

Article

Leveraging Transfer Learning and U-Nets Method for Improved Gap Filling in Himawari Sea Surface Temperature Data Adjacent to Taiwan

Dimas Pradana Putra ¹  and Po-Chun Hsu ^{1,2,*} 

¹ Center for Space and Remote Sensing Research, National Central University, Taoyuan 320317, Taiwan; dimaspradana@g.ncu.edu.tw

² Institute of Hydrological and Oceanic Science, National Central University, Taoyuan 320317, Taiwan

* Correspondence: hpochun@csrsr.ncu.edu.tw

Abstract: Satellite sea surface temperature (SST) images are valuable for various oceanic applications, including climate monitoring, ocean modeling, and marine ecology. However, cloud cover often obscures SST signals, creating gaps in the data that reduce resolution and hinder spatiotemporal analysis, particularly in the waters near Taiwan. Thus, gap-filling methods are crucial for reconstructing missing SST values to provide continuous and consistent data. This study introduces a gap-filling approach using the Double U-Net, a deep neural network model, pretrained on a diverse dataset of Level-4 SST images. These gap-free products are generated by blending satellite observations with numerical models and in situ measurements. The Double U-Net model excels in capturing SST dynamics and detailed spatial patterns, offering sharper representations of ocean current-induced SST patterns than the interpolated outputs of Data Interpolating Empirical Orthogonal Functions (DINEOFs). Comparative analysis with buoy observations shows the Double U-Net model's enhanced accuracy, with better correlation results and lower error values across most study areas. By analyzing SST at five key locations near Taiwan, the research highlights the Double U-Net's potential for high-resolution SST reconstruction, thus enhancing our understanding of ocean temperature dynamics. Based on this method, we can combine more high-resolution satellite data in the future to improve the data-filling model and apply it to marine geographic information science.

Keywords: sea surface temperature; Himawari satellite; gap-filling; U-Nets; leveraging transfer learning



Citation: Putra, D.P.; Hsu, P.-C. Leveraging Transfer Learning and U-Nets Method for Improved Gap Filling in Himawari Sea Surface Temperature Data Adjacent to Taiwan. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 162. <https://doi.org/10.3390/ijgi13050162>

Academic Editors: Diego González-Aguilera, Pablo Rodríguez-Gonzálvez and Wolfgang Kainz

Received: 16 February 2024

Revised: 1 May 2024

Accepted: 11 May 2024

Published: 13 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ocean, as a major component of Earth's surface, profoundly impacts the climate system, ecological balance, and global environmental changes. In the current Earth system science research, acquiring and calculating spatial data parameters in oceanography is particularly crucial. Among these, the sea surface temperature (SST) dataset serves as a fundamental indicator of oceanic heat content, measured by microwave or infrared wavelengths in the electromagnetic spectrum [1,2]. Variations in SST play a key role in the exchange of energy and moisture between the atmosphere and the ocean [3,4]. SST databases, established using geospatial artificial intelligence computations, are vital for accurately monitoring SST changes to understand and predict the dynamics of climate change, oceanic processes [3], and the overall health of the marine ecosystem [5–8]. With the advent of satellite technology, satellite-derived SST data have become an essential resource for studying global oceanic and atmospheric phenomena. Microwave sensors can observe seawater despite having lower spatial resolution, as they can penetrate clouds, fog, and dust in almost all situations except during the heaviest rainfall [9–11]. Infrared sensors provide higher-resolution SST observations, but their inability to penetrate cloud cover leads to data interruptions and gaps, as the sensors cannot effectively capture SST values obscured by clouds [12–16]. These data gaps can lead to inconsistencies in SST records,

affecting our ability to detect trends, limit cloud radiation feedback, and pose challenges in the precise analysis and modeling of oceanic and atmospheric processes [17]. Gaps in SST data have significant impacts on weather forecasting, climate research, and marine ecosystem management [1,18]. Complete SST data coverage is crucial for comprehensive analyses of long-term climate trends, detecting anomalies, and assessing the impacts of climate change on marine ecosystems [19,20]. Therefore, filling these data gaps is necessary for the successful use and analysis of the data.

