COMMUNICATIONS EARTH&ENVIRONMENT

ARTICLE

https://doi.org/10.1038/s43247-020-00033-2 OPEN

Pacific subsurface ocean temperature as a long-range predictor of South China tropical cyclone landfall

Nathan Sparks [1]✉ & Ralf Toumi[1]

Seasonal forecasts of the tropical cyclones which frequently make landfall along the densely populated South China coast are highly desirable. Here, we analyse observations of land-falling tropical cyclones in South China and of subsurface ocean temperatures in the Pacific warm pool region, and identify the possibility of forecasts of South China tropical cyclone landfall a year ahead. Specifically, we define a subsurface temperature index, subNiño4, and build a predictive model based on subNiño4 anomalies with a robust double cross-validated forecast skill against climatology of 23%, similar in skill to existing forecasts issued much later in the spring. We suggest that subNiño4 ocean temperatures precede the surface El Niño/Southern Oscillation state by about 12 months, and that the zonal shifts in atmospheric heating then change mid-level winds to steer tropical cyclones towards landfall in South China. We note that regional subsurface ocean temperature anomalies may permit atmospheric predictions in other locations at a longer range than is currently thought possible.

[1] Imperial College London, London, UK. ✉email: nsparks@ic.ac.uk

Tropical cyclones (TCs) are among the most harmful natural hazards in terms or both economic and human cost. There is, therefore, great interest in the prediction of the seasonal activity and particularly landfall of TCs among many organisations, including governments and business. There is demand for forecasts with longer lead times of landfall to permit planning and preparation[1]. The South China Sea region is of special importance as the highest value at-risk region to tropical cyclones in the world. The region includes the largest mega-city in the world[2], an exposed GDP of over 1 Trillion US$ and accounts for a third of global shipping[3]. Globally, it is also one of the coastlines most affected by TCs. The seasonal forecasting of TC activity in this region is, therefore, of particular interest. Existing forecasts are typically issued only just before the start of the TC season in March/April and are often only on the basin scale rather than landfall[4].

TC activity in the western North Pacific has significant inter-annual variability[5] that is related to the El Niño-Southern Oscillation (ENSO) phenomenon[6,7]. While it is generally accepted that ENSO is not strongly correlated with contemporaneous basin TC counts[8], a delayed suppression of TC activity in response to strong El Niño events linked to changes in large-scale circulation has been reported[9]. All the current seasonal forecasting models for South China landfall frequency are based on some level of ENSO sea surface temperature (SST) predictability[10–13]. However, none of these forecasts can be issued before the Spring. Subsurface ocean temperature is a known important variable in seasonal prediction and has been assimilated into dynamical ocean-atmosphere forecasting models since the 1990s[14]. However, the subsurface ocean temperature has only been used with short lead times of 3 months or less for Indian Summer Monsoon rain[15] and seasonal Atlantic hurricane activity[16]. Here, we show a step change in the forecast lead time of an atmospheric phenomenon, landfalling TCs, to 1 year using equatorial Pacific subsurface ocean temperature.

Results

Subsurface ocean temperature and landfall. We find that Pacific equatorial subsurface ocean temperature anomaly around the date line, between the depths of 100–300 m, is strongly anti-correlated with the following year's South China tropical cyclone landfall count. The anomaly in prior summer (June–August, JJA-1) was found to have the strongest anti-correlation and is shown in Fig. 1a. The region of significant anti-correlation is largely reduced outside the depth range 100–300 m. This shows that the predictive subsurface ocean temperature signal is isolated and not connected with the surface at that time. Temperature anomaly information from below 100 m near the thermocline depths is needed. In Fig. 1b, c, we show the mean and inter-annual standard deviation of equatorial Pacific summer subsurface temperature, respectively. The maximum anti-correlation occurs at the depth of the thermocline. At this depth the vertical gradient of temperature is a maximum separating the Pacific warm pool water above from cold water below. This is also the depth at which the maximum inter-annual variability occurs.

To further investigate the location of predictability we calculate the correlation of Pacific summer equatorial subsurface ocean temperature anomaly (JJA-1) averaged between depths of 100–300 m with South China TC landfall count (Fig. 2a). The region of correlation significant at the 99% confidence level extends widely in the equatorial Central Pacific centred near the date line, it has a similar lateral extent to the standard Niño 4 index definition, which includes the Pacific warm pool. In the vertical the subNiño4 region is also bounded by the eastward Pacific equatorial undercurrent at the same depths as identified here[17]. We define the subNiño4 index as the mean JJA-1 subsurface ocean temperature anomaly at the bottom of the warm pool averaged over depths 100–300 m and the Niño 4 region, 160°E–150°W, 5°S–5°N. The Pearson's correlation coefficient (r) of subNiño4 with the next year's seasonal (May–November)

