Journal of Marine Science and Technology

Volume 31 | Issue 2 Article 7

Projection of Sea Surface Temperature Based on Empirical Mode Decomposition and Its Application in Typhoon Number Forecast

Wen-Cheng Huang
Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan, ROC,
b0137@mail.ntou.edu.tw

Jung-Chen Lee
Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan, ROC

Recommended Citation

Huang, Wen-Cheng and Lee, Jung-Chen (2023) "Projection of Sea Surface Temperature Based on Empirical Mode Decomposition and Its Application in Typhoon Number Forecast," *Journal of Marine Science and Technology*: Vol. 31: Iss. 2, Article 7.
DOI: 10.51400/2709-6998.2693
Available at: https://jmstt.ntou.edu.tw/journal/vol31/iss2/7

RESEARCH ARTICLE

Projection of Sea Surface Temperature Based on Empirical Mode Decomposition and Its Application in Typhoon Number Forecast

Wen-Cheng Huang*, Jung-Chen Lee

Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan, ROC

Abstract

The primary objective of this paper is to demonstrate the applicability of empirical mode decomposition (EMD) in projecting non-stationary daily sea surface temperature (SST) series, and the other is to apply the projected series of SST to estimate the number of typhoons. The study focuses on an area ranging from 5°N~25°N and 110°E~170°E, collecting daily sea temperature data and typhoons records within this region. The findings conclusively indicate that the EMD-based data generation method proposed in this paper can effectively addresses the limitations of previous hydrological models in generating non-stationary series, signifying a significant advancement in hydrological time series modeling for data generation. Furthermore, the study demonstrates the feasibility of using a combination of EMD synthesis and statistical regression analysis to assess typhoon frequency.

Keywords: Empirical mode decomposition, Sea surface temperature, Number of typhoons, Non-stationary series, Data generation

1. Statement of the problem

In system analysis, an extensive dataset is often employed for system simulation to assess the credibility of the system [1—3]. For instance, in reservoir operations [4—6], formulating reservoir operating rules based on the incoming water sequence is essential. However, data collection is often constrained by observation periods. Therefore, hydrologists initially employ time-series modeling to reproduce data, expanding the range of potential hydrological sequences to ensure system reliability. It is crucial to note data reproduction assumes the historical data belongs to a stationary sequence, where statistical characteristics remain constant over time. For instance, the widely used Box—Jenkins model leverages time series correlation for data generation or forecasting [7]. In the case of seasonal hydrological phenomena, statistical characteristics differ across seasons, constituting a non-stationary sequence. Nevertheless, data can be transformed into a stationary state through data standardization techniques, as demonstrated by the Thomas-Fiering model [1,7,8].

However, as a result of environmental changes, numerous hydrological variables such as rainfall, flow, temperature, evaporation, etc., are undergoing gradual transformations characterized by increasing or decreasing trends. While the ARIMA model [6] can smooth and forecast this non-stationary time series data, it is incapable of generating data. Therefore, when data generation is necessary in research, system analysis must first address the challenge of reproducing sequences with non-stationarity.

The gradual increase in sea surface temperature (SST) due to the effects of climate change is already highly significant [8—12]. If SST data is intended for related research but subject to changes in statistical characteristics, employing the Box—Jenkins model

Received 7 February 2023; revised 10 April 2023; accepted 30 May 2023.
Available online 30 June 2023

* Corresponding author.
E-mail address: b0137@mail.ntou.edu.tw (W.-C. Huang).

https://doi.org/10.51400/2709-6998.2693

for data generation becomes feasible. Additionally, when examining the relationship between seawater temperature and typhoon frequency, this study encounters non-stationary seawater temperature series. Hence, this article aims to propose a novel method for synthesizing non-stationary data and applying it to the estimation of typhoon numbers. The proposed approach in this paper involves utilizing Empirical Mode Decomposition (EMD) as a non-stationary sequence generation method to achieve this research objective.

2. Typhoon affecting Taiwan

Typhoons have a significant impact on human survival! Taiwan experiences an average annual rainfall of 2515 mm, with approximately 52% occurring during the typhoon season from July to October. Being situated along the western Pacific typhoon path, Taiwan typically encounters an average of 4 typhoons each year, which also brings substantial rainfall that affects the region [13].

For instance, a severe drought struck northern Taiwan from March to June 2002, resulting in a serious water shortage in the Taipei metropolitan area. The entire irrigated area in Taoyuan, covering 35,000 ha, was left fallow, and the Hsinchu Science Park (Taiwan's Silicon Valley) faced a dire shortage of industrial water. By the end of June, both the Shimen and Feitsui Reservoirs had completely depleted their water reserves. Thankfully, Typhoon Rammasun arrived on July 2, delivering abundant rainfall. Over the course of two days, the catchment area received 443 mm of rainfall for two days, rapidly alleviating the drought conditions.

In contrast, Typhoon Morakot, which struck on August 8, 2009, unleashed over 2000 mm of rainfall to southern Taiwan within a span of 72 h, resulting in 681 fatalities, 18 missing individuals, and agricultural losses amounting to 20 billion Taiwan dollars. The catastrophic event also led to the accumulation of approximately 100 million cubic meters of silt in Zengwen Reservoir and Nanhua Reservoir, exacerbating the severity of the overall disaster. Consequently, typhoon research has always been a critical focus in Taiwan. Typically, typhoon research encompasses four areas: path research, cause research, intensity research [14,15], and investigations into the relationship between the climate and typhoons [16].

Responding to public and media interest, predictions of typhoon numbers have become increasingly prevalent. Currently, statistical methods are commonly employed for seasonal forecasts of tropical cyclones in the Western North Pacific [17,18]. This study will utilize statistical methods for forecasting as well. However, the focus of this study is on annual rather than seasonal forecasting of typhoons numbers. The analysis process for this study involves: (1) gathering daily sea temperature and typhoon data; (2) employing regression analysis to examine the relationship between selected factors and the number of typhoons; (3) utilizing the EMD-based data generation method to project the subsequent year's daily seawater temperature series; (4) forecasting the number of typhoons expected to occur in the western Pacific the following year based on the projected SST sequences through regression analysis.

3. Data collection

3.1. Number of typhoons

This study gathered typhoon data from Taiwan's Central Weather Bureau typhoon database (https://rdc28.cwb.gov.tw/), encompassing the period from 1958 to 2019, which yielded a total of 1617 typhoons in the Western Pacific. Fig. 1 depicts the geographical distribution of these typhoons. Notably, 1444 typhoons occurred within the designated area of (5°N~25°N and 110°E~170°E), constituting approximately 89%. In Taiwan specifically, 98% of typhoons that prompt land warnings are concentrated within this region (5°N~25°N and 110°E~170°E), thereby serving as the focus of this study regarding typhoon impact on Taiwan.

3.2. Daily sea surface temperature (SST)

The sea temperature data utilized in this study were sourced from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) (https://icoads.noaa.gov/) provided by the National Oceanic and Atmospheric Administration (NOAA). This database encompasses ocean surface data spanning from 1800 to the present, with a resolution of $1^\circ \times 1^\circ$ since 1960. For this research, approximately 3.8 million daily sea temperature data points were collected from 1969 to 2018. The average daily sea temperature value across all points with the designated area of (5°N~25°N and 110°E~170°E) was adopted as the representative daily sea temperature value.

4. Preliminary study

4.1. Changes in the number of typhoons

Based on the analysis of typhoon occurrences from 1958 to 2019 (Fig. 2), employing the Mann−Kendall

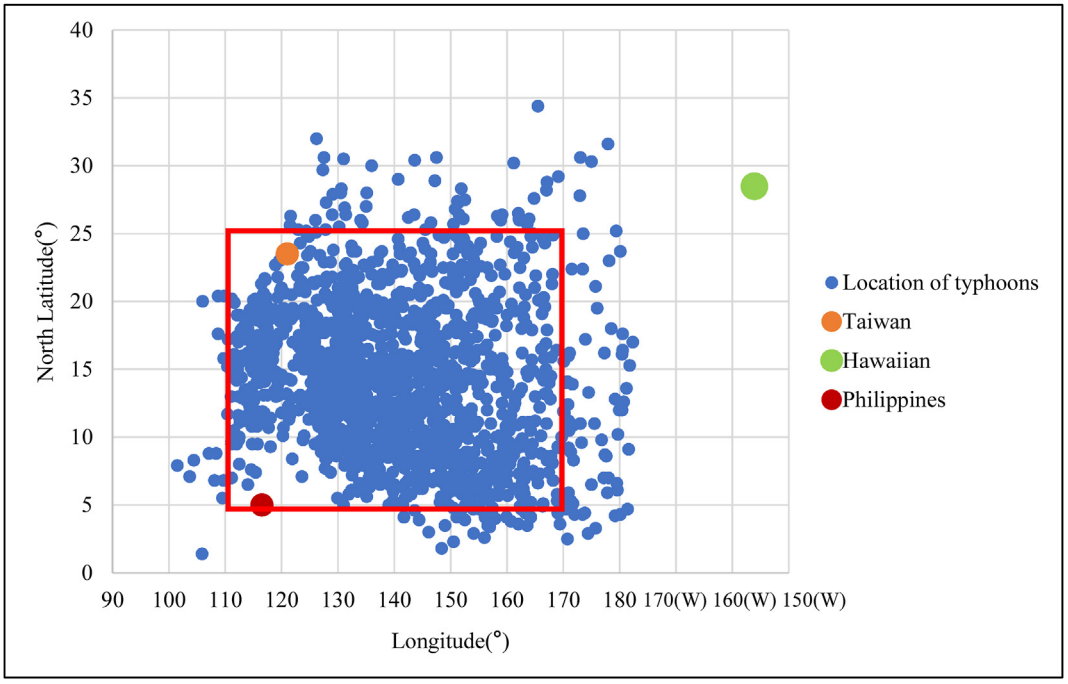


Fig. 1. Locations of typhoons in the Western Pacific from 1958 to 2019.