Traditional spatial data modeling and geostatistical interpolation techniques have historically been employed to address data gaps in satellite-based marine observations [21–24]. Early efforts in this field included the development of Optimum Interpolation (OI), Empirical Orthogonal Functions (EOFs), and the Expectation Maximization (EM) method [25], which aimed to mitigate the limitations of available data. One of the most widely recognized and effective methods for filling data gaps in oceanographic datasets is the Data Interpolating Empirical Orthogonal Functions (DINEOFs) method, proposed by Beckers and Rixen in 2003 [26]. The DINEOFs method, which utilizes EOFs, has demonstrated significant efficiency improvements, achieving speeds approximately 30 times faster than Optimum Interpolation (OI) [26]. It has been extensively applied and tested in various studies, facilitating the interpolation of single or multiple related marine variables [27,28]. Given its proven performance and widespread acceptance, DINEOFs serves as a valuable benchmark for evaluating the effectiveness of our proposed approach.

While statistical and machine learning methods, such as Self-Organizing Maps (SOMs) [29], Random Forests [30,31], and Multilayer Perceptrons [32], have been proposed for estimating data under cloud-covered pixels, they may have limitations in capturing the intricate spatial patterns and complex physical processes that drive SST dynamics [33–36]. These methods can struggle with prolonged data gaps, as accurately identifying oscillations and cycles becomes challenging [34]. Additionally, they may not fully account for the interactions among various factors affecting temperature changes, such as ocean currents, atmospheric circulation, and solar radiation [35].

Currently, advancements in deep machine learning methods have brought deep learning techniques to the forefront of filling data gaps. Barth et al. (2020) proposed using a Data-Interpolating Convolutional Auto-Encoder (DINCAE) [37], a convolutional neural network, to fill data gaps in satellite-observed SST. DINCAE, leveraging existing cloud-free pixels, surpasses traditional methods like DINEOF in reconstructing missing data. Additionally, other deep learning approaches have shown promise in reconstructing missing data and enhancing the accuracy of marine data products, particularly SST. For instance, Patil et al. (2022) introduced several deep learning models [38], including a Deep Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM), and a Spatial 2D Convolution Neural Network. These models, trained with past meteorological features, predict the next day's SST in the northeast region of Japan. The study revealed that the Spatial 2D CNN was highly skilled at coastal locations, while the Deep MLP and LSTM were equally proficient at offshore locations. Another study developed a Bi-directional Long Short-Term Memory (BiLSTM) model, featuring a novel and adjustable custom temporal penalty layer. This model, designed for filling daily night-time SST images in spatiotemporal sequences acquired by the Himawari-8 satellite, robustly reconstructed the spatial and temporal features of cloud-affected SST data. It outperformed the previous state-of-the-art Simple Spatial Gap-filler (SSGP) algorithm in error metrics, Structural Similarity Index (SSIM), and computational efficiency [39]. In recent years, deep learning techniques have emerged as a powerful tool for handling missing data and capturing complex patterns in various fields [40]. For example, in image-based tasks, deep learning models of convolutional neural networks, such as U-Net, have demonstrated exceptional performance in accurately capturing details and spatial relationships. Originally developed for medical image segmentation, the U-Net architecture has shown significant ability in solving complex image-based problems. In the context of SST data gap filling, utilizing deep learning, par-

ticularly the U-Net architecture, could potentially overcome the limitations of traditional methods, providing more precise estimates for missing SST values.

