

Combining LSTM statistical analysis with dynamical models to investigate Typhoon Mangkhut (2018)

Clifford Watkins^a and Sue Chen^b

^a National Research Council, National Academy of Sciences, USA

^b Marine Meteorology Division, Naval Research Laboratory, Monterey, California, USA

To be submit to GRL

Corresponding author: Clifford Watkins: clifford.watkins.ctr@nrlmry.navy.mil

DISTRIBUTION A: Approved for public release, distribution is unlimited.

Abstract

The study develops a long short-term memory mode (LSTM) neural network algorithm for the prediction of tropical cyclone intensity (maximum wind speed). This is achieved by combing the statistical tools of machine learning and the dynamical earth system model of Typhoon Mangkhut (2018), The Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System is used to produce a high dimensional dataset from an ensemble of runs of Mangkhut to train the LSTM. Being able to predict the behavior of tropical cyclones allows for both an early warning of behavior and as a way to analyze the complex dynamical inputs to investigate the driving forces behind periods of rapid intensification, stagnation, and weakening. For Mangkhut case, we found that the controlling dynamic for rapid intensification was the balance between asymmetry near the eyewall and the available thermal energy as measured by heat flux to the atmosphere. The input parameters to LSTM that produced the best prediction are the mean inner surface heat flux, amplitude of the first order asymmetry in the inner heat flux, the angle difference between the first order asymmetries of vorticity output at 850mb and 500mb, and the amplitude of first order asymmetry in the geopotential height at 850mb.

1 Introduction:

Forecasting the intensity of tropical cyclones (TCs) remains a complicated problem due to a combination of multiscale dynamics^{1,2} and dependence on environmental factors such as ocean temperature^{3,4}, steering forces⁵, and shear conditions⁶. The forecast of TC track has increased in accuracy over the preceding decades, however progress on TC rapid intensification (RI) forecast has lagged behind⁷. As such intensification rates of major (cat 3-5), open-ocean TCs remain a key and potentially deadly error^{1,8}. Another mechanism observed in the co-evolution of ocean and atmosphere is the rotation of the momentum fluxes associated with eye-passage that alters the turbulent mechanisms of entrainment into inertial and internal waves in the ocean surface boundary layer to

lead to prolonged deepening of the atmosphere and ocean mixed layer associated with the cold wake and cold wake interaction⁹.

The Northwest Pacific Basin (NWPB) represents a special case, where the conditions commonly focused on for intensification in the North Atlantic basin such as ocean heat content and background shear^{10,11} are more readily available in the NWPB^{12,13}. Specifically, the Western North Pacific warm pool dipole is a deep and persistent oceanic reservoir of high temperature waters to fuel TC development^{14,15}. This ready energy for strengthening intimates that the timing of intensification and the maximum sustained strength reached depend on predicting the evolution of storm-scale features as the TC progresses through the NWPB (e.g. ^{6,16,17}).