trend test or autocorrelation function test at a significance level of 5%, it is evident that the number of typhoons has exhibited no discernible increase or decrease over the past 62 years. This indicates that the frequency of typhoons represents an independent sequence. Furthermore, examining the monthly distribution of typhoons (Fig. 3a), it is observed that typhoons predominantly transpire from July to October, constituting approximately 69% of the annual count. The Mann–Kendall test confirms the absence of any notable upward or downward trend in the monthly frequency of typhoons.

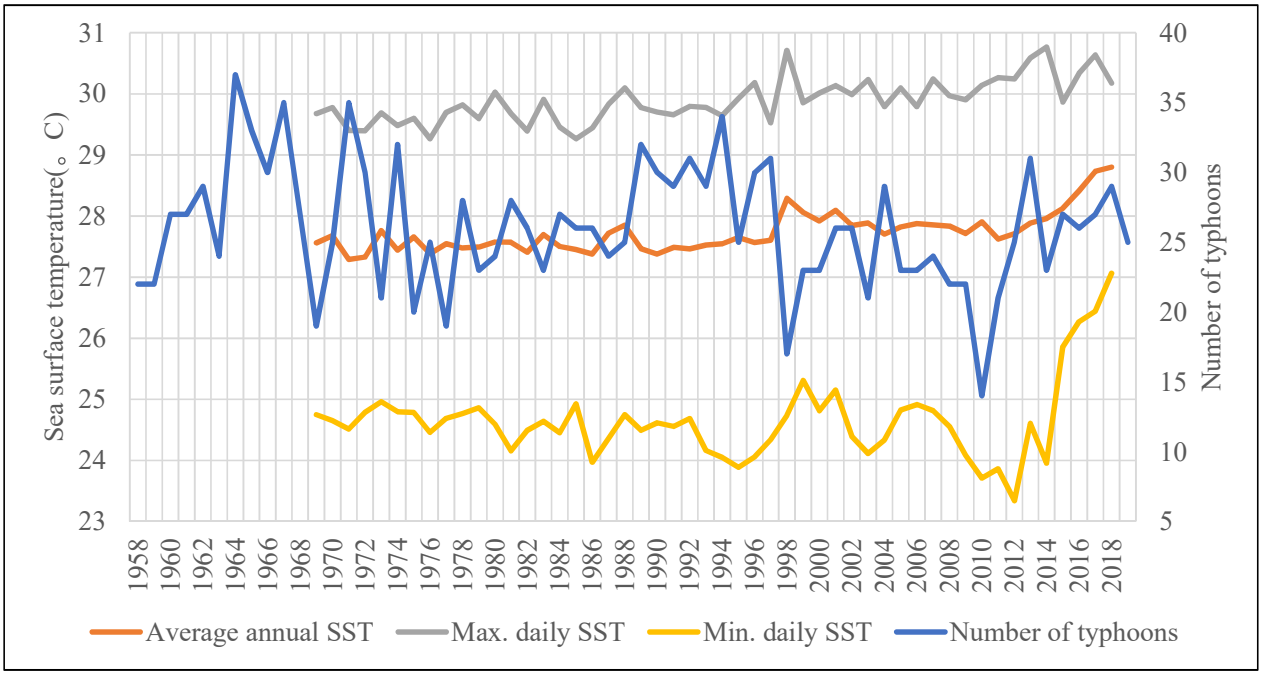


Fig. 2. Changes in the number of typhoons and sea surface temperature over the years.

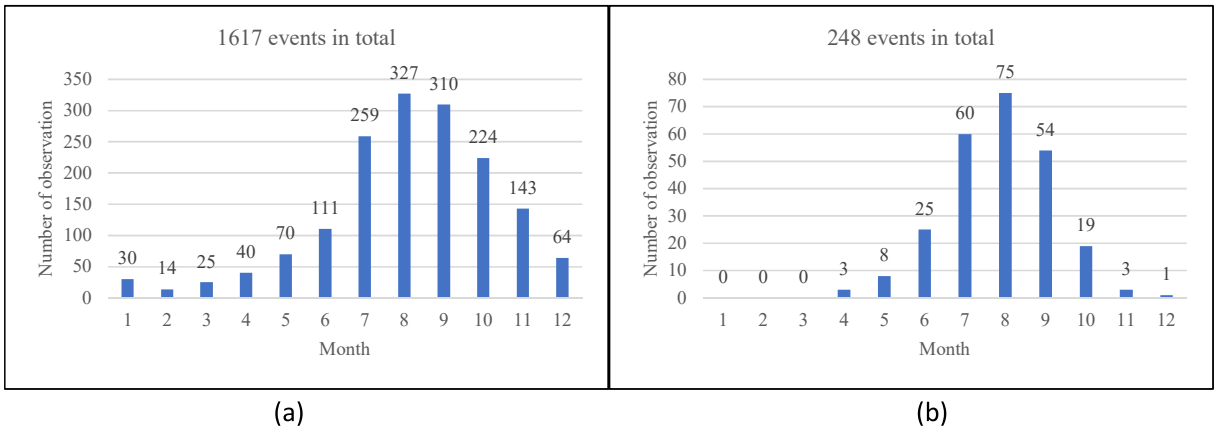


Fig. 3. (a) Monthly distribution of typhoons in the Western Pacific from 1958 to 2019; and (b) monthly distribution of land typhoon warnings issued by Taiwan from 1958 to 2019.

Assuming that the occurrence of typhoons follows a Poisson process, where typhoon events can transpire randomly at any time, the average number of typhoon occurrences per year is $\nu = 26.08$ times/year. Fig. 4 illustrates the probability mass function (PMF) for the number of typhoons in a year. Notably, in the Western Pacific, the probability of experiencing 26 or more typhoons in a year amounts to 53%. Additionally, the average number of typhoons in the months from July to October are 4.2, 5.3, 5.0, and 3.6, respectively, with probabilities of having 4 or more typhoons per month equaling 0.60, 0.75, 0.72, and 0.49, respectively.

Contrasting with the Western Pacific region, Taiwan issued a total of 248 land typhoon warnings from 1958 to 2019, with 98% of them occurring between 5°N~25°N and 110°E~170°E. As depicted in Fig. 3b, there were no typhoons affecting Taiwan from January to March, while the typhoon season (July to October) witnessed 208 typhoons, accounting for 84% of the overall count. On average, Taiwan experiences $\nu = 4$ typhoons per year, with 1.0, 1.2, 0.9 and 0.3 on average from July to October, respectively. According to the Poisson process, the probability of experiencing 4 or more typhoons in Taiwan within a year reaches 57%, while the probabilities of typhoons from July to October are 0.62, 0.70, 0.58, and 0.26, respectively.

4.2. Changes in sea surface temperature

The average annual sea surface temperature changes from 1969 to 2018 demonstrates a significant increase in the western Pacific (Fig. 2), as confirmed by the Mann–Kendall trend test, even at a 1% significance level. Notably, substantial rise in sea temperature occurred after 2014.

Fig. 5 showcases the distribution of interquartile ranges for sea temperature in each month. Sea temperatures from January to March predominantly remain below 26.5 °C, potentially indicating the threshold for tropical cyclone formation. Correspondingly, as observed in Fig. 3a, the occurrence of typhoons during this period is limited, comprising only 5% of the total. Conversely, sea temperatures in July are the warmest of the year, with lower quartile, median, and upper quartile values of 29.13 °C, 29.22 °C and 29.52 °C, respectively. The highest and lowest sea temperatures are 30.02 °C and 28.57 °C, respectively.

Fig. 6 illustrates the monthly average temperature changes. It is evident that the sea surface temperature exhibits a prominent upward trend on a yearly basis, particularly during the low temperature period. For instance, from 2014 to 2018, there was a temperature increase of more than 1.7 °C in January, February and March.

Furthermore, examining the daily SST variation, the minimum daily variation of SST from 1969 to 2014 was not significantly notable, with a range of 24–25 °C (refer to Fig. 2). These temperatures were primarily observed between December and March of the following year, with approximately 59% occurring in February. However, after 2014, a distinct rapid temperature increase became apparent, specifically in February. The recorded high-temperature surged from 25.86 °C in 2015 to 27.06 °C in 2018. The underlying reasons for this phenomenon warrant further investigation by researchers. Conversely, the annual maximum daily SST variation exhibits a significant increasing trend (see Fig. 2), occurring between June and October, with July accounting for 38% and August accounting for 26% of the occurrences.