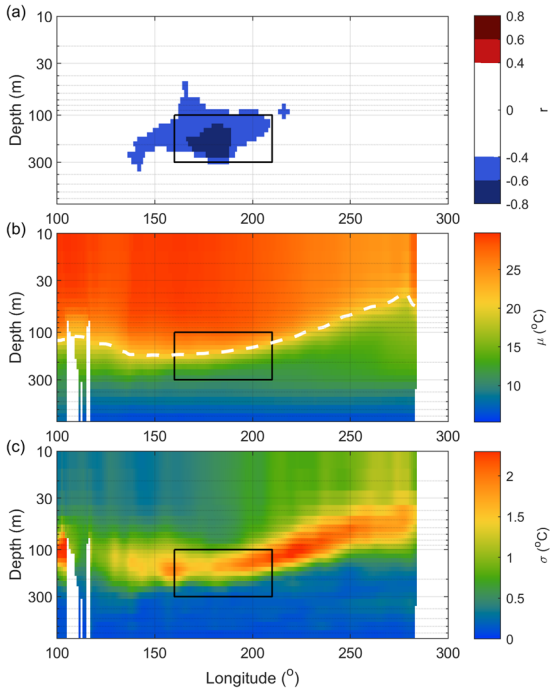


Fig. 1 Equatorial (5°S–5°N) Pacific longitude-depth profiles. Pearson's correlation coefficient (r) of May–November South China landfall counts and the preceding year's June–August (JJA-1) subsurface ocean temperature anomaly in **a**. Correlations significant at the 99% confidence level using a two-tailed t-test are shaded. Climatological **b** mean, and **c** inter-annual standard deviation of JJA subsurface ocean temperature. Black box shows region selected as subNiño4 predictor index. Dotted white line is the thermocline defined at 20 °C.

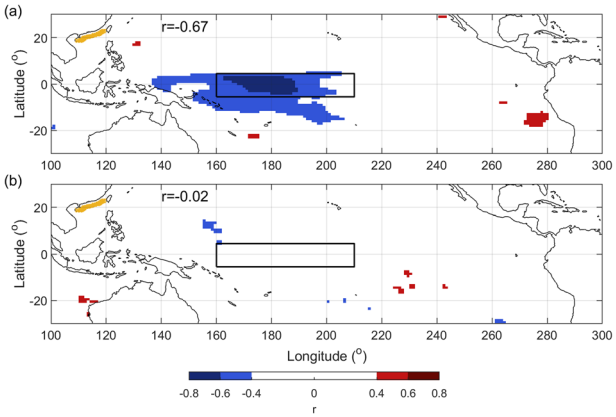


Fig. 2 Correlation of tropical Pacific summer subsurface ocean temperature and SST with following year's South China landfall counts. Pearson's correlation coefficient (r) between May–November South China landfall counts and the preceding year's June–August (JJA-1) **a** 100–300 m subsurface ocean temperature, and **b** sea surface temperature. The correlation coefficient for each variable averaged across the black box (Niño 4) is shown. Correlations significant at the 99% confidence level using a two-tailed t-test are shaded. Yellow markers show position of the landfall events in south China.

South China landfall count is $r = 0.67$ ($p < 0.001$). Therefore, cold (warm) summer subsurface ocean temperatures one year prior lead to high (low) South China landfall counts the next year. The correlation of subNiño4 with landfall counts derived from other TC warning agencies are also all highly significant ($p < 0.01$) (Supplementary Table 1). We note that the correlation is sensitive to the lifetime maximum intensity of the TCs making landfall (Supplementary Table 1). The prior summer sea surface temperature (SST) in the Pacific is not correlated with the following year's South China landfall count (Fig. 2b). The basin wide accumulated cyclone energy (ACE) is also correlated with the prior summer subNiño4 index ($r = 0.45$, $p = 0.006$). However, the correlation of the subNiño4 index with total number of tropical cyclones in the North West Pacific basin is weak ($r = -0.26$, $p = 0.12$).

The source of predictability a year in advance lies in the temperature below the ocean surface. The depth of the equatorial thermocline on the date line is an alternative good predictor of the next year's South China landfall count ($r = 0.63$, $p < 0.001$) although not as powerful as the subNiño4 index. The magnitude and location of the correlation is robust across three different subsurface ocean temperature products we have examined (Supplementary Fig. 1).

Subsurface ocean temperature and landfall factors. The difference of May–November tropical Pacific SST anomaly for those years following cool and warm subNiño4 anomalies is shown in Fig. 3a. Prior summer negative subNiño4 anomalies lead to a central La Niña type SST anomaly pattern. A cool region extends from the equatorial East Pacific into the Central Pacific with warmer SSTs to the East of the Philippines. This pattern has very high spatial correlation ($r = 0.95$) to the May-Nov Niño 3.4 SST anomaly field over the mapped region (Supplementary Fig. 2).