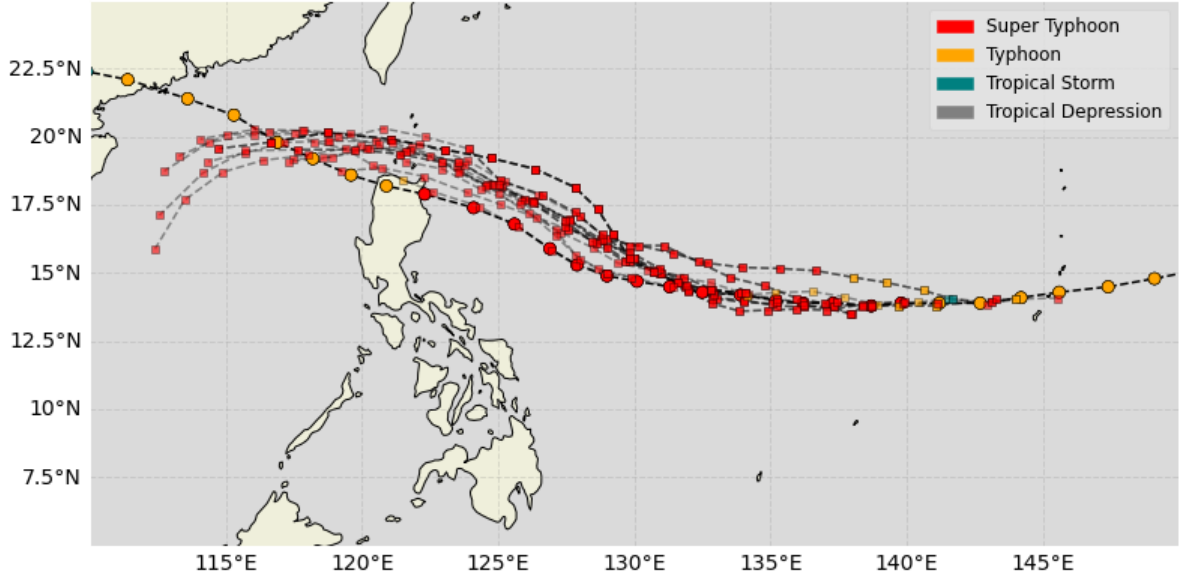


Figure 1 shows the best track for Typhoon Mangkhut (2018) located in the NWPB and strength with circular markers and the tracks and strengths from the COAMPS ensemble.

Typhoon Mangkhut represents a good test-bed for predictive studies in the WNPB causing damage in Guam, the Philippines, and Southern China in September 2018¹⁸. After a period of almost a week in conditions of low shear and high ocean heat content transiting from the Marianas Islands, Mangkhut made landfall in North of the Philippines at 1800 UTC on the 14th of September as a category 5 equivalent super typhoon. Understanding and predicting the intensity of such dangerous storms is imperative and we seek to narrow the gap in rapid intensification forecast by leveraging machine learning techniques.

Recent studies have shown that under strong TC forcing the ocean evolves at time-scales as to influence TC intensity predictions⁹. And where it is commonly known that sea surface temperature raises the potential for rapid

intensification^{3,19}, the co-evolution of the ocean under a rapidly-intensifying TC represents a key pillar of not only the initial phase of rapid intensification, but the ability of the ocean-atmospheric system to sustain the newly-strengthened storm. Utilizing the machine learning technique of a long short-term memory mode (LSTM) neural network model²⁰ in conjunction with coupled dynamical model output for TCs in the NWPB allows for the analysis of large amounts of physically relevant data to illuminate processes that modulate TC intensity.

2 Data and Methods:

By utilizing the Navy’s Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) model^{21,22}, we generated an ensemble of COAMPS simulated Mangkhut TC events, starting from 0600 UTC September 10th, 2018 to 1800 on the 12th every 6 hours. The capacity for COAMPS to accurately predict TC intensity in the NWPB is competitive with other dynamical models²². By setting ensemble members start at different times, the TC was able to have multiple different starting locations, intensities, and sizes, thus increasing the range variables. The initial ocean state variables are taken from the GOFS 3.5: 41-layer HYCOM²³ and the initial and lateral boundary atmospheric conditions are from Global Forecast System (GFS). The atmospheric model is coupled with the Navy Coastal Ocean Model (NCOM^{24,25}) with 45 vertical layers. The horizontal resolution of the atmospheric and ocean models is 4 km with 10 min coupling frequency. Mangkhut’s ocean evolution was captured by the Scripps Institute of Oceanography float array²⁶ observation during the 2018 Propagation of Intra-Seasonal Tropical Oscillations (PISTON) Office of Naval Research filed campaign²⁷ but is not assimilated into the Navy Coupled Ocean Data Assimilation system²⁸ (NCODA).

The COAMPS run initialized on 2018/9/10 at 0000 is used as the ground truth against the LSTM, once trained on the other 11 members of the ensemble, is evaluated. Figure 1 illustrates the consistent grouping of TC tracks from the COAMPS ensemble sweeping North over Luzon, differing in a key manner from the best track where the TC went over land in the Philippines²⁹. The land interaction caused both extreme damage and lowered the intensity of the observed storm³⁰. Despite the model track and intensity biases, the model ensemble was self-consistent in track and intensity biases due to consistent boundary conditions and land-interactions in all the runs. Thus the storms from the ensemble represented good training for the validation run without having to account for differential atmosphere and ocean pre-storm environment encountered due to variability of track forecasts. Moreover, COAMPS permits for a much more dense data output than contemporary machine learning models of TC behavior from the Statistical Hurricane Intensity Prediction Scheme^{2,7} (SHIPS) and coarser-resolution reanalysis data^{11,31,32}. Hence by increasing the density of training data, we can evaluate whether our model can parse more of the physical forces and interactions underpinning the development of TCs use one TC ensemble.