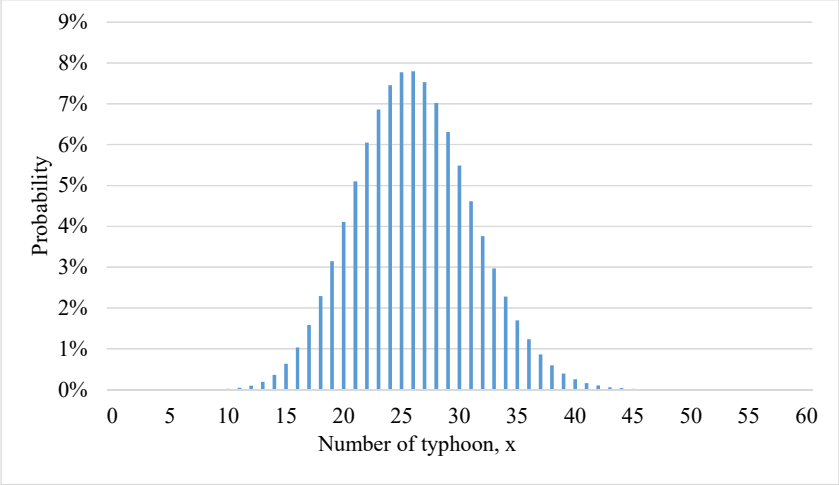


Fig. 4. Probability of the number of typhoons in a year based on the Poisson process.

Based on the data, from 1969 to 2018, there were 1131 typhoons that occurred in the [5°N~25°N, 110°E~170°E] region when the initial daily SST was between 25 °C and 31.5 °C. The quartiles for the initial SSTs are 28.33 °C (lower), 28.91 °C (median) and 29.30 °C (upper). This means 75% of typhoons have an initial sea surface temperature higher than 28.33 °C.

4.3. Statistical methods for estimating the number of annual typhoons

Various factors can influence the formation of typhoons, such as sea temperature, convective instability, relative humidity, vertical wind shear, Coriolis force, and relative vorticity [19–23]. Since seawater temperature plays a significant role in typhoon formation, and consecutive high-temperature days can facilitate typhoon development, this study considers daily sea surface temperature and the number of consecutive days exceeding a threshold temperature as key factors. Regression analysis using statistical methods will be employed to analyze the relationship between these factors and the number of typhoons. The variables selected for analysis are defined as follows:

$$X_1 = \frac{D_{max}}{D} \quad (1)$$

$$Y = \frac{N_a}{N_T} \quad (2)$$

D_{max}: The maximum number of consecutive days greater than the cutting temperature in a year;
D: Days of a year (D = 365);
N_a: The number of typhoons that occurred during the maximum number of consecutive days in a year;
N_T: The total number of typhoons that occurred in [5°N~25°N, 110°E~170°E] in one year.

Based on the operational theory [7], the quartiles of the initial temperature (28.33 °C, 28.91 °C, 29.30 °C) that give rise to typhoons can be used as threshold levels. By calculating the values of X_1 and Y above the threshold temperature for each year, a correlation can be established between the two variables. This article focuses on simple linear regression to investigate the relationship $Y = f(X_1)$.

However, for the $Y = f(X_1)$ relationship, the dependent variable Y (i.e. N_a/N_T) is the percentage of typhoon occurrences, not the total number of typhoon occurrences (N_T) in the year. Therefore, it becomes necessary to further explore the relationship between N_a and Y, denoted as $N_a = f(Y)$. In practical application, Y can be estimated from the independent variable X_1 first, then N_a can be determined from Y, and finally N_T can be obtained.

4.4. Regression-based results

Initially, a simple linear regression is performed using the typhoon records from 1969 to 2018. The resulting model serves as calibration, which is then used to project the potential number of typhoons based on the EMD reproduction sequence. The historical records are used to verify the projected number of typhoons.

Regression analysis is conducted the daily SST for a given year, the specified threshold value, and the corresponding number of typhoons that occurred in that year. Based on the data from 1969 to 2018, Fig. 7

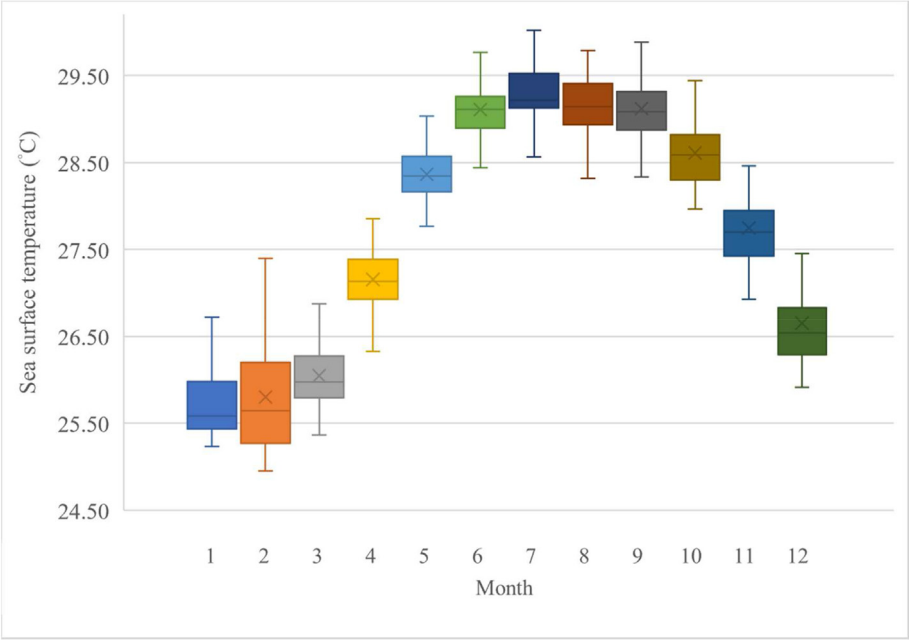


Fig. 5. Box plots of monthly sea surface temperature in the study area.

illustrates the linear relationship between X_1 and Y as well as Y and N_a for different daily sea temperature thresholds are 28.33 °C, 28.91 °C and 29.30 °C, respectively. The relevant regression analysis results are presented in Table 1.

It is evident that as the temperature threshold increases, the range of values for three variables X_1, Y, and N_a becomes smaller. The regression equations demonstrate that due to the constraint of $0 < Y \leqq 1$, the feasible range of X_1 and N_T values varies with

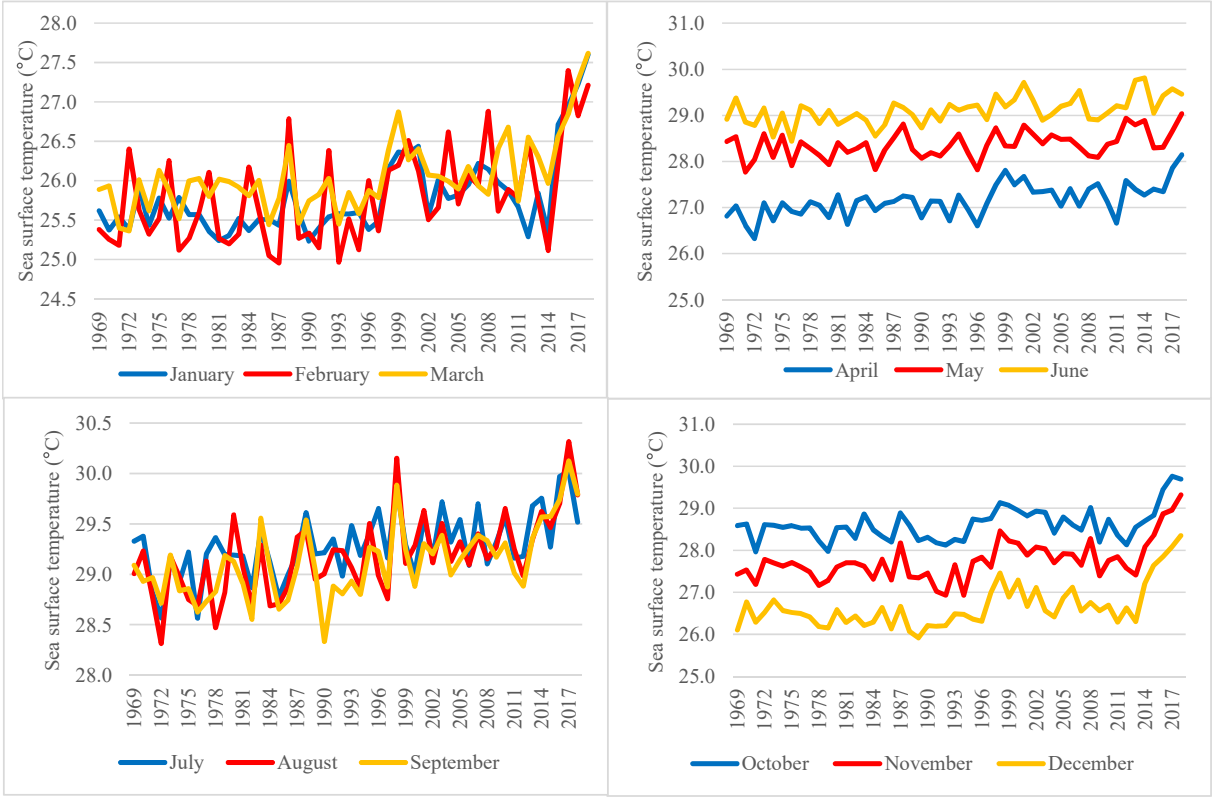


Fig. 6. Changes in the average monthly sea surface temperature of the study area from 1969 to 2018.

