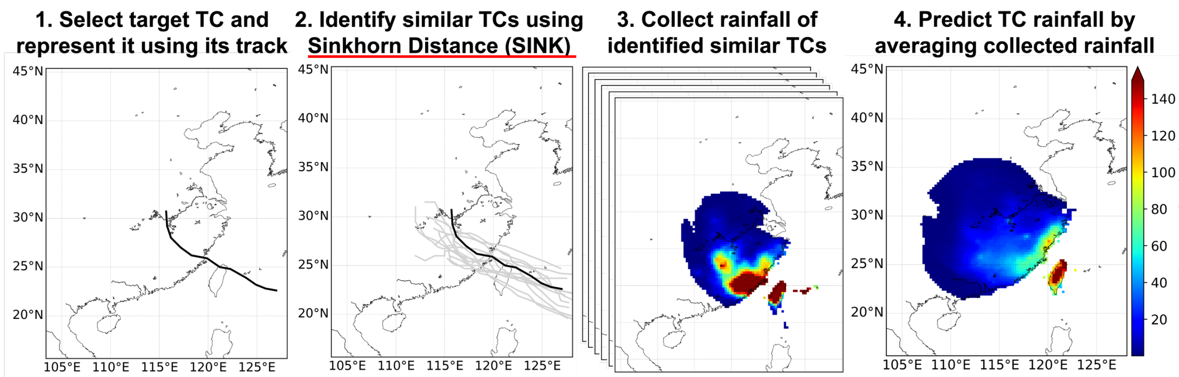

Figure 2. Four-step methodology for predicting TC for every TC. Step 2 uses Sinkhorn Distance instead of the Fuzzy C Means (FCM). The area of simulation for China and Taiwan is used as example.

The prediction of rainfall for each target TC involves a four-step process (Figure 2), adapted from previous studies (Kim et al., 2019; Kim et al., 2020; Hokson & Kanae, in press-b; Wang et al., 2023) with a modification in step 2. First, a target TC is selected and represented by its 6-hourly positions (track) within the area of simulation. Second, similar TCs are identified using the Sinkhorn Distance (Section 3.1). Third, the rainfall values (Section 3.2) of all identified similar TCs are collected. Finally, the prediction is obtained by calculating the simple average of the collected rainfall values.

3.1 Identification of similar TCs using Sinkhorn distance

To identify similar TCs, we propose the use of Sinkhorn Distance. It is derived from the optimal transport theory, which studies the most efficient way to transform one distribution into another. For this study, probability distributions are represented by TC track positions.

The Sinkhorn Distance algorithm for computing distance between two tracks, as described by Cuturi (2013) is given as follow:

1. Define the track data as:

$$\boldsymbol{A} = [(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)] \text{ with } n \text{ elements}$$
$$\boldsymbol{B} = [(u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m)] \text{ with } m \text{ elements}$$

   where x and u are longitude, and y and v are latitude.
2. Define the cost matrix (based on Euclidean distance)

$$C_{ij} = \sqrt{(x_i - u_j)^2 + (y_i - v_j)^2}$$

where C_{ij} is the pairwise distance between point i in track $\boldsymbol{A}$ and point j in track $\boldsymbol{B}$.

3. Initialize the scaling/weighting factors:
   $\boldsymbol{g} = [1, 1, \dots, 1]$ with n elements for points in track $\boldsymbol{A}$
   $\boldsymbol{h} = [1, 1, \dots, 1]$ with m elements for points in track $\boldsymbol{B}$

4. Perform Sinkhorn iterations: (repeat until convergence or a maximum number of iterations is achieved)
   a. Update the scaling factors

$$g_i = \frac{1}{\sum_{j=1}^{m} h_j^{-\frac{C_{ij}}{\varepsilon}}} \text{ for } i = 1, 2, \dots n$$

$$h_j = \frac{1}{\sum_{i=1}^{n} g_j^{-\frac{C_{ij}}{\varepsilon}}} \text{ for } j = 1, 2, \dots m$$

   where ε is the regularization parameter controlling the trade-off between accuracy and computational stability in Sinkhorn iterations.
   b. Normalize the scaling factors

$$\boldsymbol{g} = \frac{\boldsymbol{g}}{\sum_{i=1}^{n} g_i}$$

$$\boldsymbol{h} = \frac{\boldsymbol{h}}{\sum_{i=1}^{m} h_i}$$

5. Compute the optimal transport plan P_{ij}.

$$P_{ij} = g_i^{-\frac{C_{ij}}{\varepsilon}} h_j$$

   where P_{ij} represents the probability of transporting mass from point i in track $\boldsymbol{A}$ to point j in track $\boldsymbol{B}$. Values for P_{ij} are within the range 0 – 1.
6. Calculate the Sinkhorn Distance values.

$$sdist = \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} C_{ij}$$

Using the calculated Sinkhorn Distance values of similar TCs per each target TC, the similar TCs are ranked based on similarity.

