Geophysical Research Letters

RESEARCH LETTER
10.1029/2023GL104171

Predicting Tropical Cyclone-Induced Sea Surface Temperature Responses Using Machine Learning

Hongxing Cui[1,2], Danling Tang[1], Wei Mei[3], Hongbin Liu[1,2], Yi Sui[1,4], and Xiaowei Gu[5]

[1]Guangdong Remote Sensing Center for Marine Ecology and Environment, Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China, [2]Department of Ocean Science, Hong Kong University of Science and Technology, Hong Kong, China, [3]Department of Earth, Marine and Environmental Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [4]Department of Oceanography, Dalhousie University, Halifax, NS, Canada, [5]School of Computing, University of Kent, Canterbury, UK

Key Points:

- A machine learning-based model is built to predict the spatiotemporal evolution of the tropical cyclone-induced sea surface temperature response
- The model well predicts the spatial structure and temporal evolution of the observed response and captures the observed cross-case variance
- Feature scores are computed to assess the relative importance of the predictors in determining the magnitude of the SST response

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Tang and W. Mei,
lingzistdl@126.com;
wmei@email.unc.edu

Citation:

Cui, H., Tang, D., Mei, W., Liu, H., Sui, Y., & Gu, X. (2023). Predicting tropical cyclone-induced sea surface temperature responses using machine learning. *Geophysical Research Letters*, *50*, e2023GL104171. https://doi.org/10.1029/2023GL104171

Received 18 APR 2023
Accepted 27 AUG 2023

Abstract

This study proposes to construct a model using random forest method, an efficient machine learning-based method, to predict the spatial structure and temporal evolution of the sea surface temperature (SST) cooling induced by northwest Pacific tropical cyclones (TCs), a process of the so-called wind pump. The predictors in use include 12 predictors related to TC characteristics and pre-storm ocean conditions. The model is shown to skillfully predict the spatiotemporal evolutions of the cold wake generated by TCs of different intensity groups, and capture the cross-case variance in the observed SST response. Another model is further built based on the same method to assess the relative importance of the 12 predictors in determining the magnitude of the maximum cooling. Computations of feature scores of those predictors show that TC intensity, translation speed and size, and pre-storm mixed layer depth and SST dominate, depending on the area where the cooling is considered.

Plain Language Summary

While many studies have been devoted to understanding the processes and mechanisms underlying the sea surface temperature (SST) cooling induced by tropical cyclones (TCs), few studies have attempted to predict the spatial and temporal evolution of the sea surface temperature (SST) cooling triggered by TCs. In this study, we proposed to achieve this goal by building a model using an efficient and robust machine learning-based method. The constructed model uses 12 predictors associated with TC characteristics (e.g., intensity, and translation speed) and pre-storm ocean states (e.g., mixed layer depth). The model performs well in producing the TC-induced spatial structure and temporal evolution of the cold wake and can capture most of the variance in the observed SST response. We quantified the relative importance of the 12 predictors, and found that TC intensity, translation speed and size, and pre-storm mixed layer depth and SST dominate in deciding the magnitude of the SST response. The results and proposed method have important implications for predicting the response of ocean primary production to the TC wind pump effects.

1. Introduction

When a tropical cyclone (TC) passes over the ocean, it extracts a considerable amount of heat out of the ocean for its development, cooling down the ocean surface (e.g., Cione & Uhlhorn, 2003; Emanuel et al., 2004; Price, 1981; Shay et al., 1992). Meanwhile, the strong winds associated with the TC inject a huge amount of momentum into the upper ocean, driving strong ocean currents (e.g., Fan et al., 2009; Jarosz et al., 2007; Sun et al., 2022). When the vertical shear in the ocean currents is large enough to overcome the vertical stratification set up by temperature and salinity differences, vigorous vertical mixing takes place and stirs the surface warm water with colder water in the subsurface ocean, further reducing sea surface temperature (SST) (e.g., Shay et al., 1992; Sun et al., 2021; Zhang et al., 2021; Zhu & Zhang, 2006). The surface ocean cooling generated by TCs has been well evidenced by satellite images of SST (e.g., Bender et al., 1993; Chen & Tang, 2012; Hart, 2011; Liu et al., 2020; Price, 1981; Yang & Tang, 2010). The produced surface cooling in turn can diminish the heat flow from the ocean to the TC, and thereby slow down the storm's intensification or cause it to decay (the so-called negative SST feedback) (e.g., Cione & Uhlhorn, 2003; Emanuel, 1999; Liu, Wei, et al., 2019; Mei et al., 2012; Mei, Lien, et al., 2015; Mei, Xie, et al., 2015; Mei & Pasquero, 2013; Palmen, 1948; Wang & Wu, 2004; Zhang et al., 2019; Zhu & Zhang, 2006). The cooling may also serve as a predictor for the accompanying changes in chlorophyll-a (Chl-a) concentrations in the surface ocean influenced by the TC wind pump effects (e.g., Chen & Tang, 2012; Cui et al., 2023; Liu et al., 2020; Zheng & Tang, 2007). Thus, substantial observational, modeling and theoretical efforts have been

devoted to improving our understanding of the processes involved in TC-ocean interaction and the contributions of those processes to TC-induced sea surface cooling (e.g., Balaguru et al., 2022; Huang et al., 2009; Liu, Wei, et al., 2019; Mei & Pasquero, 2012; Pasquero et al., 2021; Sriver & Huber, 2007; Sui et al., 2022; Vincent, Lengaigne, Vialard, et al., 2012; Vincent, Lengaigne, Madec, et al., 2012; Wang et al., 2016; Zhang et al., 2015).

The key factors that have been identified to affect the magnitude of the TC-induced sea surface cooling include both TC and ocean characteristics, such as TC intensity and translation speed, and pre-storm ocean mixed layer depth (MLD) and stratification below mixed layer (e.g., Balaguru et al., 2012; Huang et al., 2015; I. I. Lin, 2012; Lin et al., 2013; Mei et al., 2012; Mei & Pasquero, 2013; Zheng et al., 2008). The relative importance of these factors may differ significantly from location to location (e.g., Liu, Zhang, & Guan, 2019; Liu, Wei, et al., 2019; Vincent, Lengaigne, Madec, et al., 2012; Wang et al., 2016). For a specified basin or area, identifying the dominant factors and quantifying their roles in TC-generated SST cooling can help improve the prediction of SST changes caused by future TCs.

With the development of artificial intelligence technology in recent years, several machine learning-based models have been proposed to predict TC-induced SST cooling (e.g., Jiang et al., 2018; Wei et al., 2017; Zhao et al., 2022). For example, Wei et al. (2017) introduced a neural network model to predict the TC-generated sea surface cooling, and showed that a neural network is more efficient than linear regression. Jiang et al. (2018) developed a deep learning-based algorithm to improve the stability of the prediction of TC-induced surface cooling. While these studies have greatly advanced the application of machine learning in predicting TC-generated SST response, they focused primarily on the cooling generated by a limited number of TC cases. For instance, Jiang et al. (2018) focused on four cases (including Soulik, Neoguri, Halong, and Muifa) and considered only five predictors, that is, 10-m winds ($U10$ and $V10$), ocean temperature at 100-m depth, sea surface height (SSH), and SST. Recently, a long short-term memory (LSTM) neural network has been proposed to predict the SST after TC passage using long-term time series data. For instance, Shao et al. (2021) introduced a hybrid model that combines a LSTM model with a statistical method to forecast SST and SSH anomalies after TC passage in the South China Sea (SCS). Zhang et al. (2023) developed a multi-variate LSTM model using the same predictors as those used in Jiang et al. (2018) to predict TC-induced SST cooling in the northwest Pacific Ocean (NWP).