We now examine how the prior summer subNiño4 index relates to key variables in the landfall track region during the TC season. The only regionally consistent signal over the entire track region is found in the winds at the steering level (approximately centred at 500 hPa). Wind anomalies over both the South China Sea and Philippine Sea enhance the northwest steering towards landfall on the South China coast (Fig. 3b). The zonal wind anomaly for low-high subNiño4 JJA-1 and May–November Niño 3.4 (Supplementary Fig. 2) also have very high spatial correlation ($r = 0.95$) over the mapped region for the 500 hPa zonal wind

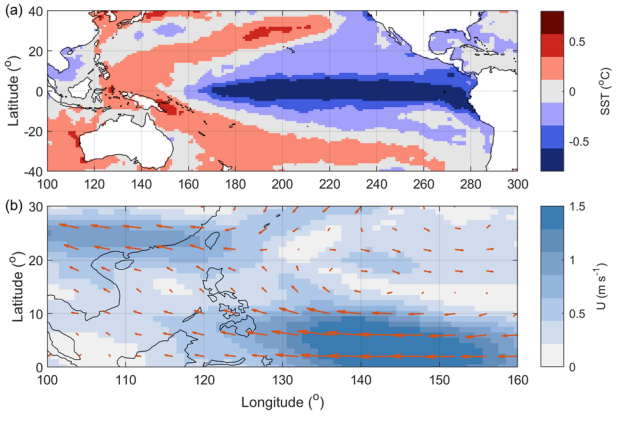


Fig. 3 Tropical Pacific anomalies associated with increased South China landfall. May–November anomalies of **a** sea surface temperature, and **b** wind at 500 hPa. Anomalies calculated as difference between mean lower and upper subNiño4 JJA-1 tercile years. Lower and upper tercile thresholds are -0.15 and 0.30 °C, respectively. Magnitude of wind vector anomaly is shaded, arrows show direction.

anomaly. Prior summer subNiño4 index anomalies lead to diverse regional changes in SST, vertical wind shear, low-level relative vorticity and mid-level relative humidity (Fig. 4). These other factors are variable and even oppose each other in the two seas. There is TC favourable enhanced mid-level relative humidity and low-level vorticity, but unfavourable lower SST and enhanced vertical wind shear over the the South China Sea. In contrast, the Philippine Sea is characterised by favourable warmer SST and reduced vertical wind shear in the South, but unfavourable reduced low-level vorticity and some reduced mid-level relative humidity in the West. Only the 500 hPa geopotential height and zonal wind anomalies are consistent across the entire region with the enhanced westward steering to South China.

We find that negative subNiño4 anomalies lead to a small shift in genesis location of landfalling storms toward the South China coast (Supplementary Fig. 3) although this difference is not statistically significant ($p > 0.1$). This lack of significant genesis shift is different to known basin wide genesis shifts under ENSO conditions[6,18]. The relationship of landfall with the basin storm total itself is weak ($r = 0.28$, $p = 0.09$), suggesting why previous studies on basin wide activity may give little insight into South China landfall statistics.

Predictive model performance. We now test the performance of the subNiño4 index as a predictor of South China Landfall count in a regression model validated using a leave-one-out cross-validation. Figure 5 shows time series of observations and cross-validated linear model predictions and an 11-year rolling correlation between the two showing decadal variation of r. The correlation coefficient after cross-validation is $r = 0.61$ ($p < 0.001$). The linear model outperforms the Poisson regression model in terms of correlation, root mean square error (RMSE) and skill. This is likely because the relationship between the subNiño4 index and landfall count is more linear than log-linear (Supplementary Fig. 4).

We also perform a double cross-validation as a further test of the robustness of subNiño4 as a predictor of South China landfall. This technique is designed to remove biases due to predictor selection, as well as model fitting, being performed using data also used in validation. The double cross-validation process reveals that the optimum choice for a predictor box (of the same size) is actually 8° East and 4° North of the Niño 4 box, but still at the original subNiño4 100–300 m depth range and JJA-1 time window. The double cross-validated correlation is 0.65 ($p < 0.001$). This is higher than the simple single cross-validation correlation because the subNiño4 box was not chosen using optimisation methods.

We find that after double cross-validation the model predictions have a skill of 23% against a climatology baseline. This is similar to a previously reported 37% skill in forecasting South China landfall issued in the April just before the TC season[13]. However, that method was not subjected to double cross-validation and is likely over-fitted with a lower true skill. Our forecast has only a slightly lower skill, but it is much simpler, likely more robust and crucially can be issued a year in advance. This is a step change.