To identify the optimal number of similar typhoons, n_opt, to be used for TC rainfall prediction, the prediction error values across different numbers of similar TCs are computed. The number of similar typhoons with the least prediction error is considered the optimal number of similar typhoons. All similar TCs that are part of the n_opt most similar typhoons are used for rainfall prediction.

3.2 Computation of rainfall values

We used the conventional circles of 500 km radius (Guzman and Jiang 2021) to extract TC rainfall at each of its 6-hourly positions. These extracted values were then summed up to compute the distributed total rainfall for each TC. Subsequently, the total rainfall values of all the similar typhoons (i.e., the n_opt most similar TCs) were collected and averaged to obtain the predicted rainfall value of each target TC.

4 Results

4.1 Identified similar TCs based on track similarity

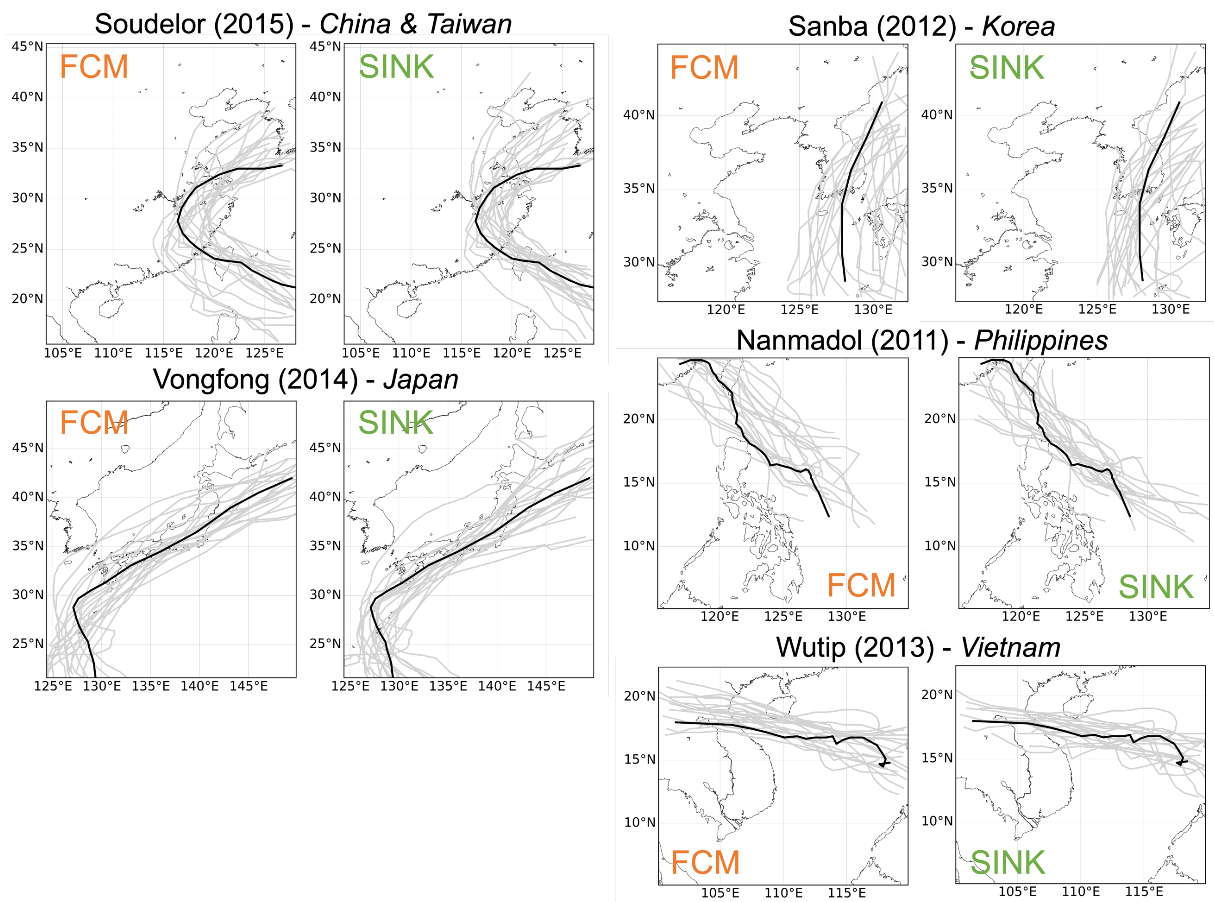

Figure 3. Top 20 most similar TCs identified through SINK and FCM for the five typhoon cases in five simulation areas. Black represents the target TC track, and gray represents identidied similar TC tracks.

To illustrate the effectiveness of Sinkhorn Distance, hereinafter referred to as SINK, we conducted an analysis employing five distinct Tropical Cyclones (TCs), each serving as a representative sample of specific simulation areas. These TCs encompass Soudelor (2015) for China & Taiwan, Vongfong (2014) for Japan, Sanba (2012) for

Korea, Nesat (2011) for the Philippines, and Wutip (2013) for Vietnam. In order to demonstrate the capacity of SINK in identifying similar TCs, we present the top 20 TCs that exhibit the highest similarity as identified by the algorithm. Overall, SINK identifies similar TC tracks wells, as portrayed in Figure 3.