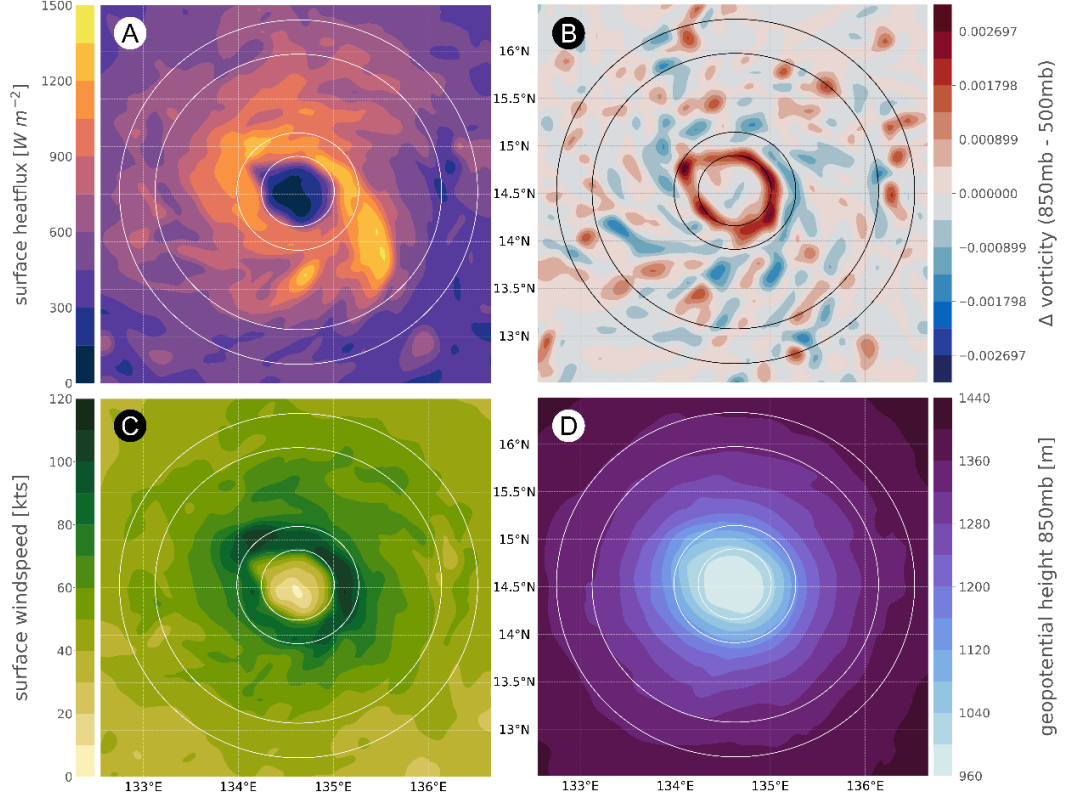


Figure 2 shows the TC-centered maps of COAMPS fields used in the final LSTM: a) latent and sensible surface heat flux (W m^{-2}), b) the difference in vorticity between the mid and low levels as output at 500mb and 850mb respectively (S^{-1}), c) the wind field (kt), and d) the geopotential heights (m). In all plots the inner two circles encompass the region near the eyewall set between 0.8 and 1.2 times the radius of maximum winds and the further two concentric rings bound the far field set at 160-200km from the eye regions are shown, with the goal of finding asymmetry in eye structure and environmental boundaries respectively.

The LSTM neural network is suitable to assess the state conditions with a TC intensity predictor as presented by Li et al. (2017) and Yang et al. (2020). Additionally, we can fully assess the influence of the ocean that has both ahead-of and after eye passage as well as on the right and left side of the TC because COAMPS atmosphere is two-way coupled with a 3D NCOM ocean model that evolves and interacts with the TC^{33,34}. The variables at the ocean surface, which include sea surface temperature, salinity, and heat flux, are input as potential variables for the LSTM as well as the reciprocal fluxes in the atmospheric data.