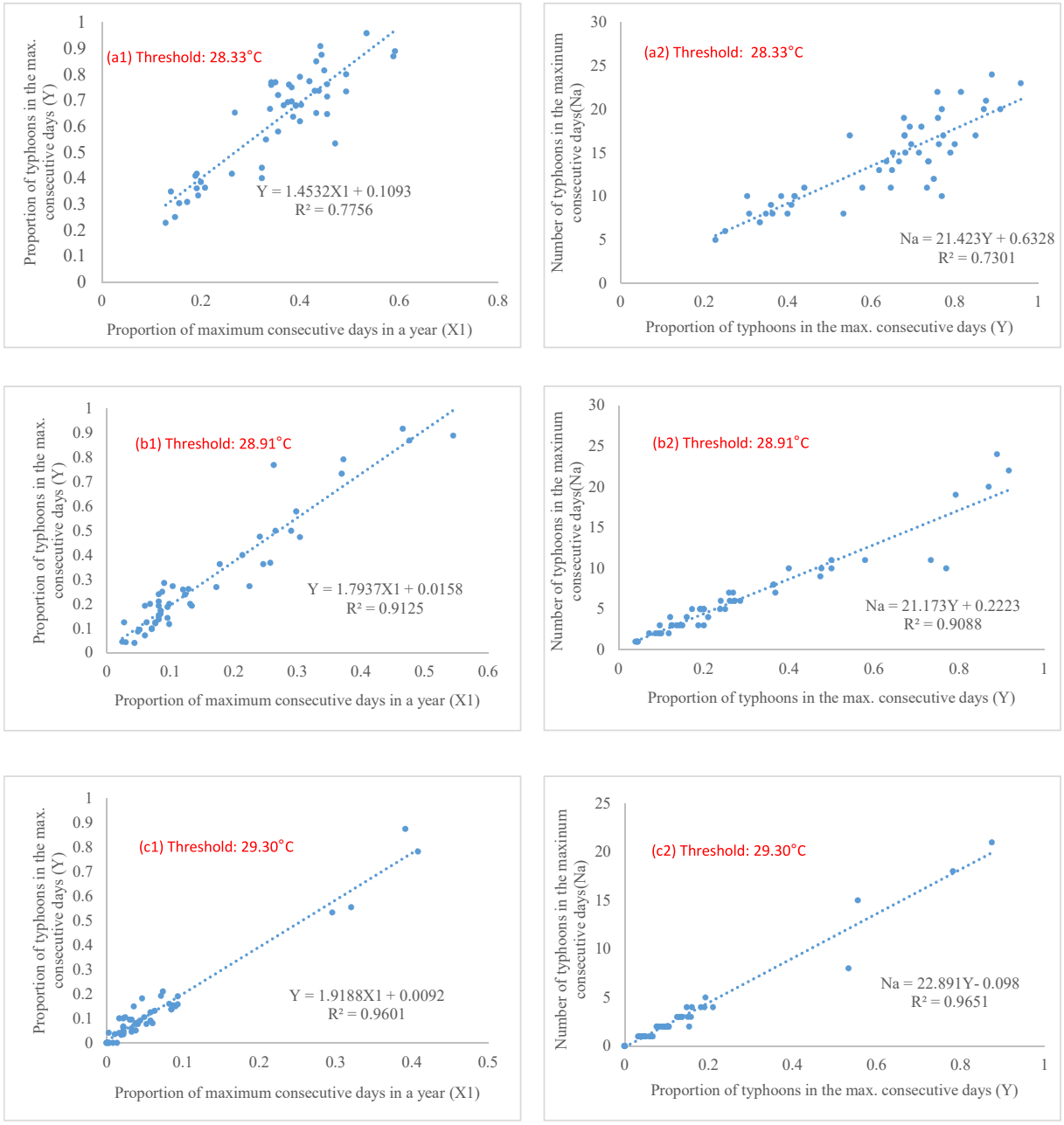


Fig. 7. Linear relationship between X_1 and Y as well as Y and N_a based on different daily sea temperature thresholds.

Table 1. Regression results under different sea temperature thresholds.

Threshold	$Y = f(X_1)$	R^2	$N_a = f(Y)$	R^2
28.33 °C	$Y = 1.4532X_1+0.1093$	0.7756	$N_a = 21.423Y+0.6328$	0.7301
28.91 °C	$Y = 1.7937X_1+0.0158$	0.9125	$N_a = 21.173Y+0.2223$	0.9088
29.30 °C	$Y = 1.9188X_1+0.0092$	0.9601	$N_a = 22.891Y-0.0980$	0.9651

X_1: percentage of the maximum number of consecutive days in a year (see Eq. (1)).
Y: percentage of typhoon occurrence in the maximum number of consecutive days (see Eq. (2)).
N_a: number of typhoons that occurred during the maximum number of consecutive days in a year.
R^2: coefficient of determination.

different sea temperature thresholds. For a threshold of 28.33 °C, it lies within $1/365 \leqq X_1 \leqq 223/365$ and $22 \leqq N_T \leqq 27$. With a threshold of 28.91 °C, the range is $1/365 \leqq X_1 \leqq 200/365$ and $21 \leqq N_T \leqq 32$. At 29.30 °C, the feasible range is $1/365 \leqq X_1 \leqq 188/365$ and $16 \leqq N_T \leqq 23$.

Notably, the estimated range of the number of typhoons using the 28.33 °C threshold falls within the estimated range of 28.91 °C. Conversely, the feasible range of typhoon occurrences at 29.30 °C differ significantly from that at 28.91 °C. The usable range for estimating typhoon frequency based on both thresholds is $16 \leqq N_T \leqq 32$, which aligns closely to the historical record range. Therefore, this study will utilize both the 28.91 °C and 29.30 °C thresholds to estimate the number of typhoons.

5. Briefing of empirical mode decomposition

Based on regression analysis discussed earlier, it is evident that if the annual daily sea temperature change is provided, the regression relationship can be employed to estimate the potential number of typhoons in that year. However, the daily sea temperature exhibits non-stationary behavior and cannot be generated using conventional hydrological models like ARIMA. This article focuses on the utilization of EMD to generate non-stationary SST sequences and predict the number of typhoons based on the regression relationship.

In order to gain a deeper understanding of the characteristics of daily sea temperature data, this study utilizes EMD to partition the historical daily sea temperature data into multiple intrinsic mode functions (IMFs) and a residue. The IMF must satisfy two conditions: (1) Throughout the entire dataset, the number of extremum points (local maxima and minima) must be equal or differ by at most one from the number of zero-crossing points; (2) At any given point in time, the mean value of the upper envelope defined by the local maxima and the lower envelope defined by the local minima must be zero [24,25]. In essence, the original sea temperature data SST(t) can be decomposed into the sum of p IMFs and a residual r(t). That is,

$$SST(t) = \sum_{i=1}^{p} IMF_i(t) + r(t) \forall t \tag{3}$$

Where the number p of IMFs is approximately equal to $(\log_2 n)$-1 [26]. The residue is a monotonic function or the residual is at most one extreme. This will show the changing trend of daily sea temperature.

Since the length of daily data over a period 50 years is 18,250, up to 13 IMFs and 1 residue can be decomposed. Fig. 8 illustrates IMF1, IMF7, IMF13 and the residue. The plot clearly demonstrates a gradual increase in the residual from 27.30 °C to 28.25 °C. This indicates a progressive upward trend in daily sea temperature, with an average annual increase of 0.019 °C. Table 2 provides information on the period and weight of each IMF. The period exhibits a consistent increase with each IMF level. Analyzing the weight variations reveals that IMFs with a period of less than one-year account for 96.8% of the total, implying that IMF1~IMF7 exert the most influence on the data reorganization. Notably, IMF7 has an average period of 365.5 days (equivalent to one-year), and a weight as high as 90.2%. Hence, IMF7 plays a dominant role in the data reconstruction process. Additionally, Fig. 8 illustrates that the cumulative value of IMFs and residue closely approximates the original sea surface temperature data. The discrepancy between the two generates an error signal characterized as white noise with a mean of zero and finite variance. The definition of the weight is as follows [26]:

$$E_i = \sum_{\forall t} |A_t|^2 \text{ for } i = 1, 2, \ldots, p \tag{4}$$

$$W_i = \frac{E_i}{\sum_{\forall i} E_i} \tag{5}$$

Where $|A_t|$ = vertical distance of the IMF value; E_i = total energy; W_i = weight of energy.