In light of previous studies, our research introduces a U-Net model for filling gaps in sea surface observational data. Building upon existing knowledge and methodologies, we aim to provide new insights into accurately estimating marine geographic data, leveraging the strengths of our proposed method to capture the complex dynamic features of the sea surface. Specifically, we propose a deep learning model with a Double U-Net architecture to fill missing values in satellite SST data. To evaluate its performance, we conducted a comparative analysis with the widely used DINEOFs method, assessing the model's ability to capture SST geospatial patterns and address data gaps in the waters around Taiwan. This region, with its complex oceanographic features, serves as an ideal testing ground for evaluating the performance of our pre-trained U-Net model. Grounded in geoinformation science techniques and methodologies, this study addresses the following four questions: (1) whether high spatial resolution geostationary Himawari satellite SST data are suitable for data gap filling; (2) whether the U-Net architecture and transfer learning techniques can enhance data usability; (3) how the spatial information of SST filled by deep learning compares with traditional DINEOFs methods and in situ observational data; (4) whether the gap-filled SST data can be used for monthly and long-term SST analysis and if spatial features can successfully capture the SST dynamics mentioned in previous research. Through the research on leveraging transfer learning and U-Nets for improved gap filling in Himawari SST data, the findings of this study are significant for advancing our understanding of dynamic marine geographic data and contribute to methods for filling gaps in satellite-derived SST datasets in other sea areas using similar approaches.

2. Data and Methods

2.1. Data

2.1.1. In Situ Buoy Observation

The in situ observations utilized in this research were obtained from 10 sites situated around Taiwan, generously provided by the Taiwan Central Weather Administration (CWA). These in situ measurements serve as ground-truth data, against which we validate the predictions of our model. Table 1 presents detailed information about the ground-truth observation data collected from each of the 10 sites, including the corresponding number of missing satellite data points in their respective areas. Additionally, Figure 1 depicts the geographical locations of the in situ SST data collection, providing a visual representation of the spatial coverage of these measurements.

Table 1. Details of buoy locations and number of data samples of gapped satellite data at each in situ location.

Buoy	Co-Ordinates		Data Sample Number
	Lat	Lon	
Pengjiayu	25°37'13" N	122°04'06" E	400
Fugui	25°18'14" N	121°32'01" E	404
Long Dong	25°05'52" N	121°55'21" E	495
Guishandao	24°50'54" N	121°55'31" E	468
Hsinchu	24°45'47" N	120°50'37" E	257
Taichung	24°12'56" N	120°24'48" E	211
Hualien	24°01'54" N	121°37'55" E	481
Chimi	23°11'14" N	119°39'22" E	213
Xiao Liuqiu	22°19'00" N	120°22'24" E	166
Lanyu	22°04'31" N	121°34'58" E	301