3 LSTM Setup and Variables:

The LSTM we developed uses the hourly output of COAMPS variables as se-

quential inputs. Rather than requiring the machine learning model to deal with the evolution of 4D fields output from the numerical models over time, the data is compressed to regions of interest as shown in recent research of the inner core ocean interaction and structure^{16,35,36} and the outer region TC response to the regional-scale conditions^{12,18}. Thus, the variables used in the LSTM are garnered by taking the radial averages and first order perturbation at the regions around the radius of maximum winds ($0.8 - 1.2$ times r_{\max}) and farther afield (160-200 km) from the center of the storm determined by minimum atmospheric pressure at sea level. By using the variable eye-radius as the inner metric, we were able to minimize the effect of diurnal cycles, eye-wall replacement cycles, and other higher frequency perturbations of eyewall radius that added non-predictive signals to averaging across any predesignated radius. Then in order to ascertain the first order perturbations of the variable fields, the circumference of the storm was segmented into 15 wedges. A sine function was then fitted to the residuals of the variable after the mean was taken out, outputting an amplitude and phase of the perturbation. This permits the known spatial variability of physically meaningful processes^{5,9,19} to be assessed and used in lieu of oversampling the RI events as proposed in Yang et al. (2020). In order to aid the LSTM model in dealing with large difference in magnitude of the variables, we normalized each to range of zero to one across all simulation runs of Mangkhut.

The advantage of the LSTM methodology follows from the ability to generate state-functions, which act like wave functions that can be superimposed to maximize the probability of predicting intensity, while the forgetting gates can minimize overfitting^{31,32}. The sequential nature of the data yields a data structure that looks at time t and each hour's data from the 5 hours previous to forecast the maximum winds at time t' .

4 Variable selection and Results:

Using a mean absolute error (MAE) as the metric to evolve the LSTM parameters and allowing the data to be processed over the ensemble for 30 epochs, the LSTM performed the fitting. We ran the model with every selection of 3, 4 and 5 variables as predictive for maximum wind speed. The resultant predictive models were compared using Akaike information criterion, which seeks to weigh the advantages of including more variables against the predictive accuracy³⁷. Thus, the best model for Mangkhut was developed from including the following variable inputs: mean inner surface heat flux, amplitude of the first order asymmetry in the inner heat flux, the angle difference between the first order asymmetries of vorticity output at 850mb and 500mb, and the amplitude of first order asymmetry in the geopotential height at 850mb. The resultant MAE across the evolution of the storm is 2.75 kts. Figure 3 shows the resultant predictions developed from the LSTM, illustrating the models ability to capture both periods of rapid intensification and arresting of intensification into a steady state storm. Allowing us to grasp that the physical constraints most controlling the intensity of Mangkhut represent a balance of the symmetry of the storm and the total heat flux, which dropped precipitously as the TC crossed the warm

pool dipole into the South China Sea. Despite this drop in available energy, the TC was able to maintain strength due to the symmetry of core processes.

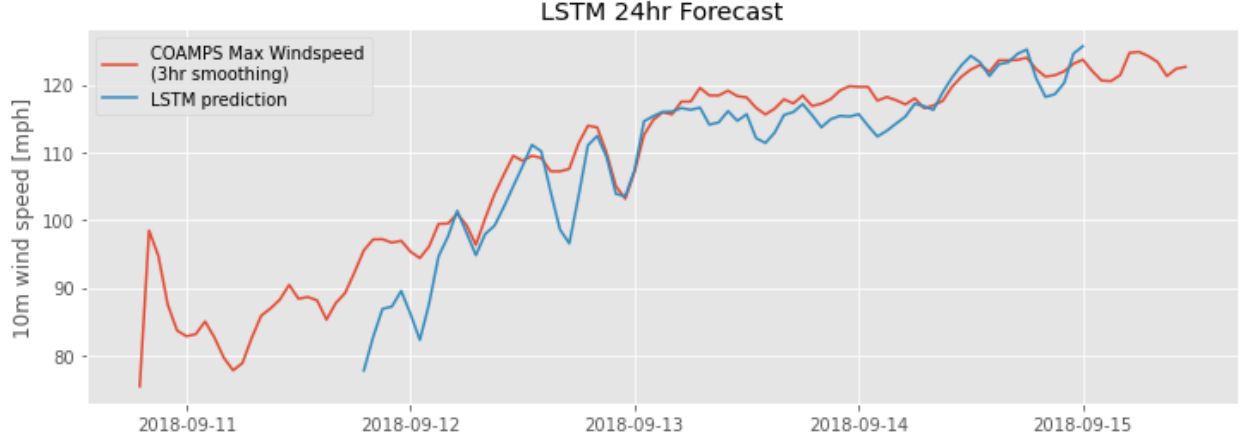


Figure 3 shows the predictive power of the LSTM. The red line is the smoothed maximum wind speed from the COAMPS run and the blue line represents the value as predicted 24 hours in advance based on the time and five hours previous from the LSTM inputs.