In a comparative context against the conventional FCM approach, it becomes evident that TCs identified as similar by SINK exhibit an enhanced spatial proximity to the target TCs. Specifically, in the case of China & Taiwan, TCs deemed similar to Soudelor (2015) demonstrate closer alignment in the southern and eastern sectors, albeit without the same degree of proximity in the northern trajectory. Similarly, for Japan, TCs identified as similar to Vongfong (2014) display notable spatial closeness, with the exception being the northeastern portion of the target TC track. While exploring the results for Korea, it is noteworthy that SINK identifies analogous TCs primarily closer to the east of the target, although some tracks exhibit distinctive and unconventional shapes. Furthermore, when considering the Philippines, TCs identified as similar to Nanmadol (2011) vividly illustrate the proximity of TC tracks, particularly in the eastern vicinity of the target TC. Lastly, for Vietnam, SINK identifies a greater number of similar TCs situated to the south of the target, consequently leading to a more centralized alignment in comparison to the FCM methodology. These findings highlight the difference in TC similar identification between the new SINK method and the conventional FCM approach.

4.2 Optimal numbers of similar TCs for TC rainfall prediction and performance of SINK

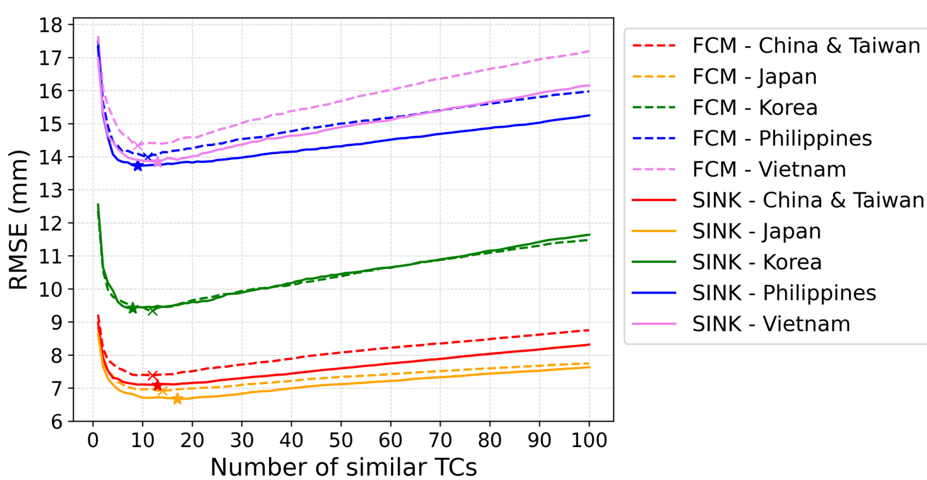

Figure 4. Average rainfall prediction error of TCs across different number of similar TCs for the five simulation areas. Only values inside rainfall calculation areas (shown in the Appendix) area are considered in the prediction error. Values in bold red represent the best values.