The above-mentioned studies open a new avenue for understanding and studying TC-induced SST cooling, and provide insights into the application of machine learning in predicting the cooling. However, those studies did not consider TC characteristics (e.g., intensity, translation speed, and size) that have been shown to strongly affect the TC-induced SST cooling. Furthermore, no studies have been devoted to predicting the temporal evolutions of the spatial structure of the TC-induced SST response and assessing the relative importance of the various factors that have been identified to modulate the cooling.

In the present study, we aim to further research in this field and develop a machine learning-based model to predict the spatial structure and temporal evolution of the SST anomalies generated by TCs in the NWP, the basin where TCs are most active. We first construct the model using a random forest regressor and 12 variables related to TC characteristics and pre-storm ocean states. We show that the model can skillfully predict the spatial structure and temporal evolution of the TC-induced SST cooling for TCs of different intensity groups (e.g., the dependence of the cooling on TC intensity, strong asymmetry of the cooling with respect to TC track, and leftward movement of the local maximum cooling with time after TC passage). We then build another model using the same machine learning method to quantify the relative importance of different predictors in determining the magnitude of the TC-induced maximum cooling. After describing the data and methods in Section 2, we present the results from the models built on machine learning in Section 3. Section 4 contains discussions and conclusions. For clarity, a list of key abbreviations and their meanings is given in Table S4 in Supporting Information S1.

2. Data and Methods

2.1. Data

We focus our work here on the NWP, which is defined as the region between equator and 30°N, 100°, and 160°E (Figure 1a). The following datasets are used in this study. The best-track data for TCs were produced by the Japan Meteorological Agency (JMA) and obtained from the International Best Track Archive for Climate Stewardship (IBTrACS) archive (Knapp et al., 2010), which provides TC intensity, location, translation direction and speed, and the shortest radius of 30-kt winds or greater at 6-hr intervals. SST data are from the Tropical

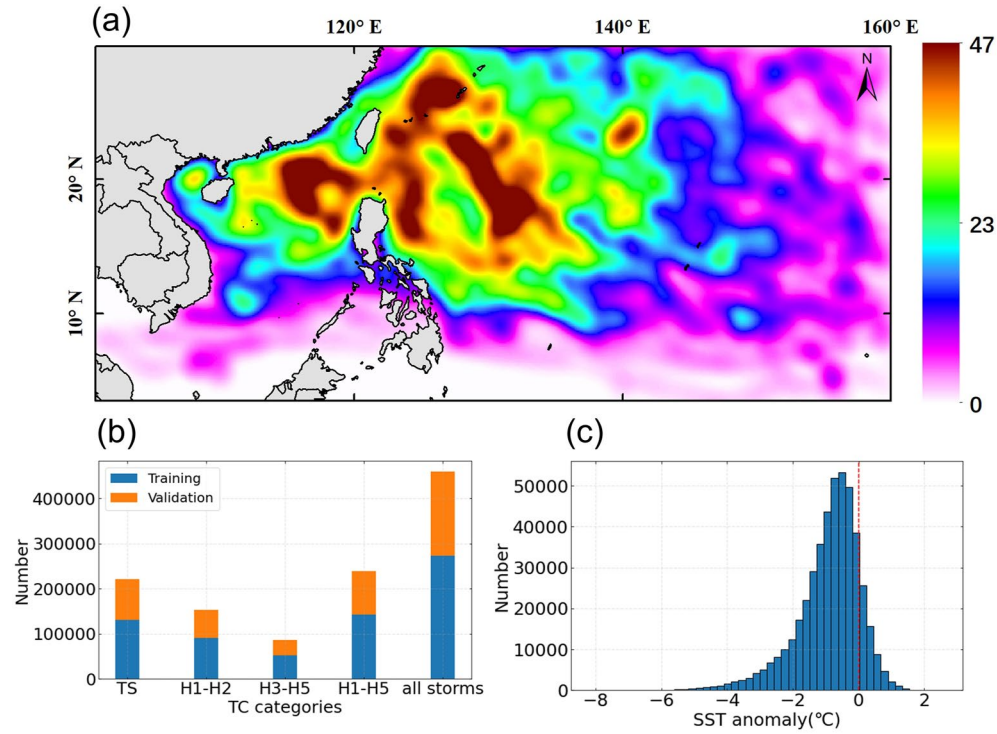

Figure 1. Study area and distribution of data points in the training and validation sets during January 1998–December 2018. (a) TC occurrence density, calculated as the number of TC occurrence during the entire study period within individual 0.1° × 0.1° grids. (b) Histogram of the number of data points used in the training stage as a function of TC intensity. (c) Histogram of TC-induced SST anomalies on day 1 after TC passage in the training stage. Abbreviations: TS, tropical storm; H1–H2, category-1–2 hurricane; H3–H5, category-3–5 hurricane; and H1–H5, category-1–5 hurricane.

Rainfall Measuring Mission (TRMM) Microwave Imager (Wentz et al., 2000), and have a spatial resolution of 0.25° × 0.25° and a daily temporal resolution including ascending and descending orbit segments; to test the robustness of the results, we also use another satellite-derived SST data, microwave and infrared (MW_IR) which have a spatial resolution of 0.1° × 0.1° and daily temporal resolution (Reynolds et al., 2007). Daily altimeter satellite gridded sea level anomalies (SLA) are from the EU Copernicus Marine Environment Monitoring Service (CMEMS). They have a spatial resolution of 0.25° × 0.25°, and are estimated by Optimal Interpolation, merging the measurement from multiple altimeter missions available (e.g., Jason-1, Jason-2, Jason-3, Sentinel-3A, and HY-2). Monthly data of MLD and potential temperature are from the World Ocean Atlas (WOA) and have a spatial resolution of 0.25° × 0.25° (Boyer et al., 2019). Eastward current speed (U) and northward current speed (V) of the ocean are provided by the Daily Global Reanalysis Multi-Ensemble Product GREP reanalysis, which has 75 levels vertically (from 0 to 5,500 m) and a horizontal resolution of 0.25° × 0.25°. Because TRMM SST data are available starting from 1998 and U and V data are available only before 2018, we limit our work here over the period of 1998–2018.