Discussion

Our prior summer subNiño4 index is correlated with the following year's TC season (May–November) mean Niño 3.4 index ($r = 0.53$, $p < 0.001$). The mean TC season Niño 3.4 SST index has the strongest correlation of all traditional ENSO indices with the seasonal South China landfall count ($r = -0.46$, $p = 0.004$). The Modoki index has no significant correlation ($r = -0.16$, $p = 0.33$). We therefore propose that the subNiño4 index has predictive

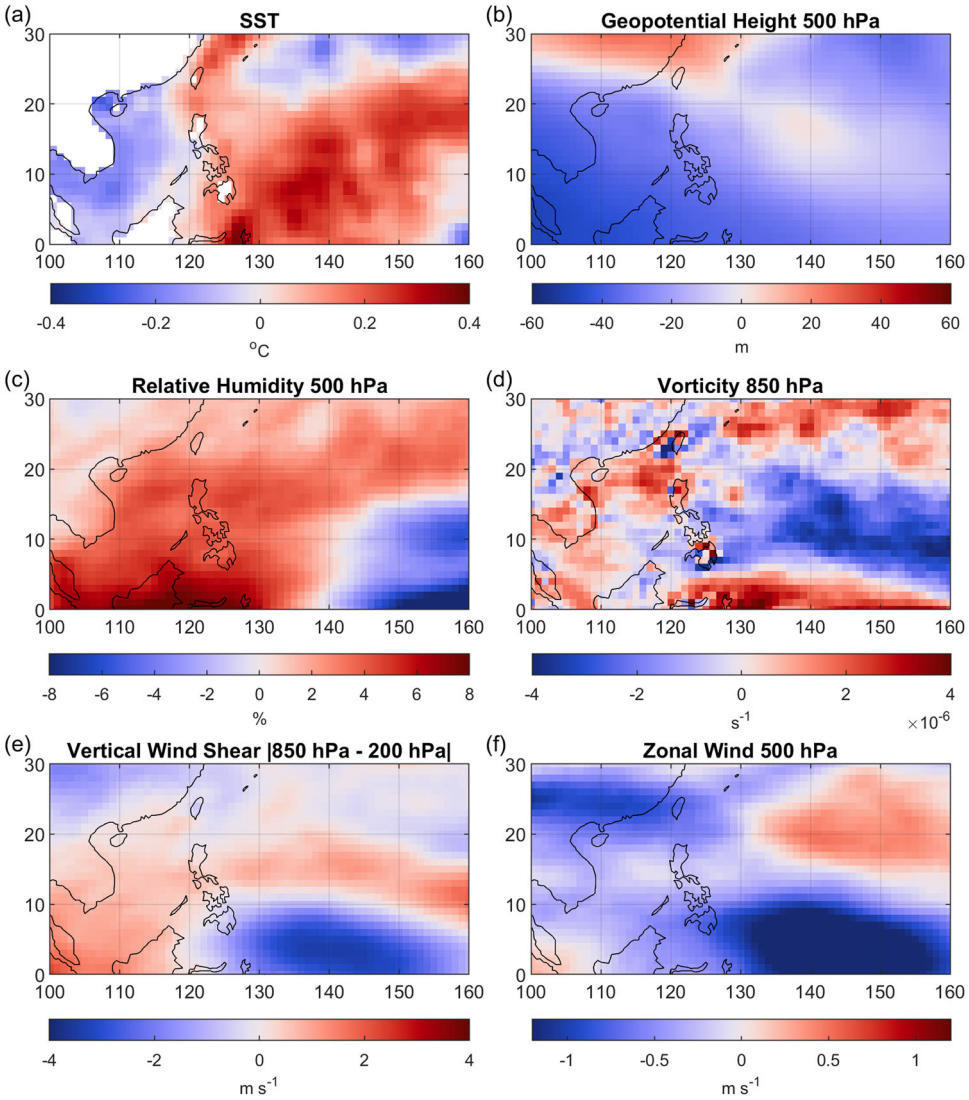


Fig. 4 Regional anomalies associated with low subNiño4 JJA-1 years. May–November anomalies of **a** SST, **b** geopotential height at 500 hPa, **c** relative humidity at 500 hPa, **d** relative vorticity at 850 hPa, **e** magnitude of vertical wind shear between 850 hPa and 200 hPa, and **f** zonal wind at 500 hPa. Anomalies calculated as difference between mean lower and upper subNiño4 JJA-1 tercile years. Lower and upper tercile thresholds are −0.15 and 0.30 °C, respectively.

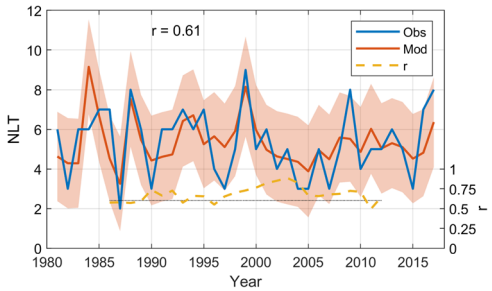


Fig. 5 Time series of number of tropical cyclones making landfall in South China (NLT). Observed value (blue) and cross-validated linear model prediction made using subNiño4 with 90% confidence bounds (orange). An 11-year rolling correlation coefficient (*r*) the two time series is shown in dashed yellow. The thin dashed line shows the *r* at which the 11-year rolling correlation is significant at the 95% confidence level using a two-tailed *t*-test ($r = 0.60$).

power because it precedes surface ENSO conditions during the TC season, which in turn influence tropical cyclone steering towards landfall in South China. A more negative summer subNiño4 precedes a La Niña the next year and more landfall in South China.