As in previous studies (Kim et al., 2019; Kim et al., 2020; Hokson & Kanae, in press-a, in press-b; Wang et al., 2023), we determined the optimal numbers of similar TCs using TC rainfall prediction error. In such, the numbers of similar TCs with the least

RMSE values for the five simulation areas are considered optimal numbers of similar TCs for simple averaging of rainfall for prediction of rainfall.

The RMSE values for all simulation areas for any number of similar TCs are between 6 mm to 18 mm for SINK (Figure 4). This same range is applicable to FCM, which is also computed for comparative and in-depth analysis. Notably, these values are markedly lower than those reported by Kim et al. (2019) and Kim et al. (2020), possibly due to the larger dataset of past TCs employed in this study. Moreover, they are also lower than those indicated by Hokson & Kanae (2023a and 2023b), as well as Wang et al. (2023), which may be attributed to the more localized approach adopted in here. Comparing SINK with FCM for all simulation areas except for Korea, RMSE values are lower for SINK than FCM across different number of similar TCs. For Korea, the RMSE values for SINK are almost the same with FCM, sometimes higher sometimes lower.

Table 1. Minimum RMSE values and corresponding optimal no. of similar TCs based on Figure 4 for the five simulation areas.

Simulation Area	No. of target TCs	RMSE [mm]		Optimal no. of similar TCs	
		FCM	SINK	FCM	SINK
China & Taiwan	768	7.38	**7.09**	12	13
Japan	672	6.92	**6.68**	14	17
Korea	234	**9.34**	9.42	12	8
Philippines	759	13.99	**13.73**	11	9
Vietnam	405	14.34	**13.84**	9	13

The minimum values of RMSE are in between 8 to 17 number of similar TCs and ranges 6 to 15 mm (Figure 4 and Table 1). The optimal numbers of similar TCs for SINK are 13 for China & Taiwan, 17 for Japan, 8 for Korea, 9 for the Philippines, and 13 for Vietnam based on the least RMSE. For FCM, the optimal numbers of similar TCs are 12 for China & Taiwan, 14 for Japan, 12 for Korea, 11 for the Philippines, and 9 for Vietnam.

4.3 Rainfall prediction

Using the optimal number of similar TCs determined in Section 4.2, the TC rainfall are predicted. In this analysis, we continue to utilize the five test TCs introduced in Section 4.1, showcasing the influence of the similar TCs identified through SINK.

We replotted the similar TCs shown and discussed in Section 4.1 (Figure 3) to reflect the optimal number of TCs for each simulation area (Figure 5). Generally, SINK identified similar TCs closer to the target than FCM, like those involving top 20 most similar. Compared to those with top 20 most similar TCs however, the top 8 most similar TCs for Korea identified through SINK are mostly on the east side of the target TC.