2.2. Preprocessing

As described in Introduction, the magnitude of TC-induced SST cooling is primarily determined by TC characteristics and pre-storm upper-ocean conditions (e.g., Lin et al., 2003a, 2003b; Mei et al., 2012; Mei & Pasquero, 2013; H. Zhang, 2023). Here we consider 12 predictors as input features for the model described in Section 2.3. Six of them are related to TC characteristics, including intensity (V_{max}), translation speed (V_{trans}), translation direction (V_{dir}), the shortest radius of 30-kt or greater ($R30$), TC-centered longitude (Lon), and TC-centered latitude (Lat). To apply the Saffir-Simpson hurricane wind scale, we convert the JMA TC wind speeds from 10-min mean values to 1-min mean values following Mei and Xie (2016) and using the Koba table (Hagiwara et al., 1990), and use the obtained 1-min mean values to represent TC intensity. The other six predictors are related to pre-storm upper-ocean state: mixed layer depth ($\overline{MLD}$), SSH ($\overline{SSH}$), SST ($\overline{SST}$), ocean temperature at 75-m depth ($\overline{T75}$),

vertical shear of zonal current speed ($\overline{UgRes}$), and vertical shear of meridional current speed ($\overline{VgRes}$). They are obtained by removing the climatologically seasonal cycle and averaging in time during seven to four days prior to the TC passage.

2.3. Model Construction Using Random Forest Method

We employ the random forest method to build a model for the prediction of TC-generated SST response from 3 days before TC passage to 14 days after the passage (referred to as day −3, day −2, …, and day 14, respectively), using the above-mentioned 12 predictors (i.e., $\overline{MLD}$, $\overline{SSH}$, $\overline{SST}$, $\overline{T75}$, $\overline{UgRes}$, $\overline{VgRes}$, V_{max}, V_{trans}, V_{dir}, $R30$, Lon, Lat). Random forest is currently one of the most widely used machine learning methods for prediction, and has been applied to oceanic and atmospheric research in recent years (e.g., Jergensen et al., 2020; Kumar et al., 2021; Rolnick et al., 2022) because of its good performance and high computational efficiency. The random forest method was developed within a bagging ensemble framework that combines a group of weak decision tree-based learners to construct a stronger prediction model (Breiman, 2001). During the training stage, each regression tree is fed with a subset of training data created by randomly sampling with replacement to construct a predictive model. Then, during the testing stage, each regression tree produces a prediction for each input sample. The final output of random forest is the average of the outputs of individual regression trees (Figure S1 in Supporting Information S1). To avoid overfitting, we use a grid search method of 5-fold cross-validation to determine the major parameters for the model (see Text S1 and Figure S2 in Supporting Information S1).

We divided all TC cases (627,400 in total) into two groups, with 2/3 used in the training stage and 1/3 used in the testing stage. We further divided the data used to train the model (Figure 1c) into two subgroups: 60% for training and 40% for validation (Figure 1b). After training and validation, the constructed model is used to predict the spatial structure and temporal evolution of the cold wakes induced by the remaining 1/3 TC cases (i.e., TC cases in the testing group). The predicted cold wakes are then compared with those in the observations to evaluate the performance of the constructed model. To assess the model performance for TCs of different intensities, the testing data are further stratified by TC intensity that is measured using the Saffir-Simpson wind scale [that is, TS: tropical storm (17.5 m/s $\leq V_{max} <$ 33 m/s), H1: category-1 hurricane (33 m/s $\leq V_{max} <$ 43 m/s), …, H5: category-5 hurricane ($V_{max} \geq$ 70 m/s), H1–H2: category-1–2 hurricane (33 m/s $\leq V_{max} <$ 50 m/s), H3–H5: category-3–5 hurricane ($V_{max} \geq$ 50 m/s), H1–H5: category-1–5 hurricane ($V_{max} \geq$ 33 m/s)].

2.4. Accuracy Assessment and Feature Importance

We use the Pearson correlation coefficient (r, Equation 1), root mean square error (RMSE, Equation 2), and mean absolute error (MAE, Equation 3) to measure the performance of the constructed model:

$$r = \frac{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)\left(P_i - \overline{P}\right)}{\sqrt{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2 \sum_{i=1}^{N}\left(P_i - \overline{P}\right)^2}}, \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2}, \tag{2}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|O_i - P_i|, \tag{3}$$

where O_i is the observed value, P_i is the predicted value, $\overline{O}$ is the mean of O_i, $\overline{P}$ is the mean of P_i, N is the total number of cases, and i = 1, 2,…,N. Smaller values of MAE and RMSE and larger values of r indicate better model performance.

To determine the relative importance of the 12 predictors in modulating the TC-induced SST cooling, we examined the feature scores of these predictors (Tuv et al., 2009). The relative importance of those predictors may depend on the area over which the spatial averaging is computed. Accordingly, we evaluate the feature scores of the predictors for SST cooling averaged over TC-centered boxes of three different sizes (i.e., 1° × 1°, 3° × 3°, and 5° × 5°, respectively) and examine the possible dependency of the dominant predictors.

3. Results

3.1. Spatial Structure and Temporal Evolution of TC-Induced Cold Wake

Figure 2a displays the spatial structure and temporal evolution of the composite cold wake induced by TCs of all intensities in the observations. On average, the cooling begins around 2 days prior to the TC passage, intensifies during the day of the TC passage, and reaches its peak of above 1.3°C on the day right after the TC passage. The cooling also expands in space from day −2 to day 1. In line with previous studies (e.g., Mei & Pasquero, 2013), the cooling exhibits strong asymmetry with respect to the TC track owing to the resonance effect (e.g., Jansen et al., 2010; Liu, Wei, et al., 2019; Price, 1981): On day 1, the local maximum cooling is located around 50 km to the right of the TC track.

Figure 2f shows the cross-track profiles of the SST cooling generated by TCs of various intensities on the day right after the TC passage (i.e., on average the day when the cooling peaks). The asymmetry and rightward bias exist in the SST cooling generated by TCs of different intensities. Figure 2f also illustrates the strong dependence of the cooling on TC intensity. For example, the average local maximum cooling produced by a tropical storm is around 1°C, whereas that left behind a category-3–5 hurricane is about 2°C.

After day 1, the cold wake decays rapidly in terms of both the magnitude and spatial coverage of the cooling (Figure 2a). On day 14, the local maximum cooling has weakened to less than 0.4°C. Meanwhile, the location of the local maximum cooling move leftward owing to the westward propagation of Rossby waves (e.g., Jansen et al., 2010), and as a result the cold wake becomes more symmetric.

Figure 3a shows the temporal evolution of the composite SST anomalies averaged within a TC-centered 3° × 1° box (3° in the cross-track direction and 1° in the along-track direction) induced by TCs of different intensities. Consistent with Figure 2a, the SST drops quickly during day 0 and day 1, experiences a quick rebound during days 2–4, followed by a recovery at a slower rate. These evolutions of the cooling are independent of TC intensity, and the decay *e*-folding times are about 10 days.

Figure 2b shows the composite map of the SST cooling generated by TCs of all intensities in the predictions by the constructed model, with its difference from the observations shown in Figure S3 in Supporting Information S1. Overall, the model is skillful at predicting the spatial structure and temporal evolution of the cooling in the observations, for example, the quick strengthening and spatial expansion of the cooling during day 0 and day 1 and the strong asymmetry in the cooling. We further calculated r, RMSE and MAE to quantify the performance of the model in predicting the TC-induced SSTA. The model captures more than 90% of the cross-case variance (i.e., $r \geq 0.95$) in the cooling within 200 km of the TC track during day −3 and day 14 (Figure 2c). In general, the stronger the cooling, the larger the error as estimated using RMSE and MAE (Figures 2d and 2e). For instance, the MAE is above 0.2°C at the location of the local maximum cooling on day 1. The predictions for the cooling induced by TCs of hurricane intensities gives similar conclusions (Figure S4 in Supporting Information S1).