5 Conclusion and Discussion:

An LSTM model for forecasting the intensity of tropical cyclones within the model realm of COAMPS has been developed with a high resolution coupled ensemble of Typhoon Mangkhut (2018). By utilizing the information from high resolution dynamical models, our predictive model of TC intensity captures physical dynamics in a way that previous work^{31,32}, which used single storm values in SHIPS could not. Thus our LSTM, while similar in statistical technique and only single storm, could intimate more of the processes behind the TC intensity trends. In the case of Mangkhut, this led to an illumination of the role of eyewall asymmetry balancing the total heat flux in the inner core region. The model variables offer physical insights into the development and sustained intensity of Mangkhut. The mean inner surface heat flux reflects the available energy for the TC determined mainly by the air-sea surface temperature difference and wind speeds. The first order asymmetry of the inner heat flux, the correlated deviation of the eyewall winds with the available ocean energy, was dominated by the asymmetry in the wind field for the majority of Mangkhut but also included the variance due to ocean conditions. Similarly to the asymmetry in the heat flux, the difference between the angle of vorticity asymmetry between 850mb and 500mb in the inner region illuminates a portion of lower and middle atmosphere vortex alignment and inertial stability. Asymmetry of the geopotential height at 850mb, the last of the input variables, effectively shows the variance in the boundary layer static stability around the TC³⁸. And by using the asymmetries in heat flux, vortex angle, and boundary layer height, the

underlying influence of the atmosphere-ocean boundary is included along with the rotational stability of the TC, allowing for the model to perceive both atmospheric and ocean influence in TC development^{16,17,39,40}, which in combination led to a better fitness of model than any asymmetry individually. In all, we see the intensification of Mangkhut is limited by the lower heat fluxes once the TC is over the South China Sea, but the intensity sustains as the storm becomes more symmetrical in slightly different ways in each COAMPS ensemble run.

Additionally, the model was developed using a framework that is repeatable and can be extend in the scope of many storms and multiple basins. Being able to rapidly predict the behavior of TC intensity, even with model environments, will allow for more rapid accurate warnings to be developed. Additionally, the use of physical variables and intuitive structure allowed for a deeper understanding of the dynamics controlling the TC. Specifically, Mangkhut illustrates the importance of inner core dynamics in the stability and growth of developed TCs. By combining the developments in machine learning with dynamical models, we can garner more insight than either alone.

6 Acknowledgements

This research was performed while the corresponding author held a National Research Council Research Associateship Award at the U.S. Naval Research Lab, Monterey. Chen is supported by the Office of Naval Research grand N0001422WX00342. The computation resource was provided by the Department of Defense (DoD) High Performance Computing Modernization Program (HPCMP).

Data Availability Statement

All data used in this study is stored at the U.S. Navy DoD Supercomputing Resource Center (DSRC). Access to the DSRC may be obtained through a request to HPCMP (<https://www.hpc.mil/>). Once an account has been established, the authors may be contacted for information to access the archived data. The final processed model data used to generate the figures in the study can be accessed at <https://doi.org/10.6084/m9.figshare.17267396.v1>.

Figure Captions.

1. Figure 1 shows the best track for Typhoon Mangkhut (2018) located in the NWPB and strength with circular markers and the tracks and strengths from the COAMPS ensemble.
2. Figure 2 shows the TC-centered maps of COAMPS fields used in the final LSTM: A) latent and sensible surface heat flux (W m^{-2}), B) the difference in vorticity between the mid and low levels as output at 500mb and 850mb respectively (S-1), C) the wind field (kt), and D) the geopotential heights (m). In all plots the inner two circles encompass the region near the eyewall set between 0.8 and 1.2 times the radius of maximum winds and the further two concentric rings bound the far field set at 160-200km from

the eye regions are shown, with the goal of finding asymmetry in eye structure and environmental boundaries respectively.

3. Figure 3 shows the predictive power of the LSTM. The red line is the smoothed maximum wind speed from the COAMPS run and the blue line represents the value as predicted 24 hours in advance based on the time and five hours previous from the LSTM inputs.