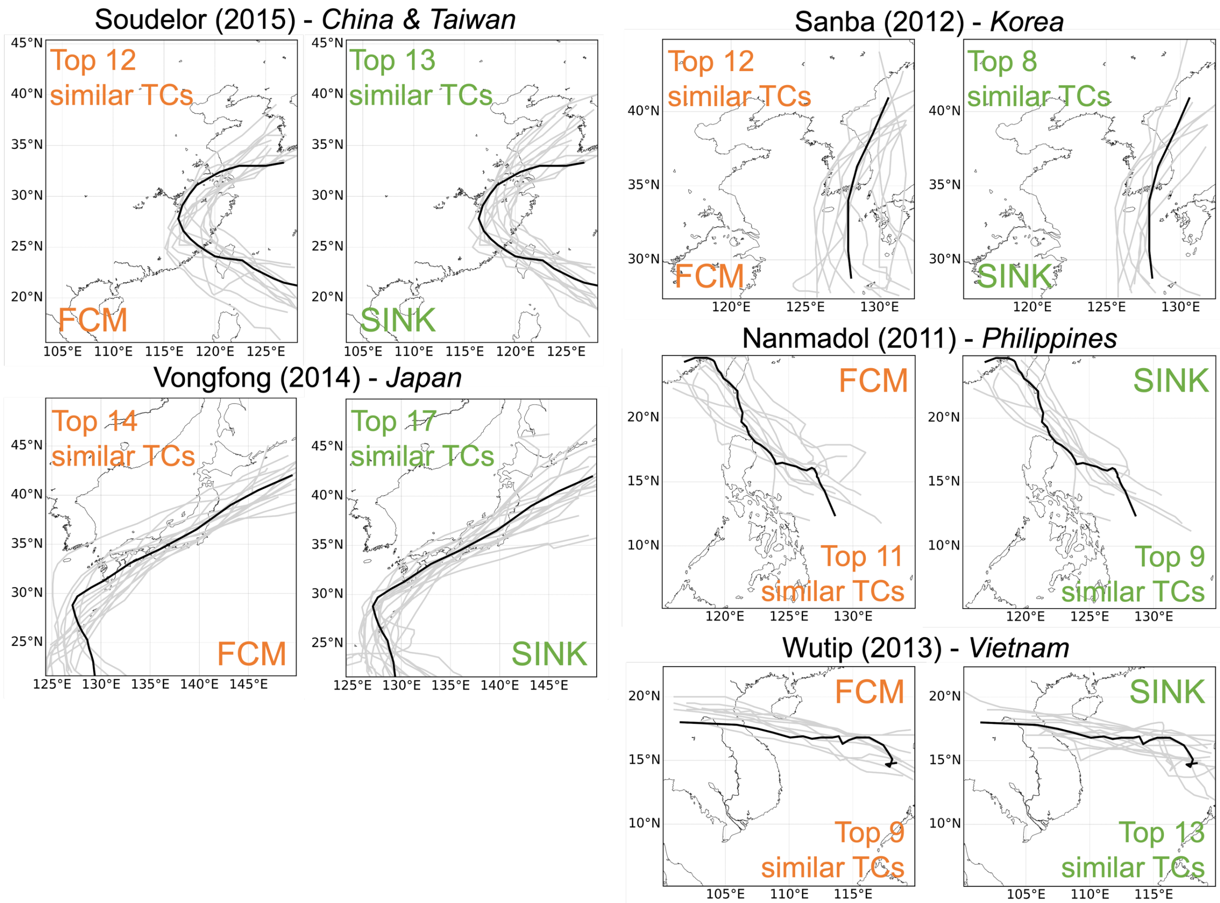

Figure 5. Similar tracks of optimal number identified through SINK and FCM for five typhoon cases in five simulation areas.

Overall, the identification of similar TCs through Sinkhorn Distance resulted in comparable predictions of TC rainfall (see Figure 6). Visually, the rainfall prediction using SINK appears to be superior when compared to predictions made using Fuzzy C-Means (FCM). In four out of the five typhoon cases – Soudelor (2015), Vongfong (2014), Sanba (2012), and Nanmadol (2011) – the results from SINK show higher rainfall values than those from FCM in areas where significant rainfall was observed (highlighted in red in Figure 6). In contrast, for the remaining typhoon case (Wutip), SINK predicts lower rainfall values than FCM in areas with high observed rainfall (also in red in Figure 6).

These values from SINK generally appear to be closer to the observed values than those from FCM.