Figure 2g shows the cross-track profiles of the predicted TC-induced SST cooling for TCs of various intensities on the day right after the TC passage; the comparisons between the prediction and the observation for individual TC intensity groups are shown in Figure S5 in Supporting Information S1. The model well predicts the prominent features in the observations (Figure 2h), including the asymmetry and rightward bias of the SST cooling for all TC intensity groups and strong dependence of the magnitude of the cooling on TC intensity. In addition, we note that on average the model underpredicts the magnitude of the local maximum cooling for all intensity groups, with larger RMSEs and MAEs for stronger TCs (Figures 2i and 2j). Figures S6 and S7 in Supporting Information S1 compare the spatial pattern of the composite TC-induced SST cooling on day 1, and show that the model well captures both the cross- and along-track structure of the TC-induced SST cooling. Figure S8 in Supporting Information S1 provides a detailed comparison of the cooling averaged within a TC-centered 3° × 1° box on day 1 induced by all TCs in the predictions with that in the observations. Although overall the predicted cooling is weaker than that in the observations, the model well captures the case-to-case variations in the maximum cooling, with a correlation coefficient of 0.97.

Figure 3b displays the predicted temporal evolutions of TC-induced SSTA averaged within a TC-centered 3° × 1° box for TCs of various intensities; the comparisons between the prediction and the observation for individual TC intensity groups are shown in Figure S9 in Supporting Information S1. The predicted temporal evolutions highly resemble those in the observations for all intensity groups, with the cross-case correlation coefficients above 0.95

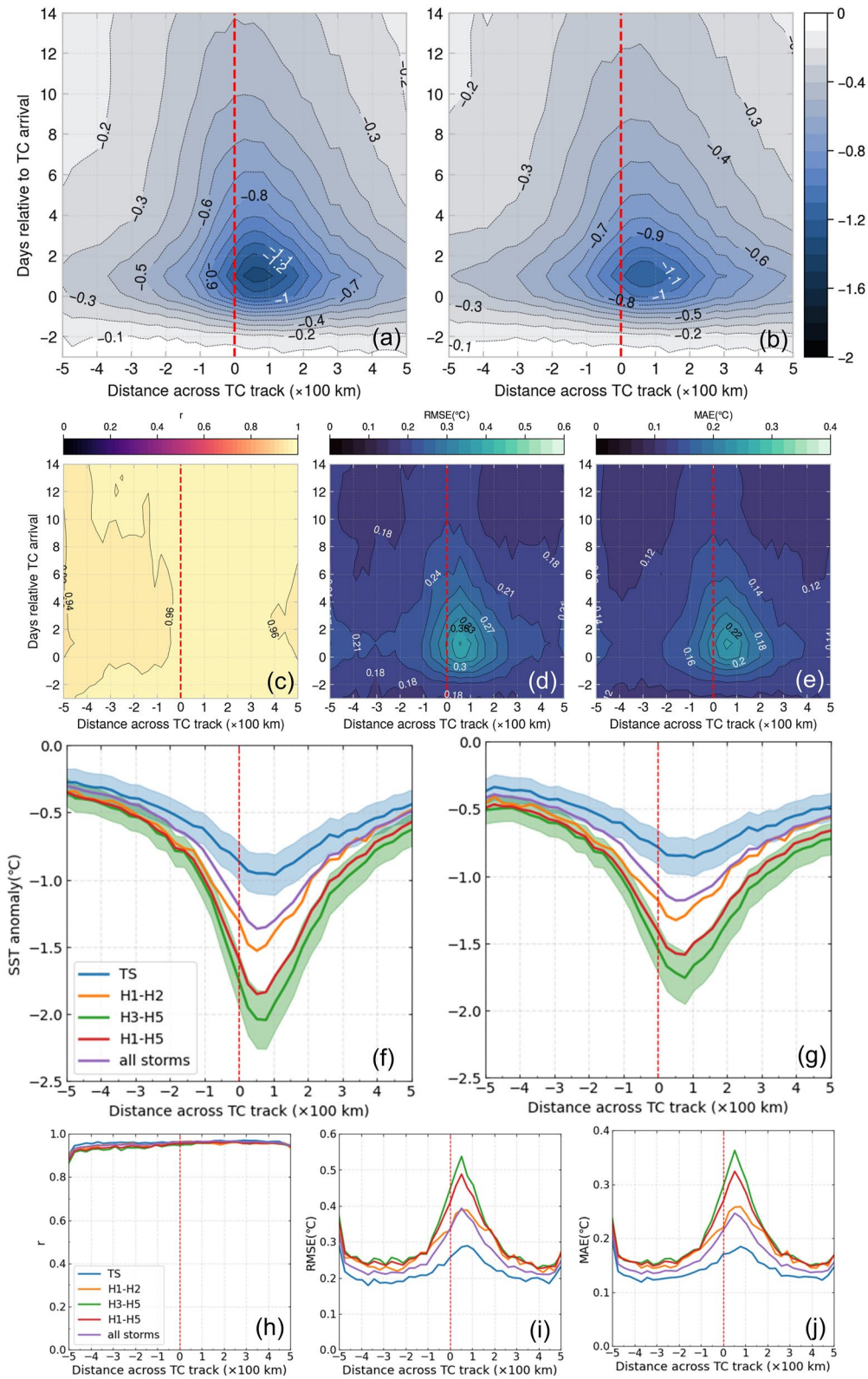

Figure 2.