References.

1. Alessandrini, S., Delle Monache, L., Rozoff, C. M. & Lewis, W. E. Probabilistic Prediction of Tropical Cyclone Intensity with an Analog Ensemble. *Mon. Weather Rev.* **146**, 1723–1744 (2018).
2. DeMaria, M. A Simplified Dynamical System for Tropical Cyclone Intensity Prediction. *Mon. Weather Rev.* **137**, 68–82 (2009).
3. Rudzin, J. E., Shay, L. K. & Johns, W. E. The Influence of the Barrier Layer on SST Response during Tropical Cyclone Wind Forcing Using Idealized Experiments. *J. Phys. Oceanogr.* **48**, 1471–1478 (2018).
4. Park, M.-S. Vertical Wind Shear and Ocean Heat Content as Environmental Modulators of Western North Pacific Tropical Cyclone Intensification and Decay. **1**, 10 (2012).
5. Hobgood, J. S. The Effects of Climatological and Persistence Variables on the Intensities of Tropical Cyclones over the Eastern North Pacific Ocean. *WEATHER Forecast.* **13**, 8 (1998).
6. Lee, T.-Y., Wu, C.-C. & Rios-Berrios, R. The Role of Low-Level Flow Direction on Tropical Cyclone Intensity Changes in a Moderate-Sheared Environment. *J. Atmospheric Sci.* (2021) doi:10.1175/JAS-D-20-0360.1.
7. DeMaria, M., Sampson, C. R., Knaff, J. A. & Musgrave, K. D. Is Tropical Cyclone Intensity Guidance Improving? *Bull. Am. Meteorol. Soc.* **95**, 387–398 (2014).
8. Wang, Y. & Wu, C.-C. Current understanding of tropical cyclone structure and intensity changes? a review. *Meteorol. Atmospheric Phys.* **87**, 257–278 (2004).
9. Chen, S., Elsberry, R. L. & Harr, P. A. Modeling Interaction of a Tropical Cyclone with Its Cold Wake. *J. Atmospheric Sci.* **74**, 3981–4001 (2017).
10. Grimes, A. & Mercer, A. E. Synoptic-Scale Precursors to Tropical Cyclone Rapid Intensification in the Atlantic Basin. *Adv. Meteorol.* **2015**, 1–16 (2015).
11. Mercer, A. & Grimes, A. Atlantic Tropical Cyclone Rapid Intensification Probabilistic Forecasts from an Ensemble of Machine Learning Methods. *Procedia Comput. Sci.* **114**, 333–340 (2017).
12. Ying, M. Impacts of Climate Change on Tropical Cyclones in the Western North Pacific Basin. Part II: Late Twenty-First Century Projections. **1**, 11 (2012).
13. Yu, H., Chen, P., Li, Q. & Tang, B. Current Capability of Operational Numerical Models in Predicting Tropical Cyclone Intensity in the Western North Pacific. *Weather Forecast.* **28**, 353–367 (2013).
14. Wang, C. & Wang, B. Tropical cyclone predictability shaped by western Pacific subtropical high: integration of trans-basin sea surface temperature effects. *Clim. Dyn.* **53**, 2697–2714 (2019).
15. Zhan, R., Wang, Y. & Wen, M. The SST Gradient between the Southwestern Pacific and the Western Pacific Warm Pool: A New Factor Controlling the Northwestern Pacific Tropical Cyclone Genesis Frequency. *J. Clim.* **26**, 2408–2415 (2013).
16. Bhalachandran, S. *et al.* Characterizing the Energetics of Vortex-Scale and Sub-Vortex-Scale Asymmetries during Tropical Cyclone Rapid Intensity Changes. *J.*