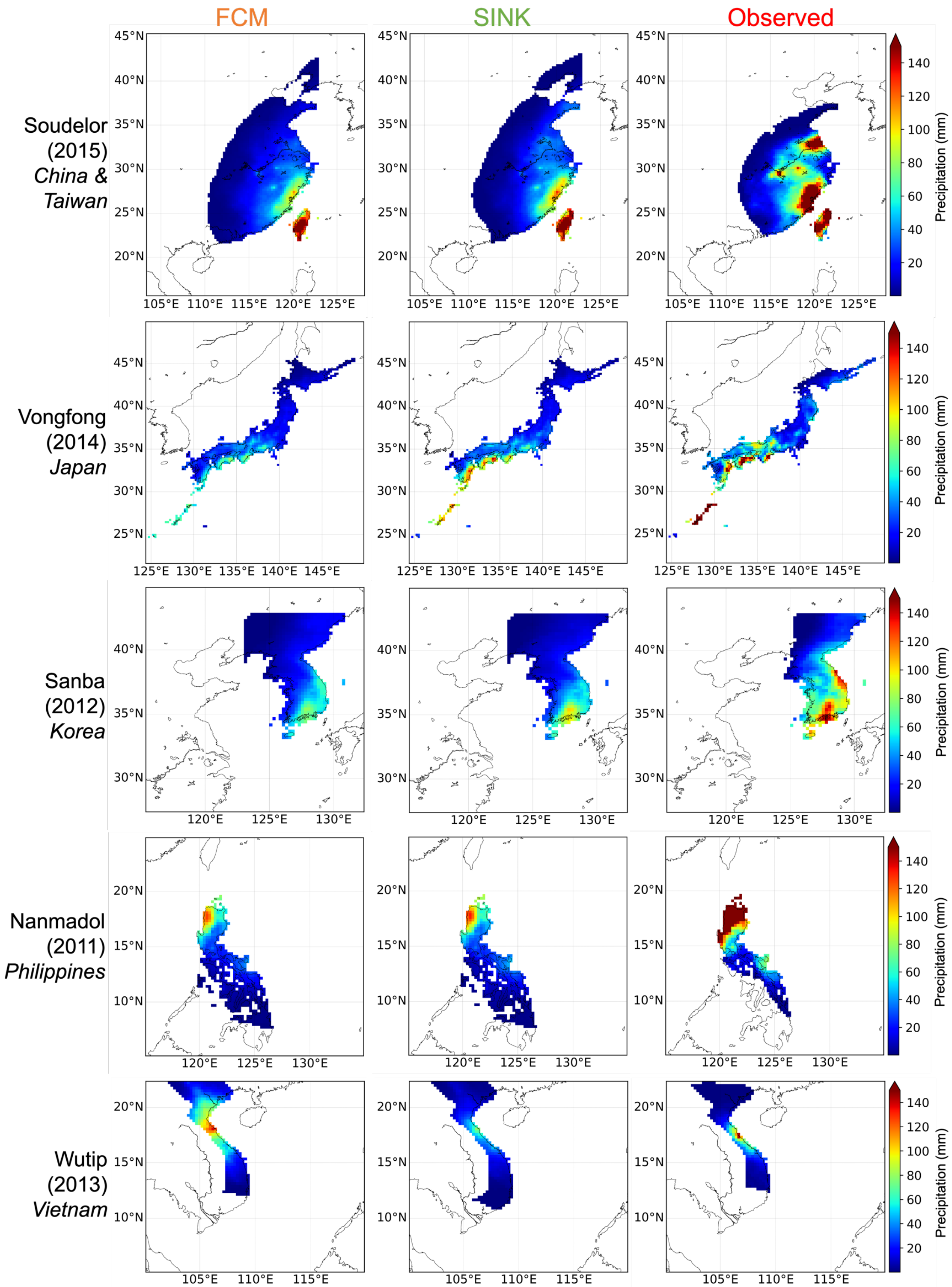

Figure 6. Spatially distributed rainfall prediction values based on similar TCs identified through SINK and FCM (Figure 5) for five typhoon cases in five simulation areas.

To quantitatively assess the prediction performance illustrated in Figure 6, we computed the Root Mean Square Error (RMSE) and correlation coefficient values. In general, the results from SINK exhibit superior RMSE and correlation values. The RMSE values are consistently better for all five typhoon cases in the results from SINK compared to those from FCM. The correlation coefficient of the results from SINK outperforms that of FCM for Soudelor (2015), Vongfong (2014), and Wutip (2013), remains similar for Nanmadol (2011), but is slightly worse for Sanba (2012). Notably, Sanba (2012) serves as an illustrative example of the degradation introduced by using SINK specifically in the context of Korea.

Table 2. RMSE and correlation coefficient values for the five simulation cases in five simulation areas. Values in bold red represent the best values. Only values inside rainfall calculation areas (shown in the Appendix) area are considered in the prediction error.

Simulation area	TC (Year)	No. of similar TCs used		RMSE [mm]		Correlation coefficient	
		FCM	SINK	FCM	SINK	FCM	SINK
China & Taiwan	Soudelor (2015)	12	13	23.48	**22.01**	0.82	**0.84**
Japan	Vongfong (2014)	14	17	17.87	**14.28**	0.85	**0.87**
Korea	Sanba (2012)	12	8	18.51	**17.13**	**0.97**	0.95
Philippines	Nanmadol (2011)	11	9	45.58	**44.85**	**0.96**	**0.96**
Vietnam	Wutip (2013)	9	13	24.95	**9.17**	0.64	**0.88**

5 Summary

Our research introduces the Sinkhorn Distance as a novel TC similarity measure in the statistical prediction of TC rainfall. By incorporating this metric into an established methodology (Kim et al., 2019; Hokson & Kanae, in press-b), we have demonstrated its potential effectiveness. Our investigation revealed that, in general, the utilization of Sinkhorn Distance leads to accurate predictions of TC rainfall across five simulation areas – namely China & Taiwan, Japan, Korea, Philippines, and Vietnam.