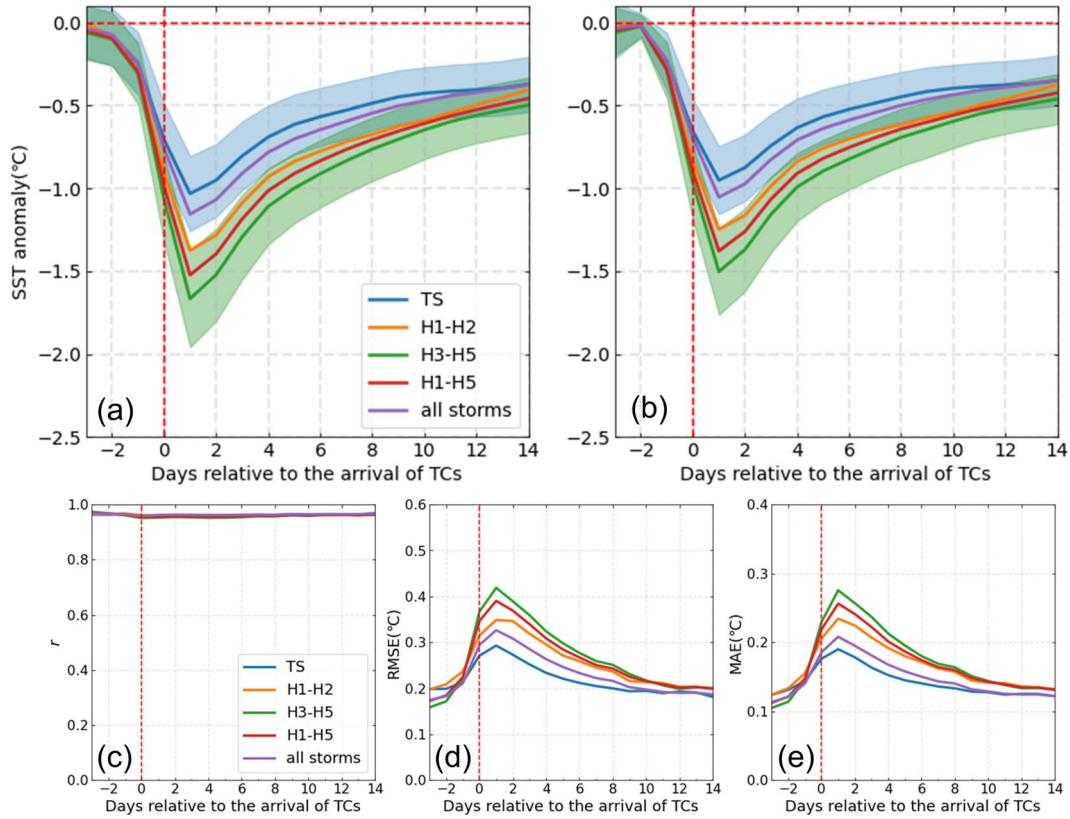

Figure 3. (a) Temporal evolution of the composite SST anomaly averaged within a TC-centered 3° × 1° box in association with the passage of TCs of different intensity groups in the observations from the testing set. The standard errors for TS and H3–H5 are calculated and shown as in Figure 2f. (b) As in (a), but for the predictions. (c)–(e) Evaluation of the model performance for (b), that is, (c) correlation coefficient *r*, (d) RMSE (°C) and (e) MAE (°C), for different TC intensity groups.

for all days (Figure 3c). Both the RMSE and MAE rapidly increase from day −3 to day 1, peak on day 1, and gradually decrease after that (Figures 3d and 3e).

We also tested all the results shown above using a different SST data set (i.e., MW_IR SST), and found no significant differences. As an example, Figure S10 in Supporting Information S1 compares the corresponding spatial structure and temporal evolution of the SST cooling induced by all TCs in the observations and predictions. We conclude that the model is skillful at simulating and predicting the TC-produced cold wake and its evolution, independent of the SST observations in use.

3.2. Quantification of Relative Importance of Predictors

To assess the relative importance of the 12 factors considered here in producing the SST cooling, we built another model to reconstruct the strongest cooling induced by all TCs and then calculate the feature scores of the 12 factors using the feature permutation; note that the strongest cooling could occur during day 0 to day 4 after TC passage although on average the strong cooling takes place on day 1. Because the cooling is inhomogeneous in space as shown in Figure 2a, averaging the cooling over area of different sizes may produce different results. Accordingly, we examine the cooling averaged over TC-centered boxes of three different sizes: 1° × 1°, 3° × 3°, and 5° × 5°, in order to test the sensitivity of the results to the size of the box in use.

Figure 2. (a) Spatial structure and temporal evolution of the TC-induced SST anomalies (°C) in the observations from the testing set. (b) As in (a), but for the predictions. (c)–(e) Evaluation of the model performance for (b), that is, (c) correlation coefficient *r*, (d) RMSE (°C), and (e) MAE (°C). (f) TC-induced SST anomalies in the observations as a function of the distance across TC track on the day right after the cyclone passage when the SST anomaly near the storm center peaks. Only the standard errors for TS and H3–H5 are shown (light blue and light green shading, respectively), and are calculated as the standard deviation divided by the square root of the number of observations used in the testing stage (i.e., standard error of the mean); the size of the standard errors for the other three intensity groups is in between these two. (g) As in (f), but for the predictions. (h)–(j) Evaluation of the model performance for (g), that is, (h) correlation coefficient *r*, (i) RMSE (°C) and (j) MAE (°C), for different TC intensity groups.

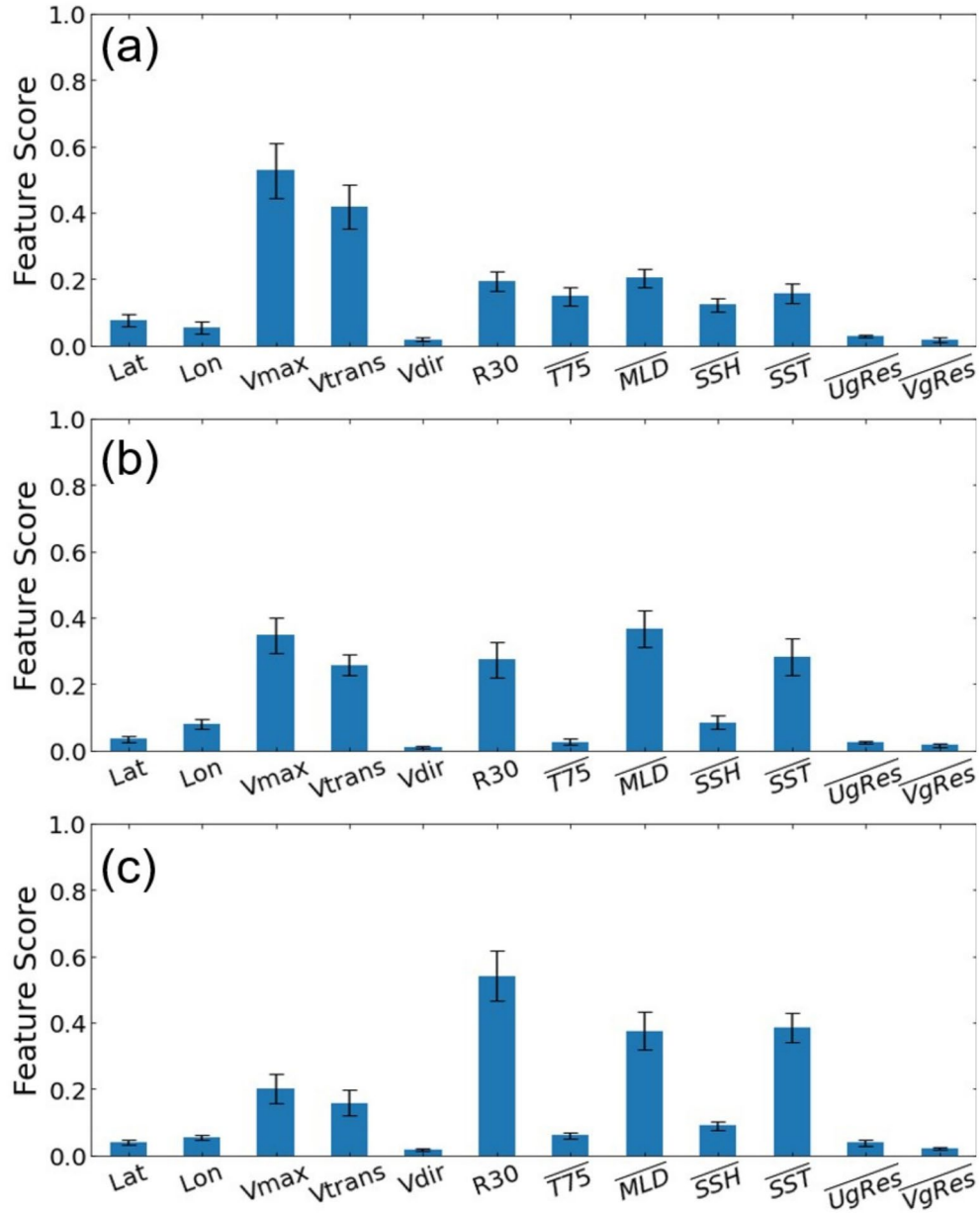

Figure 4. (a) The feature scores of the 12 predictors used in the model constructed to predict the maximum SST cooling induced by TCs averaged within a TC-centered 1° × 1° box. (b), (c) As in (a), but for the model constructed to predict the maximum SST cooling induced by TCs averaged within TC-centered 3° × 3° and 5° × 5° boxes, respectively. Vertical lines show the standard deviations of the feature scores.