Historical records indicate a significant rise in sea surface temperatures since 2014. Fig. 8 further illustrates that the trough of IMF7 has continued to rise after 2014, with the peak and trough getting closer together. Given that IMF7 plays a leading role in the data reconstruction process (as shown in Table 2), the cumulative effect of IMFs contributes to higher SSTs post-2014.

Analyzing the amplitude change of IMF7 reveals that the peak and trough amplitudes were close to ±2 °C from 1969 to 1997. From 1998 to 2014, the distance between the peak and trough reduced but still remained close to ±2 °C. However, from 2015 to 2018, the peak and trough amplitudes decreased significantly, and fluctuated between ±1.5 °C. The pattern in IMF7's amplitude closely mirrors the change in annual sea temperatures from 1969 to 2018. Historical data indicates that the average annual sea temperature was 27.54 °C from 1969 to 1997, which rose to 27.89 °C from 1998 to 2014, and then experienced a rapid increase from 28.12 °C in 2015 to 28.79 °C in 2018. The reduced amplitude of IMF7 indicates a convergence between high temperature and low temperatures. If the contrast

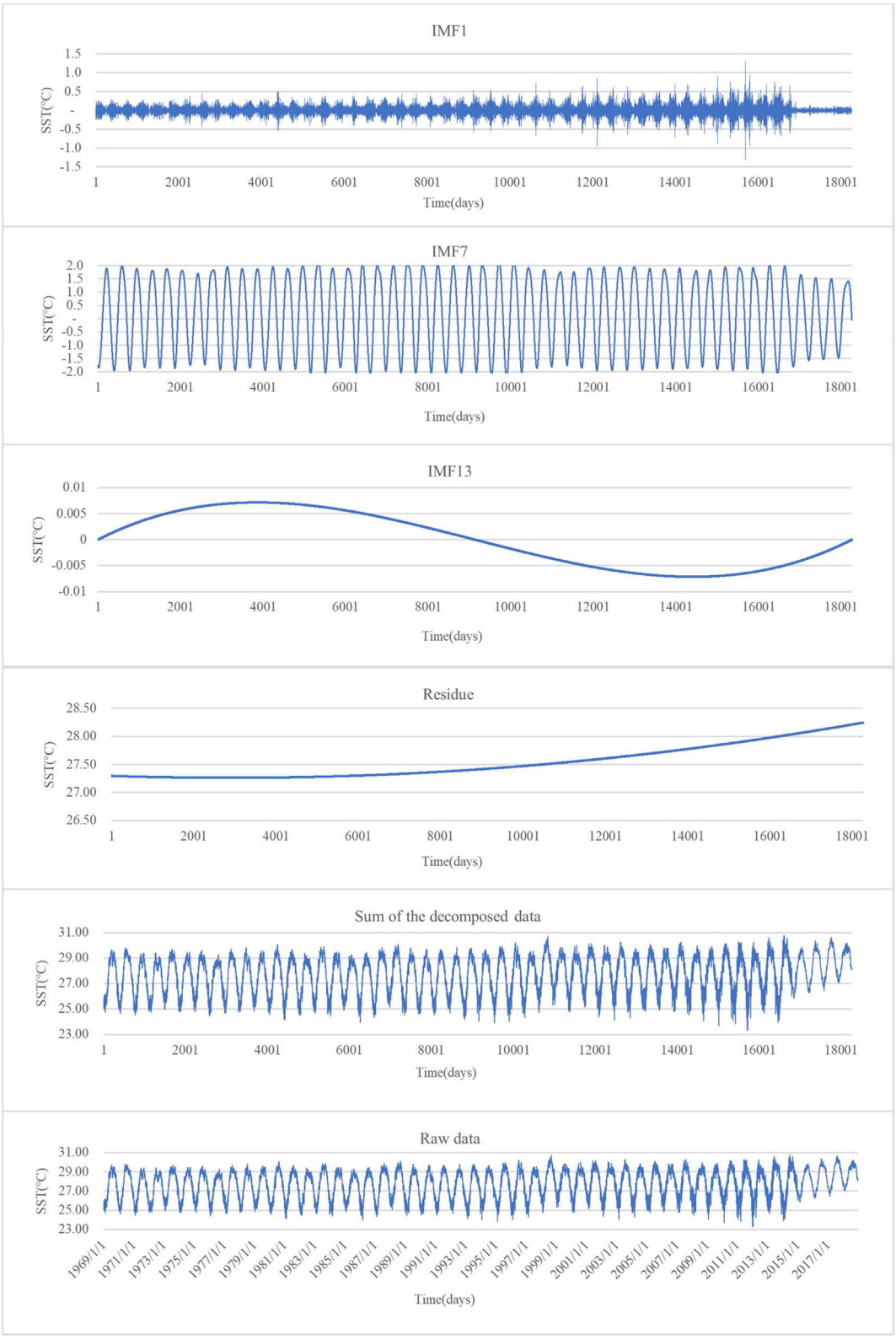


Fig. 8. Decomposition of 50-year daily SST by EMD.

between cold and hot becomes less distinct, it may have implications for the marine ecological environment and even human health. This trend must be taken seriously moving forward. Fig. 2 also illustrates this phenomenon, where the maximum daily SST after 2014 remained between 30 °C and

Table 2. Results of 50-year daily SST data by EMD.

IMF	Period (days)	Weight (%)
IMF1	2.90	1.12
IMF2	5.73	1.04
IMF3	11.15	0.91
IMF4	22.09	0.77
IMF5	43.76	0.80
IMF6	100.35	2.03
IMF7	365.52	90.15
IMF8	710.21	1.26
IMF9	1459.81	0.58
IMF10	3107.59	0.65
IMF11	7940.18	0.68
IMF12	12066.70	0.01
IMF13	N/A	0.00

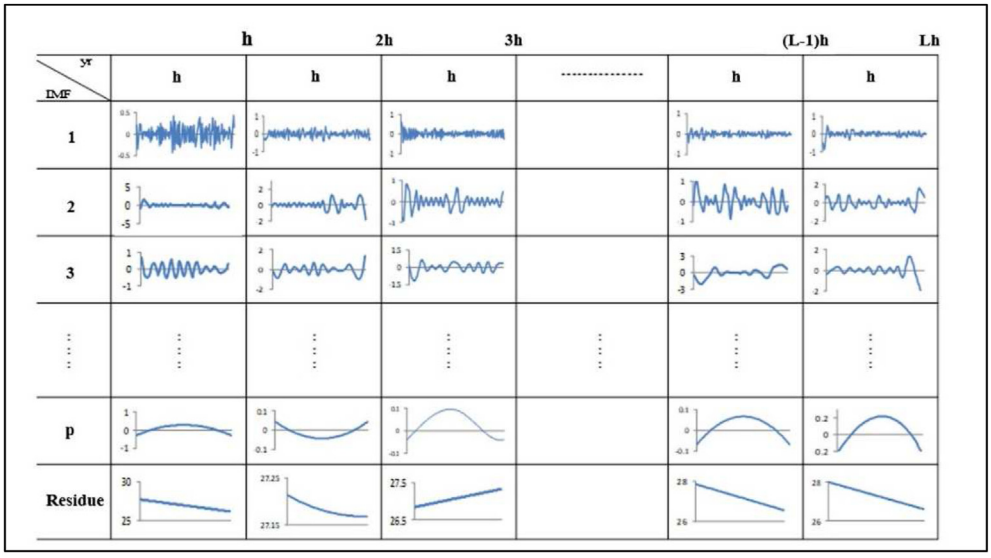
30.5 °C, while the minimum daily SST rapidly rose from 24 °C to 27 °C.

6. Concept of SST data synthesis

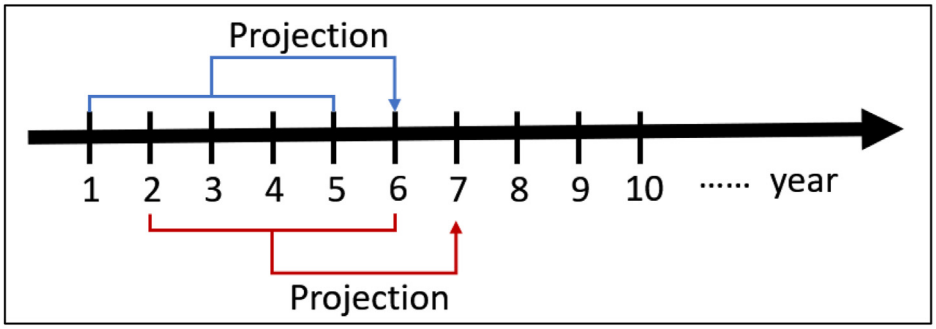
The main focus of this paper is to predict the number of typhoons based on the projected SST series. In a previous study, the authors proposed a generation method suitable for non-stationary sequences [26]. Given the evident non-stationary characteristics of sea temperature, this study employs EMD for data reproduction. The steps of the data synthesis proposed are illustrated in Fig. 9a:

(1) The original data is divided into L parts, with each part containing data from h years, equivalent to the length of the synthetic data.
(2) EMD is applied to each data part to obtain p IMFs and 1 residue, where p ranges between $\log_2 n - 1$ and $\log_2 n$ [25], and n represents the sample size of the data.
(3) Assuming similarity in the IMFs and residue of each part, the parts can be interchanged with each other. This results in a total of $L^{(p+1)}$ sequence combinations that can be reproduced.