In comparison to the conventional approach employing FCM as a TC similarity measure, our methodology employing Sinkhorn Distance yielded generally better results for the simulation areas of China & Taiwan, Japan, Philippines, and Vietnam. However, it exhibited slightly less favorable outcomes for the simulation area of Korea. These assessments were based on spatially distributed rainfall data, with performance metrics quantified using RMSE and correlation coefficients.

Taken collectively, the inclusion of Sinkhorn Distance in our study presents an additional valuable tool for discerning similar TC tracks, thereby enhancing the accuracy of TC rainfall predictions. As statistical and AI techniques continue to advance, we anticipate even more refined approaches to further enhance our predictive capabilities. This study constitutes a stride towards a much-needed enhancement in our predictions of TC rainfall.

Appendix

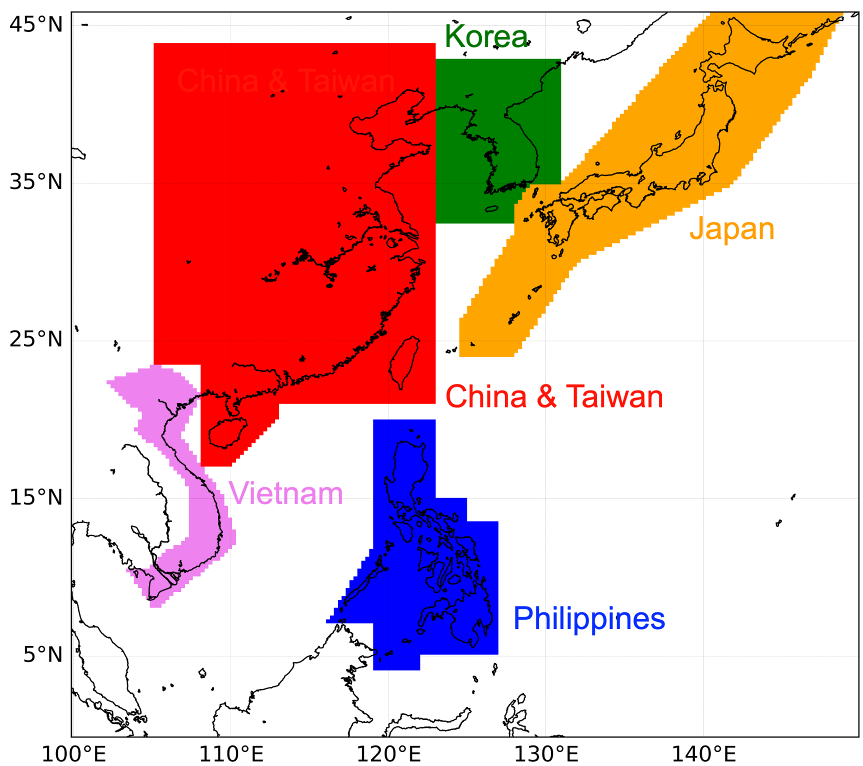

Figure A1. Rainfall calculation areas. Only values on land inside the rainfall calculation areas are considered in the analysis in the main text.

Acknowledgments

This work was supported by JSPS KAKENHI 21K18744.

Data Availability

All datasets used in this study are publicly available: RSMC Best Track Dataset is available at https://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/trackarchives.html; APHRODITE precipitation dataset is available at http://aphrodite.st.hirosaki-u.ac.jp/.

Code Availability

All codes are currently in preparation to be uploaded at a GitHub repository. These codes mainly contain code for data handling and implementation of the this paper's methodology, including implementation of Sinkhorn Distance Algorithm and FCM Algorithm.

Conflict of interest

The authors declare no conflict of interest.