Figure 4a shows the feature scores when a TC-centered 1° × 1° box is used: TC intensity (V_{max}) and translation speed (V_{trans}) are the two dominant factors, followed by $R30$, $\overline{MLD}$, and $\overline{SST}$. All these factors have been shown in the literature to play important roles in determining the magnitude of the TC-induced SST cooling (e.g., Mei et al., 2012; Mei & Pasquero, 2013). As the size of the box increases, the importance of TC intensity (V_{max}) and translation speed (V_{trans}) decreases, but TC size ($R30$) and pre-storm ocean states (i.e., $\overline{MLD}$ and $\overline{SST}$) become more important (Figures 4b and 4c). Specifically, for cooling averaged over a TC-centered 3° × 3° box, TC features and pre-storm ocean conditions are nearly equally important, while for cooling averaged over a TC-centered 5° × 5° box, TC size and pre-storm ocean conditions dominate over TC intensity and translation speed. These results are expected since the influence of the former two factors is relatively more localized around the TC track and the latter three factors affect SSTs over a larger area. The feature scores of individual factors

show some dependence on the size of the TC-centered box used in spatial averaging. It is worth noting that V_{max} always contributes more than V_{trans}, independent of the box size.

4. Discussions and Conclusions

In this study, we applied the random forest method to build models to predict the sea surface temperature (SST) responses to tropical cyclones (TCs) in the northwest Pacific using 12 factors related to TC characteristics (V_{max}, V_{trans}, V_{dir}, $R30$, Lon, Lat) and pre-storm ocean properties ($\overline{MLD}$, $\overline{SSH}$, $\overline{SST}$, $\overline{T75}$, $\overline{UgRes}$, $\overline{VgRes}$). TC characteristics have not been included as predictors in previous studies (e.g., Shao et al., 2021; Zhang et al., 2023). In this study, we have demonstrated that the performance of the model could be effectively improved by using both the TC features and pre-storm upper-ocean variables as predictors. Specifically, the constructed model well predicts the spatial structure and temporal evolution of the TC-induced SST cooling for various TC intensity groups, for example, the cross-track profile and strong asymmetry of the cooling on day 1 (the day on average when the maximum cooling occurs), the leftward movement of the local maximum cooling with time after day 1, and the temporal evolution of the SSTA averaged within a TC-centered 3° × 1° box. The model also well captures the cross-case variance in TC-generated sea surface cooling in the observations. We note, however, that more efforts are needed to improve the prediction of the local maximum of the TC-generated SST cooling (e.g., Figures 2a and 2b).

We then proceeded to assess the relative importance of the 12 predictors considered in the model in determining the magnitude of the SST cooling, which has not been quantified in previous studies. The assessment was achieved by building another model to predict the TC-induced maximum cooling averaged within TC-centered boxes and then calculating the feature scores of those predictors using the feature permutation. For SST cooling averaged within a small TC-centered box (e.g., 1° × 1°), TC intensity and translation speed are the primary contributors, followed by TC size and pre-storm mixed layer depth and SST. These results are broadly consistent with previous studies (e.g., Mei & Pasquero, 2013) in that stronger, slower-moving and larger TCs in areas with shallower mixed layer generate stronger SST cooling, although the relative importance of the factors has not been evaluated previously. As the box size increases, the importance of TC intensity and translation speed weakens, while TC size, pre-storm mixed layer depth and SST become more important. For example, for cooling averaged within a 5° × 5° box the latter three predictors dominate over the former two in shaping the SST cooling. This is consistent with our expectation, because the influence of the former two predictors is relatively more localized around the TC track and the latter three predictors exert influence over a larger area.

In short, this study has demonstrated the effectiveness of a random forest regressor in predicting the SST cooling generated by TCs. We recognize that further improvements (e.g., the underestimated SST cooling in the model predictions) are still needed for our model. In addition, we plan to apply the same methodology to other ocean basins (e.g., the North Atlantic), and compare the results with those shown in this study, such as the model performance and dominant factors. The presented results and proposed method also have important implications for predicting the response of ocean primary production to the TC wind pump effects.

Data Availability Statement

Microwave Imager SST and MW_IR SST data were provided by Remote Sensing Systems for the (https://www.remss.com/missions/tmi/). Daily altimeter satellite gridded SLA data and the daily current speed (U and V) data were provided by Copernicus Marine and Environmental Monitoring Service Center (https://resources.marine.copernicus.eu/). TC best track data was provided by National Centers for Environmental Information (https://www.ncdc.noaa.gov/ibtracs/). MLD and potential temperature data were provided by WOA (https://www.ncei.noaa.gov/data/oceans/woa/WOA18/).

Acknowledgments
This study was financially supported by Key Special Supporting Talent Team Project of the Natural Science Foundation of Guangdong (NSFGDP) (2019BT02H594), Key Special Project of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (for Introduced Talents Team, GML2021GD0810), Major Project of National Social Science Foundation of China (21ZDA097), and Remote Sensing Monitoring Study of Atmospheric Methane in the Northern South China Sea (2022ZD003).

References

Balaguru, K., Chang, P., Saravanan, R., Leung, L. R., Xu, Z., Li, M., & Hsieh, J.-S. (2012). Ocean barrier layers' effect on tropical cyclone intensification. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(36), 14343–14347. https://doi.org/10.1073/pnas.1201364109

Balaguru, K., Foltz, G. R., Leung, L. R., & Hagos, S. M. (2022). Impact of rainfall on tropical cyclone-induced sea surface cooling. *Geophysical Research Letters*, *49*(10), e2022GL098187. https://doi.org/10.1029/2022gl098187

Bender, M. A., Ginis, I., & Kurihara, Y. (1993). Numerical simulations of tropical cyclone-ocean interaction with a high-resolution coupled model. *Journal of Geophysical Research*, *98*(D12), 23245–23263. https://doi.org/10.1029/93jd02370

Boyer, T., Baranova, O., Locarnini, R., Mishonov, A., Grodsky, A., Paver, C., et al. (2019). *WORLD OCEAN ATLAS 2018 product documentation ocean climate laboratory NCEI*. NESDIS / NOAA NOAA National Centers for Environmental Information.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Chen, Y., & Tang, D. (2012). Eddy-feature phytoplankton bloom induced by a tropical cyclone in the South China Sea. *International Journal of Remote Sensing*, *33*(23), 7444–7457. https://doi.org/10.1080/01431161.2012.685976