In this study, the proposed concept of data generation is applied to the synthesis of one-year SST data, with h = 1. Additionally, the synthesis is performed in a 5-year moving process, as depicted in



(a)



(b)

Fig. 9. Concept of SST data synthesis; and (b) projection based on a 5-year moving process.

Fig. 9b, resulting in L = 5. The following exploration investigates whether the characteristics of the IMF and residue are similar for each year and examines the changes in the properties of the sequences produced through the mutual exchange of each year.

6.1. Characteristics of annual daily SST data decomposed by EMD

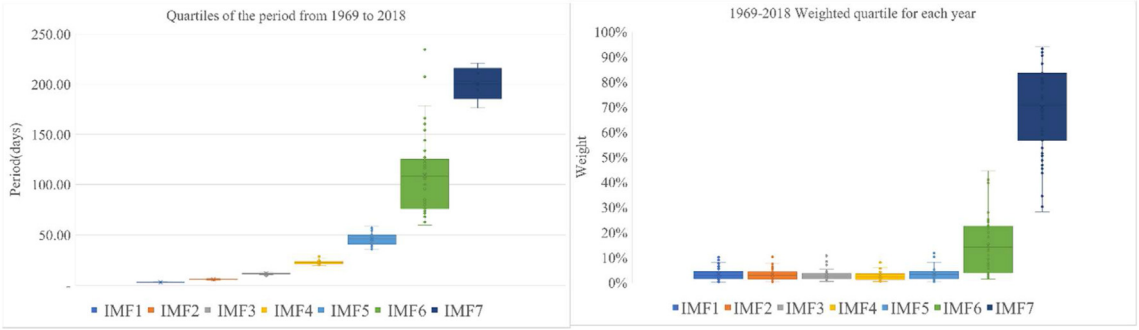
Based on the SST data spanning from 1969 to 2018, it can be divided into 50 1-year-long daily sea temperature series. Leap years are not considered in the study, so there will be 365 data points for one year. Applying EMD to these series, the number of SST samples per year (n) can be decomposed into 7 IMFs and 1 residue. The decomposition results for each year exhibit similarities in both the IMF cycle and the weight distribution compared to the 50-year decomposition. Fig. 10 presents the IMF cycles and weights of the SST data for each year. Since the 1-year data can only be decomposed into 7 IMFs, the cycle of the 7th IMF may not be less clear, but its trend suggests similarity to the 365-day cycle observed in the 50-year IMF7. Furthermore, the average weights for 50 1-year IMF7 is 0.70, indicating its dominant position in data synthesis, while IMF6 accounts for 0.15, and IMF1 to IMF5 contribute 0.03 each. As for the residues, the variations follow a quadratic function. The annual residues exhibit striking similarities (refer to Fig. 11). Over the course of 50 years, the daily sea temperature has changed from 24 °C to 29 °C with seasonal variations, and higher temperatures primarily occurring in summer. However, the residues after 2014 show a notable increase, particularly in 2018, where the residues reach their peak, with the highest value reaching 30 °C. These higher residues further indicate the ongoing rise in sea temperature post-2014.

In essence, the 1-year EMD decomposition is conducted as an independent analysis. In contrast, the 50-year EMD reveals that the amplitude difference has decreased and the residue temperature has risen after 2014. The 50-year residue provides insights into the long-term changes in sea temperature, whereas the 1-year residue only reflects seasonal variations. Typically, IMFs of the same level in a 1-year SST exhibit similarities, implying that the reorganized data should not undergo significant changes unless the residuals of the data itself differ considerably from other sample data. This is because, during data reorganization, the influence of residues outweighs the combined impact of the IMFs. Therefore, when the residues are distinctive, such as in 2018, high-temperature sequences can be synthesized.

6.2. SST projection based on a 5-year moving process

In this study, the generation of annual SST sequence will be performed using a 5-year moving process, as illustrated in Fig. 9b. By using the SST data from the previous 5 years as the base period, the EMD generation process can yield up to 390,625 one-year daily SST data (5^8). Each year's SST data will represent the potential daily SST values for that specific year. Utilizing these data, a probability distribution can be constructed, such as for the annual temperature of the following year, and the analysis result can be compared with the actual annual temperature for validation. Similarly, the sea temperature data of the most recent 5 years will be selected sequentially for EMD reproduction.

This article gathered SST data from 1969 to 2018. Initially, the SST from 1969 to 1973 was utilized for data reproduction, resulting in the generation of 390,625 sets of one-year-long SST sequences



(a) (b)

Fig. 10. (a) Period; and (b) weight of each IMF in the one-year SST.

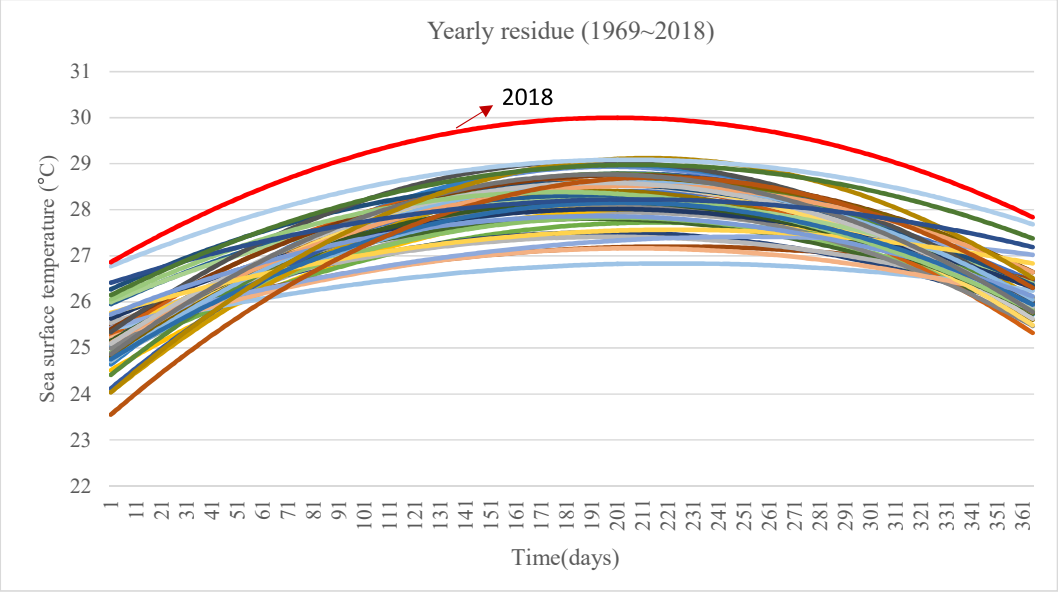
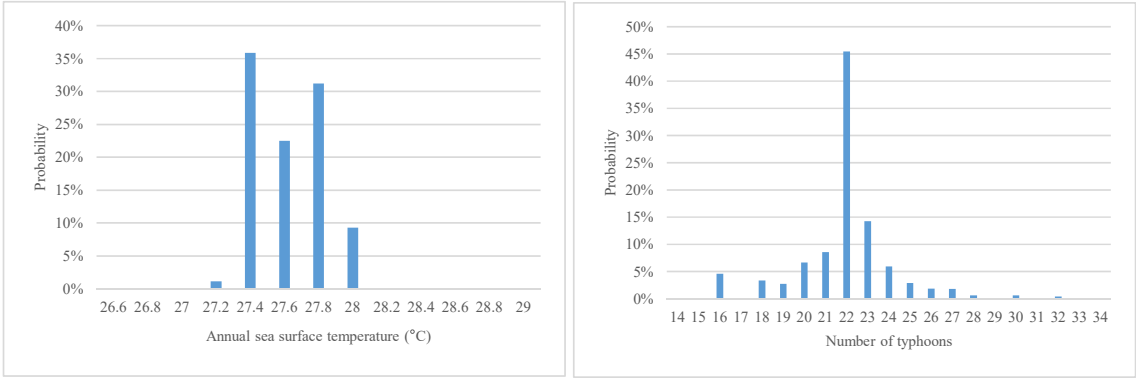


Fig. 11. Residue in each year.

representing possible SST values in 1974. The probability distribution of the average annual temperature for the reproduction sequences is depicted in Fig. 12a. The distribution ranges from 27.2 °C to 28.0 °C and exhibits a double-peak pattern. Among all the 390,625 sequences, the annual temperature distribution ranges from 27.2 °C to 28.0 °C. The figure shows the double-peak mode. The first peak is at 27.4 °C (27.3 °C–27.4 °C, 85,779 sequences in total), the ratio reaches 22%; the other peak is at 27.8 °C (27.7 °C–27.8 °C, 85,626 sequences in total), the ratio is also about 22%. In terms of data sorting, (first quartile, median, third quartile, average) equal (27.34 °C, 27.48 °C, 27.73 °C, 27.53 °C) respectively, and the actual annual sea temperature in 1974 was 27.43 °C, between the first quartile and median. Compared with the annual temperature of all the synthetic sequences, the exceedance probability, P(Temp.>27.43 °C), is equal to 57.19%, signifying that the return period of 27.43 °C is 1.75 years. In fact, based on the "random" observational data (1969–1997), the annual temperature in 1974 ranked sixth from the bottom. From a frequency analysis perspective, the temperature return period for that year is approximately 1.25 years, similar to the results obtained from the reproduction sequences. Furthermore, Fig. 13 demonstrates the close resemblance between the monthly average temperature of the projected SST and the actual SST in 1974.