Current theories of ENSO are intimately associated with subsurface ocean temperature. These originate with Bjerknes[19] who proposed a coupled ocean-atmosphere feedback mechanism, which enhances warm east Pacific SST anomalies via ocean processes, including zonal advection, local upwelling, and thermocline feedback. This led to the delayed oscillator model based on Rossby wave reflection at the ocean western boundary[20] and the recharge-discharge model[21,22] based on discharge of warm water through Sverdrup transport. Observations of subsurface temperature have been shown to lead surface ENSO signals by up to 15 months[23,24]—the lead time scale we have identified here for TC forecasting. Statistical prediction models of ENSO based on these ideas helped break through the "spring persistence barrier"[25,26].

What are the physical mechanisms by which the subNiño4 index becomes a predictor of landfall? The depth of the thermocline and its spatial integral, the warm water volume, are well known as important factors for the evolution of ENSO signals and have previously been used in predictive models[27]. Here, we are particularly concerned with the relatively shorter ENSO growth phase rather than explaining the longer term reversal. The recharge oscillator theory predicts a time of about 12 months from the West Pacific subsurface anomaly, the subNiño4 region, to reach the surface by advection and upwelling[22]. Coupled models also confirm the eastward advection and climatalogical tradewinds driving equatorial upwelling of the subsurface temperature anomalies from the subNiño4 region. Specifically they show that the summer subsurface temperature anomaly leads the Niño 3.4 surface region by about 12 months[28]. Once the anomaly has reached the surface, the SST patterns influence atmospheric circulation patterns through well known zonal shits in atmospheric diabatic heating. We have shown that La Niña type SST patterns cause more northwestward steering wind conditions, which favour landfall of TCs in South China. This steering wind anomaly is consistent with the idealised northwestward Matsuno-Gill type circulation response to enhanced equatorial atmospheric diabatic heating towards the west, which is typical of La Niña[29].

Surprisingly, there is significantly more correlation ($r = 0.67$, $p < 0.001$) using the subNiño4 index as a predictor almost a year in advance than the Niño 3.4 (the most correlated index out of the standard El Niño indices and the El Niño Modoki) SST anomaly during the season itself ($r = 0.46$, $p = 0.004$). This finding is robust for all the different TC agencies. It may be explained by considering the atmospheric response when the subsurface temperature anomaly emerges at the surface. A cool (warm) SST anomaly causes less (more) cloud cover. These cloud anomalies drive surface shortwave radiative heating anomalies and this negative feedback is largest in the summer[28]. This radiative effect dampens the in-season SST-landfall correlation. The Niño 3.4 SST index in season is not an independent predictor: the very conditions, the diabatic heating anomalies, that favour enhanced TC steering also at the same time dampen the SST through cloudiness anomalies. However, the subNiño4 index is a leading indicator and independent of TC activity. We therefore propose that temperature anomalies at depth, before they are damped at the surface, are a stronger predictor.

The prior summer subNiño4 index also has some potential for predicting the basin wide seasonal ACE. ACE is measure of intensity, duration and number of TCs. The correlation of subNiño4 index with total number is found to be small. As subNiño4 can predict more/less steering to South China it is also therefore a measure of the TC poleward fraction. As more poleward TCs typically have larger duration and higher intensities they also enhance the seasonal ACE. Tracks typically are either more poleward (high ACE) or towards South China (low ACE) so that there is an anti-correlation between basin ACE and South China landfall ($r = -0.40$, $p = 0.01$). Therefore, there is also an opposite correlation of the basin ACE compared to TC landfall with the subNiño4 index.

Conclusion

We identify a subNiño4 index as subsurface ocean temperature anomalies averaged between depths of 100–300 m in the Pacific warm pool region. The subNiño4 index is strongly anti-correlated with the following year's South China tropical cyclone landfall count. Pacific subsurface temperature anomalies precede the surface ENSO conditions and cause atmospheric heating anomalies that modify the steering of TCs onto South China. Surprisingly, the previous summer subNiño4 is more correlated with South China landfall activity than the ENSO signal during the season itself. A prediction based on the observed prior summer subNiño4 index can be issued almost a year in advance and is as skilful as those currently issued in the Spring just before the TC season. Perhaps the forecasts of other atmospheric variables can also be substantially extended further than currently thought using regional subsurface ocean temperature in dynamical and statistical models.