Cione, J. J., & Uhlhorn, E. W. (2003). Sea surface temperature variability in Hurricanes: Implications with respect to intensity change. *Monthly Weather Review*, *131*(8), 1783–1796. https://doi.org/10.1175//2562.1

Cui, H., Tang, D., Liu, H., Sui, Y., & Gu, X. (2023). Composite analysis-based machine learning for prediction of tropical cyclone-induced sea surface height anomaly. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 2644–2653. https://doi.org/10.1109/jstars.2023.3247881

Emanuel, K., DesAutels, C., Holloway, C., & Korty, R. (2004). Environmental control of tropical cyclone intensity. *Journal of the Atmospheric Sciences*, *61*(7), 843–858. https://doi.org/10.1175/1520-0469(2004)061<0843:ecotci>2.0.co;2

Emanuel, K. A. (1999). Thermodynamic control of hurricane intensity. *Nature*, *401*(6754), 665–669. https://doi.org/10.1038/44326

Fan, Y., Ginis, I., Hara, T., Wright, C. W., & Walsh, E. J. (2009). Numerical simulations and observations of surface wave fields under an extreme tropical cyclone. *Journal of Physical Oceanography*, *39*(9), 2097–2116. https://doi.org/10.1175/2009jpo4224.1

Hagiwara, T., Osano, S., & Akashi, S. (1990). Relationship between the CI-number and central pressure and maximum wind speed in typhoons. *Journal of Meteorological Research*, *42*, 59–67.

Hart, R. E. (2011). An inverse relationship between aggregate northern hemisphere tropical cyclone activity and subsequent winter climate. *Geophysical Research Letters*, *38*(1), L01705. https://doi.org/10.1029/2010gl045612

Huang, P., Lin, I.-I., Chou, C., & Huang, R.-H. (2015). Change in ocean subsurface environment to suppress tropical cyclone intensification under global warming. *Nature Communications*, *6*(1), 1–9. https://doi.org/10.1038/ncomms8188

Huang, P., Sanford, T. B., & Imberger, J. (2009). Heat and turbulent kinetic energy budgets for surface layer cooling induced by the passage of Hurricane Frances (2004). *Journal of Geophysical Research*, *114*(C12), C12023. https://doi.org/10.1029/2009jc005603

Jansen, M. F., Ferrari, R., & Mooring, T. A. (2010). Seasonal versus permanent thermocline warming by tropical cyclones. *Geophysical Research Letters*, *37*(3), L03602. https://doi.org/10.1029/2009gl041808

Jarosz, E., Mitchell, D. A., Wang, D. W., & Teague, W. J. (2007). Bottom-up determination of air-sea momentum exchange under a major tropical cyclone. *Science*, *315*(5819), 1707–1709. https://doi.org/10.1126/science.1136466

Jergensen, G. E., McGovern, A., Lagerquist, R., & Smith, T. (2020). Classifying convective storms using machine learning. *Weather and Forecasting*, *35*(2), 537–559. https://doi.org/10.1175/waf-d-19-0170.1

Jiang, G. Q., Xu, J., & Wei, J. (2018). A deep learning algorithm of neural network for the parameterization of typhoon-ocean feedback in typhoon forecast models. *Geophysical Research Letters*, *45*(8), 3706–3716. https://doi.org/10.1002/2018gl077004

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., & Neumann, C. J. (2010). The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, *91*(3), 363–376. https://doi.org/10.1175/2009bams2755.1

Kumar, C., Podestá, G., Kilpatrick, K., & Minnett, P. (2021). A machine learning approach to estimating the error in satellite sea surface temperature retrievals. *Remote Sensing of Environment*, *255*, 112227. https://doi.org/10.1016/j.rse.2020.112227

Lin, I., Liu, W. T., Wu, C. C., Wong, G. T. F., Hu, C., Chen, Z., et al. (2003). New evidence for enhanced ocean primary production triggered by tropical cyclone. *Geophysical Research Letters*, *30*(13), 10–13. https://doi.org/10.1029/2003gl017141

Lin, I. I. (2012). Typhoon-induced phytoplankton blooms and primary productivity increase in the western North Pacific subtropical ocean. *Journal of Geophysical Research*, *117*(3), C03039. https://doi.org/10.1029/2011jc007626

Lin, I.-I., Black, P., Price, J. F., Yang, C.-Y., Chen, S. S., Lien, C.-C., et al. (2013). An ocean coupling potential intensity index for tropical cyclones. *Geophysical Research Letters*, *40*(9), 1878–1882. https://doi.org/10.1002/grl.50091

Lin, I. I., Liu, W. T., Wu, C.-C., Chiang, J. C. H., & Sui, C.-H. (2003a). Satellite observations of modulation of surface winds by typhoon-induced upper ocean cooling. *Geophysical Research Letters*, *30*(3), 1131. https://doi.org/10.1029/2002gl015674

Liu, X., Wei, J., Zhang, D. L., & Miller, W. (2019). Parameterizing sea surface temperature cooling induced by tropical cyclones: 1. Theory and an application to typhoon Matsa (2005). *Journal of Geophysical Research: Oceans*, *124*(2), 1215–1231. https://doi.org/10.1029/2018jc014117

Liu, X., Zhang, D. L., & Guan, J. (2019). Parameterizing sea surface temperature cooling induced by tropical cyclones: 2. Verification by ocean drifters. *Journal of Geophysical Research: Oceans*, *124*(2), 1232–1243. https://doi.org/10.1029/2018jc014118

Liu, Y., Tang, D., Tang, S., Morozov, E., Liang, W., & Sui, Y. (2020). A case study of Chlorophyll a response to tropical cyclone Wind Pump considering Kuroshio invasion and air-sea heat exchange. *Science of the Total Environment*, *741*, 140290. https://doi.org/10.1016/j.scitotenv.2020.140290

Mei, W., Lien, C. C., Lin, I. I., & Xie, S. P. (2015). Tropical cyclone-induced ocean response: A comparative study of the South China sea and tropical northwest Pacific. *Journal of Climate*, *28*(15), 5952–5968. https://doi.org/10.1175/jcli-d-14-00651.1

Mei, W., & Pasquero, C. (2012). Restratification of the upper ocean after the passage of a tropical cyclone: A numerical study. *Journal of Physical Oceanography*, *42*(9), 1377–1401. https://doi.org/10.1175/jpo-d-11-0209.1

Mei, W., & Pasquero, C. (2013). Spatial and temporal characterization of sea surface temperature response to tropical cyclones. *Journal of Climate*, *26*(11), 3745–3765. https://doi.org/10.1175/jcli-d-12-00125.1

Mei, W., Pasquero, C., & Primeau, F. (2012). The effect of translation speed upon the intensity of tropical cyclones over the tropical ocean. *Geophysical Research Letters*, *39*(7), L07801. https://doi.org/10.1029/2011gl050765

Mei, W., & Xie, S.-P. (2016). Intensification of landfalling typhoons over the northwest Pacific since the late 1970s. *Nature Geoscience*, *9*(10), 753–757. https://doi.org/10.1038/ngeo2792

Mei, W., Xie, S.-P., Primeau, F., McWilliams, J. C., & Pasquero, C. (2015). Northwestern Pacific typhoon intensity controlled by changes in ocean temperatures. *Science Advances*, *1*(4), e1500014. https://doi.org/10.1126/sciadv.1500014

Palmen, E. (1948). On the formation and structure of tropical hurricanes. *Geophysica*, *3*(1), 26–38.