Atmospheric Sci. **77**, 315–336 (2020).17. Callaghan, J. Asymmetric Inner Core Convection Leading to Tropical Cyclone Intensification. **6**, 12 (2017).18. Liu, S. *et al.* Basin-wide responses of the South China Sea environment to Super Typhoon Mangkhut (2018). *Sci. Total Environ.* **731**, 139093 (2020).19. Yang, J. *et al.* A Comparative Study of Typhoon Hato (2017) and Typhoon Mangkhut (2018)—Their Impacts on Coastal Inundation in Macau. *J. Geophys. Res. Oceans* **124**, 9590–9619 (2019).20. Hua, Y. *et al.* Deep Learning with Long Short-Term Memory for Time Series Prediction. *ArXiv181010161 Cs* (2018).21. Chen, S. *et al.* COAMPS version 3 model description: General theory and equations. *NRL Publ NRLPU7500-03* **448**, 145 (2003).22. Doyle, J. *et al.* Tropical Cyclone Prediction Using COAMPS-TC. *Oceanography* **27**, 104–115 (2014).23. Metzger, E. J. *et al.* Simulated and observed circulation in the Indonesian Seas: 1/12° global HYCOM and the INSTANT observations. *Dyn. Atmospheres Oceans* **26** (2010).24. Barron, C. N., Kara, A. B., Martin, P. J., Rhodes, R. C. & Smedstad, L. F. Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). *Ocean Model.* **11**, 347–375 (2006).25. Jensen, T. G., Campbell, T. J., Allard, R. A., Small, R. J. & Smith, T. A. Turbulent heat fluxes during an intense cold-air outbreak over the Kuroshio Extension Region: results from a high-resolution coupled atmosphere–ocean model. *Ocean Dyn.* **61**, 657–674 (2011).26. Ni, Z. *et al.* Response of the upper ocean to tropical cyclone in the Northwest Pacific observed by gliders during fall 2018. *Acta Oceanol. Sin.* **40**, 103–112 (2021).27. Rydbeck, A. V., Jensen, T. G. & Flatau, M. Characterization of Intraseasonal Kelvin Waves in the Equatorial Pacific Ocean. *J. Geophys. Res. Oceans* **124**, 2028–2053 (2019).28. Lunde, B. N. & Coelho, E. F. Implementations of the Navy Coupled Ocean Data Assimilation system at the Naval Oceanographic Office. in *OCEANS 2009* 1–10 (IEEE, 2009). doi:10.23919/OCEANS.2009.5422232.29. Kieu, C. *et al.* Track Dependence of Tropical Cyclone Intensity Forecast Errors in the COAMPS-TC Model. *Weather Forecast.* **36**, 469–485 (2021).30. Abancó, C., Bennett, G. L., Matthews, A. J., Matera, M. A. M. & Tan, F. J. The role of geomorphology, rainfall and soil moisture in the occurrence of landslides triggered by 2018 Typhoon Mangkhut in the Philippines. *Nat. Hazards Earth Syst. Sci.* **21**, 1531–1550 (2021).31. Li, Y. *et al.* Leveraging LSTM for rapid intensifications prediction of tropical cyclones. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **IV-4/W2**, 101–105 (2017).32. Yang, Q., Lee, C.-Y. & Tippett, M. K. A Long Short-Term Memory Model for Global Rapid Intensification Prediction. *Weather Forecast.* **35**, 1203–1220 (2020).33. Martin, P. J. *Description of the NAVY coastal ocean model version 1.0*. *NRL Rep.* (2000).34. Komaromi, W. A., Reinecke, P. A., Doyle, J. D. & Moskaitis, J. R. The Naval Research Laboratory’s Coupled Ocean–Atmosphere Mesoscale Prediction System–Tropical Cyclone Ensemble (COAMPS-TC Ensemble). *Weather Forecast.* **36**, 499–517 (2021).35. Williams, G. J. Idealized Simulations of the Inner Core Boundary Layer Structure in a Landfalling Tropical Cyclone. Part I: Kinematic Structure. *Trop. Cyclone Res. Rev.* **8**, 47–67 (2019).36. Knapp, K. R., Velden, C. S. & Wimmers, A. J. A Global Climatology of Tropical Cyclone Eyes. *Mon. Weather Rev.* **146**, 2089–2101 (2018).37. Akaike, H. A new look at

the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723 (1974).38. Vigh, J. L. & Schubert, W. H. Rapid Development of the Tropical Cyclone Warm Core. *J. Atmospheric Sci.* **66**, 3335–3350 (2009).39. Ahern, K., Hart, R. E. & Bourassa, M. A. Asymmetric Hurricane Boundary Layer Structure during Storm Decay. Part I: Formation of Descending Inflow. *Mon. Weather Rev.* **149**, 3851–3874 (2021).40. Zhu, H., Ulrich, W. & Smith, R. K. Ocean Effects on Tropical Cyclone Intensification and Inner-Core Asymmetries. *J. Atmospheric Sci.* **61**, 1245–1258 (2004).