Methods

The primary tropical cyclone landfall data were based on the China Meteorological Administration (CMA) best track data set[30] from the International Best Track Archive for Climate Stewardship (IBTrACS, v04r00) World Meteorological Organisation data[31,32]. Additional TC data from the Joint Typhoon Warning Centre (JTWC), Hong Kong Observatory (HKO), and the Japan Meteorological Agency (JMA) were all obtained via IBTrACS. A South China landfall event was defined at the time and location of the first recorded timestep when the cyclone position was within 200 km of the mainland China coast in the latitude range 18°N–23°N and longitude greater than 109°E. This definition of landfall has the benefit of including cyclones, which affect coastal regions but whose centres do not necessarily reach land. This is more precise than, for example, a common method of counting the number of cyclones entering a box encompassing a coastal region. The South China landfall seasonal count is the number of landfall events occurring in the period May–November, which accounts for 99% of South China landfall events. Only TCs with lifetime maximum intensity >13 m s^{-1}, the speed at which JTWC begin issuing western North Pacific tropical cyclone warnings, are included. A climatology of TCs making landfall in South China is provided in Fig. 6. Seasonal basin total accumulated cyclone energy (ACE) was calculated as the sum of the maximum sustained wind speed squared for each 3-h interval in a given year excluding periods a cyclone is marked as extra-tropical[33].

The subsurface ocean temperatures are from Met Office Hadley Centre EN4 objective analyses[34] with time varying bias adjustments[35]. The monthly means are on a 1° by 1° grid at a variety of depth levels. The ECMWF ocean reanalysis system[36] (ORAS5) and the Global Ocean Data Assimilation System[37] (GODAS) provided alternative subsurface ocean temperature data. Thermocline depth was estimated via linear interpolation depth to 20 °C. The monthly mean sea surface temperature was from the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST)[38] on a 1° by 1° grid. The geopotential height and wind speed reanalysis data are from European Centre for Medium-range Weather Forecasts

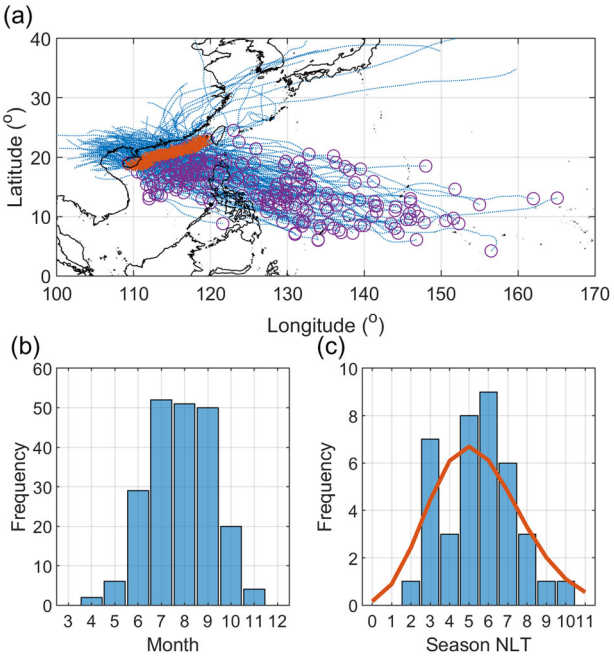


Fig. 6 Climatology of TCs making landfall in South China between 1981–2017. Tracks of typhoons making landfall in South China (blue lines) in the CMA best track database are shown in **a** with genesis locations (mauve circles) and landfall locations (orange crosses). A distribution of South China landfall event month is shown in **b**. The distribution of seasonal number of landfalling typhoons (NLT) in South China is shown in **c** with a Poisson distribution of the same mean (5.32) overlaid for comparison.

(ECMWF) reanalysis ERA5[39]. Monthly mean fields down-sampled to 1° by 1° resolution were used in this analysis.

Prediction models for regional annual count of TCs making landfall typically use Poisson regression[40–42] models with parameters estimated using a maximum likelihood technique. However, the assumption in a Poisson regression model that the logarithm of the expected value of the response variable varies with a linear combination of the predictors may not be optimal hence we also fit a linear regression model to our data. These have been used successfully in previous studies to model tropical cyclone landfall counts[13,43,44].

In addition to choosing the right model, it is important to assess its performance fairly through cross-validation. We use a double cross-validation method comprising an inner and outer layer. The outer layer is a standard leave-one-out cross-validation[43,45], whereby model fitting and validation are performed separately for every year in the data. For each year, data are split into two sets: the training set (all years except the target year), and the validation set (the target year). Model fitting is performed using only the training set and validation only on the validation set. This removes the bias due to validating the model on data which has used also been using in the fitting process and therefore improves assessment of out of sample model error. The inner layer attempts to deal with the problem of biased predictor selection by choosing the optimum predictor from an ensemble of predictors at each stage of the outer cross-validation. This process acknowledges the fact that decisions are made when choosing a predictor based on correlation analysis of the whole data set, whereas a fair analysis chooses a predictor based only on the training set.