Pasquero, C., Desbiolles, F., & Meroni, A. N. (2021). Air-sea interactions in the cold wakes of tropical cyclones. *Geophysical Research Letters*, *48*(2), e2020GL091185. https://doi.org/10.1029/2020gl091185

Price, J. F. (1981). Upper ocean response to a hurricane. *Journal of Physical Oceanography*, *11*(2), 153–175. https://doi.org/10.1175/1520-0485(1981)011<0153:uortah>2.0.co;2

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007jcli1824.1

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., et al. (2022). Tackling climate change with machine learning. *ACM Computing Surveys*, *55*(2), 1–96. https://doi.org/10.1145/3485128

Shao, Q., Li, W., Han, G., Hou, G., Liu, S., Gong, Y., & Qu, P. (2021). A deep learning model for forecasting sea surface height anomalies and temperatures in the South China Sea. *Journal of Geophysical Research: Oceans*, *126*(7), e2021JC017515. https://doi.org/10.1029/2021jc017515

Shay, L. K., Black, P. G., Mariano, A. J., Hawkins, J. D., & Elsberry, R. L. (1992). Upper ocean response to Hurricane Gilbert. *Journal of Geophysical Research*, *97*(C12), 20227–20248. https://doi.org/10.1029/92jc01586

Sriver, R. L., & Huber, M. (2007). Observational evidence for an ocean heat pump induced by tropical cyclones. *Nature*, *447*(7144), 577–580. https://doi.org/10.1038/nature05785

Sui, Y., Sheng, J., Tang, D., & Xing, J. (2022). Study of storm-induced changes in circulation and temperature over the northern South China Sea during Typhoon Linfa. *Continental Shelf Research*, *249*, 104866. https://doi.org/10.1016/j.csr.2022.104866

Sun, J., Vecchi, G. A., & Soden, B. J. (2021). Influence of vertical wind shear on the ocean response to tropical cyclones based on satellite observations. *Geophysical Research Letters*, *48*(20), e2021GL095451. https://doi.org/10.1029/2021gl095451

Sun, R., Villas Bôas, A. B., Subramanian, A. C., Cornuelle, B. D., Mazloff, M. R., Miller, A. J., et al. (2022). Focusing and defocusing of tropical cyclone generated waves by ocean current refraction. *Journal of Geophysical Research: Oceans*, *127*(1), e2021JC018112. https://doi.org/10.1029/2021jc018112

Tuv, E., Borisov, A., Runger, G., & Torkkola, K. (2009). Feature selection with ensembles, artificial variables, and redundancy elimination. *Journal of Machine Learning Research*, *10*, 1341–1366.

Vincent, E. M., Lengaigne, M., Madec, G., Vialard, J., Samson, G., Jourdain, N. C., et al. (2012). Processes setting the characteristics of sea surface cooling induced by tropical cyclones. *Journal of Geophysical Research*, *117*(C2), C02020. https://doi.org/10.1029/2011jc007396

Vincent, E. M., Lengaigne, M., Vialard, J., Madec, G., Jourdain, N. C., & Masson, S. (2012). Assessing the oceanic control on the amplitude of sea surface cooling induced by tropical cyclones. *Journal of Geophysical Research*, *117*(C5), C05023. https://doi.org/10.1029/2011jc007705

Wang, G., Wu, L., Johnson, N. C., & Ling, Z. (2016). Observed three-dimensional structure of ocean cooling induced by Pacific tropical cyclones. *Geophysical Research Letters*, *43*(14), 7632–7638. https://doi.org/10.1002/2016gl069605

Wang, Y.-q., & Wu, C.-C. (2004). Current understanding of tropical cyclone structure and intensity changes–a review. *Meteorology and Atmospheric Physics*, *87*(4), 257–278. https://doi.org/10.1007/s00703-003-0055-6

Wei, J., Jiang, G. Q., & Liu, X. (2017). Parameterization of typhoon-induced ocean cooling using temperature equation and machine learning algorithms: An example of typhoon Soulik (2013). *Ocean Dynamics*, *67*(9), 1179–1193. https://doi.org/10.1007/s10236-017-1082-z

Wentz, F. J., Gentemann, C., Smith, D., & Chelton, D. (2000). Satellite measurements of sea surface temperature through clouds. *Science*, *288*(5467), 847–850. https://doi.org/10.1126/science.288.5467.847

Yang, X., & Tang, D. (2010). Location of sea surface temperature cooling induced by typhoon in the South China Sea. *Journal of Tropical Oceanography*, *29*(4), 26–31.

Zhang, D.-L., Zhu, L., Zhang, X., & Tallapragada, V. (2015). Sensitivity of idealized hurricane intensity and structures under varying background flows and initial vortex intensities to different vertical resolutions in HWRF. *Monthly Weather Review*, *143*(3), 914–932. https://doi.org/10.1175/mwr-d-14-00102.1

Zhang, H. (2023). Modulation of upper ocean vertical temperature structure and heat content by a fast-moving tropical cyclone. *Journal of Physical Oceanography*, *53*(2), 493–508. https://doi.org/10.1175/jpo-d-22-0132.1

Zhang, H., He, H., Zhang, W.-Z., & Tian, D. (2021). Upper ocean response to tropical cyclones: A review. *Geoscience Letters*, *8*(1), 1. https://doi.org/10.1186/s40562-020-00170-8

Zhang, H., Jing, M., Zhang, H., Li, L., Zheng, Y., Tang, J., et al. (2023). Deep learning approach for forecasting sea surface temperature response to tropical cyclones in the Western North Pacific. *Deep Sea Research Part I: Oceanographic Research Papers*, *197*, 104042. https://doi.org/10.1016/j.dsr.2023.104042

Zhang, J., Lin, Y., Chavas, D. R., & Mei, W. (2019). Tropical cyclone cold wake size and its applications to power dissipation and ocean heat uptake estimates. *Geophysical Research Letters*, *46*(16), 10177–10185. https://doi.org/10.1029/2019gl083783

Zhao, Z., Zhou, J., & Du, H. (2022). Artificial intelligence powered forecast of oceanic mesoscale phenomena: A typhoon cold wake case occurring in Northwest Pacific Ocean. *Future Generation Computer Systems*, *129*, 389–398. https://doi.org/10.1016/j.future.2021.10.031

Zheng, G. M., & Tang, D. (2007). Offshore and nearshore chlorophyll increases induced by typhoon winds and subsequent terrestrial rainwater runoff. *Marine Ecology Progress Series*, *333*, 61–74. https://doi.org/10.3354/meps333061

Zheng, Z. W., Ho, C. R., & Kuo, N. J. (2008). Importance of pre-existing oceanic conditions to upper ocean response induced by Super Typhoon Hai-Tang. *Geophysical Research Letters*, *35*(20), 1–5. https://doi.org/10.1029/2008gl035524

Zhu, T., & Zhang, D.-L. (2006). The impact of the storm-induced SST cooling on hurricane intensity. *Advances in Atmospheric Sciences*, *23*(1), 14–22. https://doi.org/10.1007/s00376-006-0002-9