References

Bagtasa, G. (2021). Analog Forecasting of Tropical Cyclone Rainfall in the Philippines. *Weather and Climate Extremes, 32*, 100323. https://doi.org/10.1016/j.wace.2021.100323

Bagtasa, G. (2022). Assessment of Tropical Cyclone Rainfall from GSMaP and GPM Products and Their Application to Analog Forecasting in the Philippines. *Atmosphere, 13*(9), 1398. https://doi.org/10.3390/atmos13091398

Cuturi, M. (2013). Sinkhorn Distances: Lightspeed Computation of Optimal Transport. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 26 (NIPS 2013) (pp. 2292-2300).

Gori, A., et al. (2022). Tropical cyclone climatology change greatly exacerbates US extreme rainfall–surge hazard. *Nature Climate Change*. https://doi.org/10.1038/s41558-021-01272-7

Guzman, O., & Jiang, H. (2021). Global increase in tropical cyclone rain rate. *Sustainability, 13*(18), 10222. https://doi.org/10.3390/su131810222

Hokson, J. A., & Kanae, S. (in press-a). Spatial and attribute filtering as a complementary measure in the statistical prediction of tropical cyclone rainfall. *Atmospheric Science Letters*.

Hokson, J. A., & Kanae, S. (in press-b) The use of along-track central pressure and movement speed in similar typhoon identification for rainfall prediction. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*.

Intergovernmental Panel on Climate Change. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.)]. Cambridge University Press.

Kim, H. J., Moon, I. J., & Kim, M. (2019). Statistical Prediction of Typhoon-Induced Accumulated Rainfall over the Korean Peninsula Based on Storm and Rainfall Data. *Meteorological Applications, 27*(1). https://doi.org/10.1002/met.1797

Kim, H.-S., Kim, J.-H., Ho, C.-H., & Chu, P.-S. (2011). Pattern Classification of Typhoon Tracks Using the Fuzzy C-Means Clustering Method. *Journal of Climate, 24*(2), 488–508. https://doi.org/10.1175/2010JCLI3771.1

Kim, J.-S., Chen, A., Lee, J., Moon, I.-J., & Moon, Y.-I. (2020). Statistical Prediction of Typhoon-Induced Rainfall over China Using Historical Rainfall, Tracks, and Intensity of Typhoon in the Western North Pacific. *Remote Sensing, 12*(24), 4133. https://doi.org/10.3390/rs12244133

Li, Q., Lan, H., Chan, J. C. L., Cao, C., Li, C., & Wang, X. (2016). An Operational Statistical Scheme for Tropical Cyclone-Induced Rainfall Forecast. In Recent Developments in Tropical Cyclone Dynamics, Prediction, and Detection (pp. 217-232). IntechOpen. https://www.intechopen.com/chapters/51981

Lee, T.-C., Knutson, T. R., Nakaegawa, T., Ying, M., & Cha, E. J. (2020). Third Assessment on Impacts of Climate Change on Tropical Cyclones in the Typhoon Committee Region – Part I: Observed Changes, Detection and Attribution. *Tropical Cyclone Research and Review, 9*(1), 1–22. https://doi.org/10.6057/2019TCRR12.01

Luitel, B. P., et al. (2018). Statistical Prediction of Tropical Cyclone-Induced Rainfall over the Philippines Using Machine Learning Techniques. *Journal of Hydrology, 556*, 118–129. https://doi.org/10.1016/j.jhydrol.2017.11.044

Magee, A. D., et al. (2021). A New Approach for Location-Specific Seasonal Outlooks of Typhoon and Super Typhoon Frequency across the Western North Pacific Region. *Scientific Reports, 11*, 19439. https://doi.org/10.1038/s41598-021-98329-6

Ren, F., et al. (2018). An Objective Track Similarity Index and Its Preliminary Application to Predicting Precipitation of Landfalling Tropical Cyclones. *Weather and Forecasting, 33*(6), 1725–1742. https://doi.org/10.1175/WAF-D-18-0059.1

Tu, J., et al. (2021). Recent Global Decrease in the Inner-Core Rain Rate of Tropical Cyclones. *Nature Communications, 12*, 1–9. https://doi.org/10.1038/s41467-021-22304-y

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., & Kitoh, A. (2012). Aphrodite: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bulletin of the American Meteorological Society, 93*(9), 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1

Wang, C., et al. (2023). Statistical Prediction of Typhoon-Induced Total Accumulated Rainfall in the Western North Pacific Using Typhoon Track Similarity Indices. *Atmospheric Research, 288*, 106724. https://doi.org/10.1016/j.atmosres.2022.106724