The ensemble of predictors used in the inner layer of cross-validation was made by calculating the mean seasonal subsurface ocean temperature anomaly of a box in different positions. We sampled every position available to a box of size 50° longitude by 10° latitude (Niño 4 size) by 200 m depth and spanning 3 months within the region 146°E–224°E, −11°N–11°N (at 2° longitude, 1° latitude intervals), 50–350 m (in 50 m intervals) and the months May–September (in 1-month intervals). This creates an ensemble of 1755 predictors to choose from at each stage of the outer leave-one-out process. Extending the ranges beyond those described yields no alternative predictor choices so we believe this cross-validation to be complete.

We calculate skill scores against a climatological mean baseline using,

$$\mathrm{Skill} = (1 - \mathrm{RMSE}/\mathrm{RMSE}_{\mathrm{clim}}) \times 100\%, \quad (1)$$

where RMSE is the model prediction root mean squared error, and $\mathrm{RMSE}_{\mathrm{clim}}$ is the root mean squared error using the climatological mean as the baseline prediction.

Data availability

All data used in this study are publicly available. Sea surface temperature data is available at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html. TC track data are available through IBTrACS at https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access. ERA5 reanalysis data are provided at https://doi.org/10.24381/cds.f17050d7. The UK Met Office EN4 subsurface ocean temperature data is at https://www.metoffice.gov.uk/hadobs/en4/download-en4-2-1.html.

Code availability

All code used to perform the calculations can be provided by the author upon request.

Received: 28 April 2020; Accepted: 14 September 2020;
Published online: 15 October 2020

References

1. Camargo, S. J., Barnston, A. G., Klotzbach, P. J. & Landsea, C. W. Seasonal tropical cyclone forecasts. *WMO Bull.* **56**, 297–309 (2007).
2. World Bank Group. *East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth. Urban Development* (World Bank Publications, 2015). https://books.google.co.uk/books?id=pQ25BgAAQBAJ.
3. United Nations Publications. *Review of Maritime Transport 2016. Review Of Maritime Transport* (UN, 2017). https://books.google.co.uk/books?id=C0xRAQAACAAJ.
4. Klotzbach, P. et al. Seasonal tropical cyclone forecasting. *Trop. Cycl. Res. Rev.* **8**, 134–149 (2019).
5. Landsea, C. W. in *El Nino and the Southern Oscillation*, 149–182 (Cambridge University Press, 2010).
6. Chan, J. C. Tropical cyclone activity in the northwest Pacific in relation to the El Nino/Southern Oscillation phenomenon. *Mon. Weather Rev.* **113**, 599–606 (1985).
7. Lander, M. A. An exploratory analysis of the relationship between tropical storm formation in the western North Pacific and ENSO. *Mon. Weather Rev.* **122**, 636–651 (1994).
8. Wang, B. & Chan, J. C. How strong ENSO events affect tropical storm activity over the western North Pacific. *J. Clim.* **15**, 1643–1658 (2002).
9. Chan, J. C. L. Tropical cyclone activity over the western north pacific associated with El Niño and La Niña events. *J. Clim.* **13**, 2960–2972 (2000).
10. Huang, W. R. & Chan, J. C. L. Dynamical downscaling forecasts of Western North Pacific tropical cyclone genesis and landfall. *Clim. Dynam.* **42**, 2227–2237 (2014).
11. Lok, C. C. & Chan, J. C. Simulating seasonal tropical cyclone intensities at landfall along the South China coast. *Clim. Dynam.* **50**, 2661–2672 (2018).
12. Liu, K. S. & Chan, J. C. Climatological characteristics and seasonal forecasting of tropical cyclones making landfall along the South China Coast. *Mon. Weather Rev.* **131**, 1650–1662 (2003).
13. Goh, A. Z. C. & Chan, J. C. An improved statistical scheme for the prediction of tropical cyclones making landfall in South China. *Weather Forecast.* **25**, 587–593 (2010).
14. Schneider, E. K. et al. Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM. *Mon. Weather Rev.* **127**, 1187–1207 (1999).
15. Venugopal, T. et al. Statistical evidence for the role of southwestern Indian ocean heat content in the indian summer monsoon rainfall. *Sci. Rep.* **8**, 12092 (2018).
16. Scoccimarro, E. et al. Remote subsurface ocean temperature as a predictor of Atlantic hurricane activity. *Proc. Natl Acad. Sci. USA* **115**, 11460–11464 (2018).
17. Yu, X. & McPhaden, M. J. Seasonal variability in the equatorial Pacific. *J. Phys. Oceanogr.* **29**, 925–947 (1999).
18. Camargo, S. J., Emanuel, K. A. & Sobel, A. H. Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Clim.* **20**, 4819–4834 (2007).
19. Bjerknes, J. Atmospheric teleconnections from the equatorial pacific. *Month. Weather Rev.* **97**, 163–172 (1969).
20. Suarez, M. J. & Schopf, P. S. A delayed action oscillator for ENSO. *J. Atmos. Sci.* **45**, 3283–3287 (1988).
21. Jin, F.-F. An equatorial ocean recharge paradigm for ENSO. Part I: conceptual model. *J. Atmos. Sci.* **54**, 811–829 (1997).
22. Jin, F. F. & An, S. I. Thermocline and zonal advective feedbacks within the equatorial ocean recharge oscillator model for ENSO. *Geophys. Res. Lett.* **26**, 2989–2992 (1999).
23. Meinen, C. S. & McPhaden, M. J. Observations of warm water volume changes in the equatorial Pacific and their relationship to El Nino and La Nina. *J. Clim.* **13**, 3551–3559 (2000).
24. Lin, J. & Qian, T. Switch between El Nino and La Nina is caused by subsurface ocean waves likely driven by lunar tidal forcing. *Sci. Rep.* **9**, 13106 (2019).
25. Drosdowsky, W. Statistical prediction of ENSO (Nino 3) using sub-surface temperature data. *Geophys. Res. Lett.* **33**, L03710 (2006).
26. Clarke, A. J. & Van Gorder, S. Improving El Niño prediction using a space-time integration of Indo-Pacific winds and equatorial Pacific upper ocean heat content. *Geophys. Res. Lett.* **30**, 1399 (2003).
27. Dijkstra, H. A. The ENSO phenomenon: Theory and mechanisms. *Adv. Geosci.* **6**, 3–15 (2006).
28. Zhang, Q., Kumar, A., Xue, Y., Wang, W. & Jin, F. F. Analysis of the ENSO cycle in the NCEP coupled forecast model. *J. Clim.* **20**, 1265–1284 (2007).
29. Gill, A. E. Some simple solutions for heat-induced tropical circulation. *Q. J. R. Meteorol. Soc.* **106**, 447–462 (1980).
30. Ying, M. et al. An overview of the China meteorological administration tropical cyclone database. *J. Atmos. Ocean. Technol.* **31**, 287–301 (2014).
31. Knapp, K. R. et al. The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Am. Meteorol. Soc.* **91**, 363–376 (2010).
32. Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C. & Schreck, C. J. *International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4 (2018)*. (NOAA National Centers for Environmental Information, 2018), accessed 18 October 2019.
33. Bell, G. & Chelliah, M. The 1999 north atlantic and eastern north pacific hurricane season. "climate assessment for 1999". *Bull. Amer. Meteor. Soc.* **81**, S19–S22 (2000).
34. Good, S. A., Martin, M. J. & Rayner, N. A. EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Oceans* **118**, 6704–6716 (2013).
35. Gouretski, V. & Reseghetti, F. On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep-Sea Res. Part I: Oceanogr. Res. Pap.* **57**, 812–833 (2010).
36. Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K. & Mayer, M. The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: a description of the system and assessment. *Ocean Sci.* **15**, 779–808 (2019).
37. Behringer, D. W. & Xue, Y. Evaluation of the global ocean data assimilation system at ncep: the pacific ocean. In *Proc. Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*