(a) (b)

Fig. 12. (a) Projected probability distribution of the average annual SST in 1974; and (b) projected probability distribution of the number of typhoons in 1974.

Moving forward, the EMD synthesis process continues using the five-year SST sequence from 1970 to 1974 to predict the SST sequence in 1975, extending it to 2018. Fig. 14 presents a comparison between the actual annual temperature from 1974 to 2018 and the annual temperatures of the synthetic sequences. Observing the figure, it becomes apparent that while the annual temperatures of all synthetic sequences after 2014 exhibit a consistent increasing trend after 2014, the actual annual temperatures surpass the third quartile of the synthetic sequences. For instance, the actual annual temperature in 2018 amounts to 28.79 °C. Among the 390,625 sequences synthesized based on sea temperature data from 2013 to 2017, the exceedance probability of P(Temp.> 28.79 °C) stands at 3.08%. Thus, the return period for a temperature of 28.79 °C is estimated to be 32.43 years, signifying an extreme event. As the annual temperatures have visibly increased each year after 2014, their statistical characteristics differ from those of preceding years, indicating non-stationarity in the series of annual temperature changes. Consequently, conducting frequency analysis solely based on annual temperature variations is deemed inappropriate. This further highlights the superiority of the EMD synthesis method.

Furthermore, Fig. 13 also depicts the disparity between the EMD synthesis temperature and the observed temperature for each month in 2018. The trend of the EMD value aligns closely with the observed values, and the summer temperatures generated from June to September remain in proximity to the actual value. However, significant differences arise in other months, particularly from January to March. In reality, the monthly temperatures in 2018 were generally higher than those in 1974, exhibiting an increase of at least 0.59 °C from July to September and reaching 1.89 °C from January to March. Evidently, the impact of climate change on sea temperature is substantial, and it is crucial to acknowledge the escalating frequency of so-called extreme events.

Drawing on the SST data from 2014 to 2018, an estimate for the average annual sea temperature in 2019 is 28.41 °C. The quartiles are 27.99 °C (lower), 28.40 °C (median) and 28.73 °C (upper). These estimates can be compared with the subsequent SST observation data provided by NOAA for validation.

6.3. Forecast of the number of typhoons

In the preceding sections, this article has effectively employed EMD to generate non-stationary sequences of daily sea temperature and utilized linear regression to estimate the expected number of typhoons per year. In this section, we will combine these findings to project the probability distribution of annual typhoon frequency through a 5-year moving process.

To illustrate the projection methodology in detail, let's consider an example using historical sea temperature data from 1969 to 1973. The following steps will be followed:

(1) EMD generation: Utilize EMD to generate a total of 390,625 (5^8) sequences based on sea temperature data from 1969 to 1973.

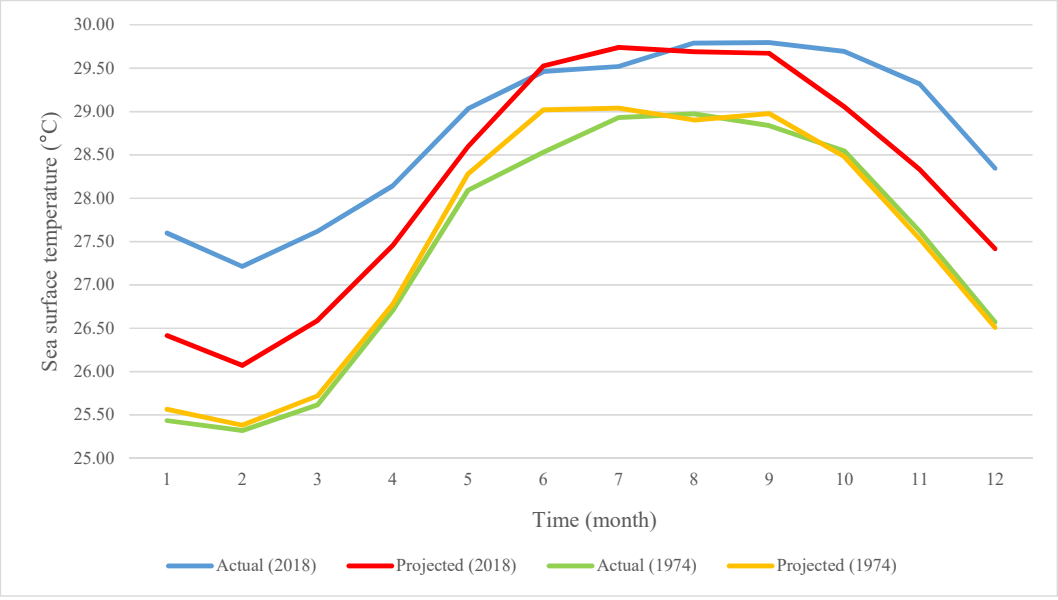


Fig. 13. Comparison of the projected SST and actual SST for each month in 1974 and 2018.

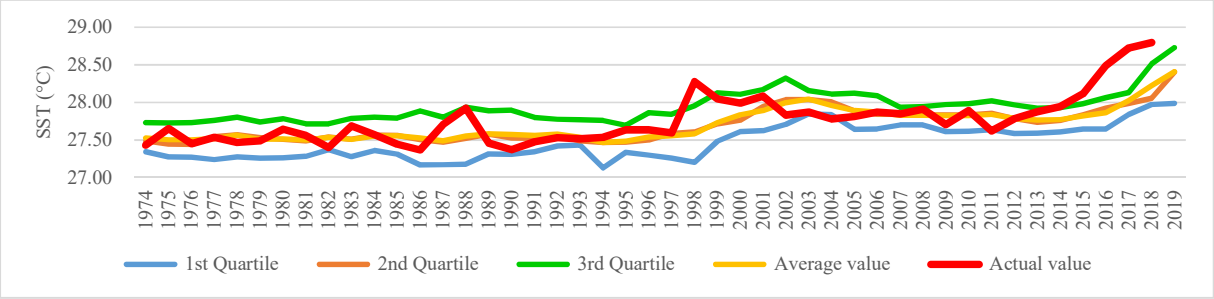


Fig. 14. Comparison of the average annual SST between the projection and observation (1974–2018).

(2) Threshold Estimation: Set cutting thresholds at 28.91 °C and 29.30 °C to estimate the number of typhoons in each generated sequence.
(3) Probability Distribution Analysis: Analyze the probability distribution of the occurrence counts across all the generated sequences. This distribution represents the range of possible typhoon frequencies for the year 1974.

Since 28.91 °C and 29.30 °C are the thresholds respectively, the number of typhoons is estimated 781,260 ($5^8 \times 2$) times. In addition, due to the limitation of $0 < Y \leqq 1$ (see Eq. (2)), X_1 (see Eq. (1)) has its feasible region. In this case, there are 716,449 valid samples that satisfy the constraints. The probability distribution diagram is depicted in Fig. 12b. The number of typhoons ranges from 16 to 32, with an average of 22. Among all the sample data, there are 325,759 instances with 22 typhoons, accounting for 45.47% of the total. The estimated number of typhoons in 1974, with a 95% confidence interval, falls between 18 and 30. The minimum estimated number is 16 (with 33,084 samples) and the maximum is 32 (with 3368 samples). The respective ratios of these extremes are 4.62% and 0.47%. However, in reality, there were 31 typhoons observed in the region between 5°N~25°N and 110°E~170°E in 1974. This value lies just outside the 95% confidence interval. Notably, among the years spanning from 1958 to 2019, 1974 ranked fourth in terms of typhoon occurrence, making it one of the years with a higher number of typhoons.

Fig. 15 displays the 95% confidence interval for the estimated number of typhoons from 1974 to 2019. The average number of typhoons per year falls with the range of 22 and 23. In most cases, the actual number of typhoons observed each year falls within this confidence interval. Out of the 46-year projected values, only 5 years of historical data exceeded the 95% confidence interval. Specifically, 1974 and 2018 exceeded the upper limit of 95%, while 1998, 1999, and 2010 were below the lower limit of 95%. Let's take the example of 2018: out of 781,013 valid samples, the estimated values were mostly concentrated between 21 and 23 occurrences, accounting for 97.15% of the total. The range of 24–26 occurrences accounted for only 0.49%. The actual number of typhoon occurrences in 2018 was 27, slightly outside the estimated range. However, the projection of 24 times in 2019 fell within the 95% confidence interval once again. The results demonstrate the feasibility of combining EMD synthesis and regression analysis for the projection of typhoon frequency.