(AMS 84th Annual Meeting, Washington State Convention and Trade Center, 2004).
38. Rayner, N. A. et al. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res. D: Atmos.* **108**, 4407 (2003).
39. Copernicus Climate Change Service (C3S). *ERA5: Fifth Generation of ECMWF atmospheric Reanalyses of the Global Climate*. (Copernicus Climate Change Service Climate Data Store (CDS), 2017), accessed 7 November 2019. https://cds.climate.copernicus.eu/cdsapp#!/home.
40. Tippett, M. K., Camargo, S. J. & Sobel, A. H. A Poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. *J. Clim.* **24**, 2335–2357 (2011).
41. Camargo, S. J., Robertson, A. W., Gaffney, S. J., Smyth, P. & Ghil, M. Cluster analysis of typhoon tracks. Part II: large-scale circulation and ENSO. *J. Clim.* **20**, 3654–3676 (2007).
42. Villarini, G., Vecchi, G. A. & Smith, J. A. Modeling the dependence of tropical storm counts in the north atlantic basin on climate indices. *Month. Weather Rev.* **138**, 2681–2705 (2010).
43. Li, X., Yang, S., Wang, H., Jia, X. & Kumar, A. A dynamical-statistical forecast model for the annual frequency of western Pacific tropical cyclones based on the NCEP Climate Forecast System version 2. *J. Geophys. Res. Atmos.* **118**, 12061–12074 (2013).
44. Tian, B. & Fan, K. Seasonal climate prediction models for the number of landfalling tropical cyclones in China. *J. Meteorol. Res.* **33**, 837–850 (2019).
45. Fan, K. & Wang, H. A new approach to forecasting typhoon frequency over the western north pacific. *Weather Forecast.* **24**, 974–986 (2009).

Acknowledgements

This work was supported by the UK-China Research and Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund. Ralf Toumi was also supported by the Horizon 2020 EU Research and Innovation programme.

Author contributions

N.S. and R.T. wrote and reviewed the manuscript text. N.S. prepared figures.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s43247-020-00033-2.

Correspondence and requests for materials should be addressed to N.S.

Peer review information Primary handling editor: Heike Langenberg.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020