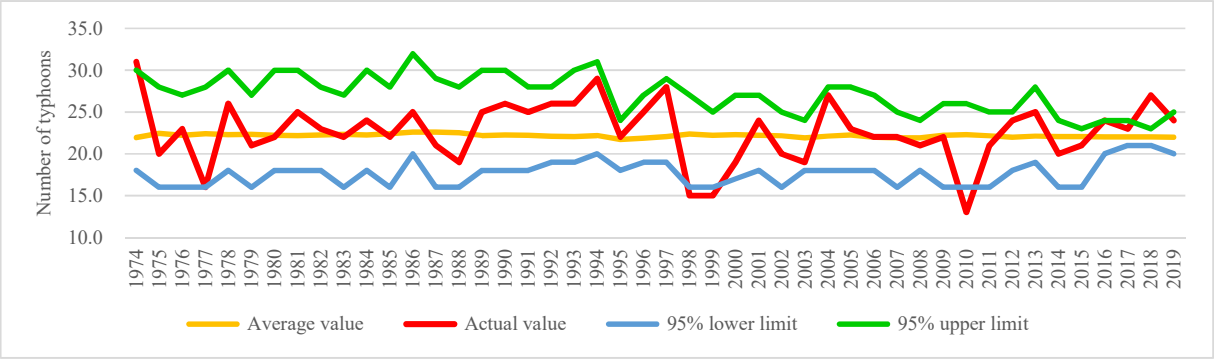


Fig. 15. Comparison of annual typhoon numbers between the forecast and observation (1974–2019).

7. Conclusion

This article successfully accomplishes two objectives. Firstly, it generates non-stationary sequences resembling sea surface temperature changes. Secondly, it utilizes these generated daily sea surface temperature sequences to forecast the annual number of typhoons in a year.

The results demonstrate the viability of the proposed EMD-based data generation method. This process can be repeated annually, following a similar moving pattern, to synthesize new sequences.

Predicting the frequency of typhoons is known to be a challenging task. This paper employs the generated SST series to forecast typhoon frequency. The resulting SST series are segmented based on various thresholds, and a two-stage simple linear regression is used to estimate the number of typhoons. The findings indicate that most of the observed values fall within the predicted 95% confidence interval. Hence, it is feasible to explore the typhoon frequency based on year-to-year sea temperature changes.

Having achieved the objective of generating non-stationary sea temperature series using EMD, this study can be further extended to other areas of sea temperature projection. For instance, it can be applied to predict phenomena like the El Niño phenomenon.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was sponsored by the Ministry of Science and Technology in Taiwan (MOST 110-2221-E-019-012-). The authors sincerely thank Taiwan's Central Weather Bureau for providing the data on the number of typhoons, and the National Oceanic and Atmospheric Administration (NOAA) for providing daily sea surface temperature data in relevant sea areas.

References

[1] Loucks DP, Stedinger JR, Haith DA. Water resources system planning and analysis. Old Tappan, N. J.: Prentice-Hall; 1981.
[2] Yeh WW-G. Reservoir management and operations models: A state-of-the-art review. Water Resour Res 1985;21(12): 1797−818. https://doi.org/10.1029/WR021i012p01797.
[3] Huang WC, Yuan LC, Lee CM. Linking genetic algorithms with stochastic dynamic programming to the long-term operation of a multi-reservoir system. Water Resour Res 2002; 38(12). 40-41-40-9. https://doi.org/10.1029/2001WR001122.
[4] Huang WC, Chang CW. Water shortage risk in Taiwan's Silicon Valley. J Chin Inst Eng 2022;45(6). https://doi.org/10.1080/02533839.2022.2078420.
[5] Oliveira R, Loucks DP. Operating rules for multireservoir systems. Water Resour Res 1997;33(4):839−52. https://doi.org/10.1029/96WR03745.
[6] Wang Z, Zhang L, Cheng L, Liu K, Ye A, Cai X. Optimizing operating rules for a reservoir system in northern China considering ecological flow requirements and water use priorities. J Water Resour Plann Manag 2020;146(7). https://doi.org/10.1061/(ASCE)WR.1943-5452.0001236.
[7] Salas JD, Delleur JW, Yevjevich V, Lane WL. Applied modeling of hydrologic time series. Littleton, Colorado: Water Resources Publications; 1985.
[8] Harms AA, Campbell TH. An extension to the Thomas-Fiering Model for the sequential generation of streamflow. Water Resour Res 1967;3(3):653−61. https://doi.org/10.1029/WR003i003p00653.
[9] Pastor F, Valiente JA, Khodayar S. A warming Mediterranean: 38 years of increasing sea surface temperature. Rem Sens 2020;12(17). https://doi.org/10.3390/rs12172687.
[10] Mohamed B, Nagy H, Ibrahim O. Spatiotemporal variability and trends of marine heat waves in the Red Sea over 38 years. J Mar Sci Eng 2021;9(8). https://doi.org/10.3390/jmse9080842.
[11] Lee MA, Huang WP, Shen YL, Weng JS, Semedi B, Wang YC, Chan JW. Long-term observations of interannual and decadal variation of sea surface temperature in the Taiwan Strait. J Mar Sci Technol 2021;29(4). Article 7. https://doi.org/10.51400/2709-6998.1587.
[12] Cheng L, Abraham J, Trenberth KE, et al. Another record: ocean warming continues through 2021 despite La Niña conditions. Adv Atmos Sci 2022;39:373−85. https://doi.org/10.1007/s00376-022-1461-3.
[13] Wu CC, Kuo YH. Typhoons affecting Taiwan: current understanding and future challenges. Bull Am Meteorol Soc 1999;80(1):67−80. https://doi.org/10.1175/1520-0477(1999)080<0067:TATCUA>2.0.CO;2.
[14] Ooyama K. Numerical simulation of the life cycle of tropical cyclones. J Atmos Sci 1969;26(1):3−40. https://doi.org/10.1175/1520-0469(1969)026<0003:NSOTLC>2.0.CO;2.
[15] Lin YF, Wu CC, Yen TH, Huang YH, Lien GY. Typhoon Fanapi (2010) and its interaction with Taiwan terrain − evaluation of the uncertainty in track, intensity and rainfall simulations. J Meteor Soc Japan 2020;98(1):93−113. https://doi.org/10.2151/jmsj.2020-006.
[16] Hung CWA. 300-year typhoon record in Taiwan and the relationship with solar activity. Terr Atmos Ocean Sci 2013; 24(4):737−43. https://doi.org/10.3319/TAO.2013.02.18.01(A). Part II.
[17] Camargo SJ, Barnston AG, Klotzbach PJ, Landsea CW. Seasonal tropical cyclone forecasts. World Meteorol Organ Bull 2007;56(4):297−309.
[18] Wang C, Wang B, Wu L. A region-dependent seasonal forecasting framework for tropical cyclone genesis frequency in the Western North Pacific. Am Meteorol Soc 2019;32: 8415−35. https://doi.org/10.1175/JCLI-D-19-0006.1.
[19] Riehl H. On the formation of typhoons. J Meteorol 1948;5(6): 247−65. https://doi.org/10.1175/1520-0469(1948)005<0247:OTFOT>2.0.CO;2.
[20] Yanai M. Dynamical aspects of typhoon formation. J Meteorol Soc Jpn 1961;39(5):282−309. https://doi.org/10.2151/jmsj1923.39.5_282.
[21] Gray WM. The formation of tropical cyclones. Meteorol Atmos Phys 1998;67:37−69.
[22] Wang B, Wu R, Lau KM. Interannual variability of Asian summer monsoon: contrasts between the Indian and Western North Pacific-East Asian monsoons. J Clim 2001;14(20): 4073−90. https://doi.org/10.1175/1520-0442(2001)014<4073:IVOTAS>2.0.CO;2.
[23] Choi JW, Kim BJ, Zhang R, Park KJ, Kim JY, Cha Y, Nam JC. Possible relation of the western North Pacific monsoon to the tropical cyclone activity over western North Pacific.

Int J Climatol 2016;36(9):3334−45. https://doi.org/10.1002/joc.4558.

[24] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. Proc Royal Soc A Math Phys Eng Sci 1998; 454(1971):903−95. https://doi.org/10.1098/rspa.1998.0193.

[25] Lu CL, Hsu CY, Huang PH. Analysis of power system low frequency oscillation with empirical mode decomposition. J Mar Sci Technol 2018;26(4). Article 1. https://doi.org/10.6119/JMST.201808_26(4).0001.

[26] Huang WC, Chu TY, Jhang YS, Lee JL. Data synthesis based on empirical mode decomposition. ASCE J Hydrol Eng 2020;25(7): 04020028. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001935.