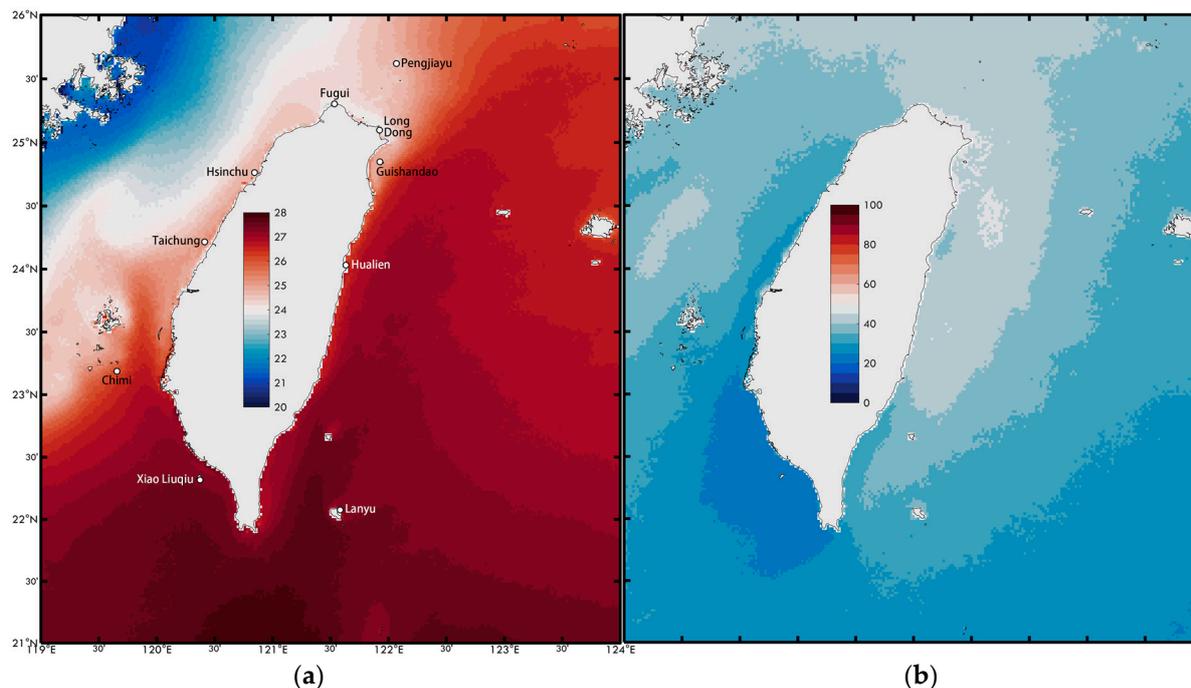


Figure 1. (a) The SST ($^{\circ}\text{C}$) surrounding Taiwan, observed by the Himawari satellite from 2016 to 2023, with white dots representing the positions of buoy stations. (b) Daily missing SST data rate (%) of spatial grid observed by the Himawari satellite, from July 2015 to December 2023.

2.1.2. Himawari-8/9 Satellite Data

We utilize the daily Level 4 (L4) Himawari SST dataset for the year 2022. This dataset encompasses the area around Japan, including the Taiwan, spanning from 117°E to 150°E longitude and from 17°N to 50°N latitude. We accessed the L4 dataset through the JAXA website (<https://www.eorc.jaxa.jp/ptree/index.html>, accessed on 20 May 2023). The L4 dataset is constructed using a data assimilation process that integrates information from various sources, including the high-resolution regional ocean model ‘JCOPE-T’, developed by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC). This model assimilates observational data from six types of satellite SST provided by JAXA, offering a comprehensive representation of SST in the region, particularly around Taiwan. We employed a cropping process on the daily L4 Himawari SST data to tailor it to our specific study area. This cropped dataset covers a latitude range of 21°N to 26°N and a longitude range of 119°E to 124°E , which aligns with the spatial extent relevant to our research objectives (see Figure 1a). The area is characterized by complex oceanographic features, such as proximity to the Kuroshio Current (KC) and diverse topography [41–43]. Notably, the daily data missing rate in the northern part of Taiwan can reach up to 50%, particularly during winter and typhoon seasons, causing significant data gaps. Even in areas with the lowest missing rates, the proportion can reach 20% (see Figure 1b). Additionally, to match the resolution of the Level 3 (L3) Himawari SST product, the original data, with a resolution of 3 km, is linearly interpolated to a higher resolution of 2 km. The resulting cropped dataset consists of SST images with dimensions of 256×256 pixels.

Before commencing the model training phase, we generate an artificial cloud dataset. This dataset is created by superimposing artificial cloud features, derived from the L3 Himawari dataset, onto the cropped L4 Himawari SST dataset within our study area. The artificial clouds, designated as value ‘0’, are designed to emulate the realistic cloud patterns and coverage observed in the region during the study period. This step enhances the model’s capability to manage missing data and incomplete observations due to cloud cover, which are common challenges in SST estimation. The target dataset used for training the model comprises free gap data extracted from the cropped L4 Himawari SST dataset

within the study area. This dataset includes regions where SST observations are missing or incomplete. By using the free gap data as the target, the model aims to estimate and fill in the gaps in SST values, serving as a ground-truth reference for evaluating the model's performance.

To further optimize the model's performance, we incorporate a supplementary dataset for fine-tuning: the L3 Himawari night-time observation daily dataset, spanning from July 2015 to December 2022. This dataset, obtained from the JAXA website, is similar to the L4 dataset. These night-time observations are particularly valuable as they provide insights into the unique characteristics and variations of SST during the darker hours, which minimizes the influence of solar radiation-induced fluctuations. Although the L3 dataset offers high temporal resolution up to 1 h, we utilize it on a daily timescale for our analysis. To determine the most representative time for daily observations, we analyze the absolute error between in situ and satellite data, as shown in Figure 2. Our analysis reveals that 7 a.m. local time (23 UTC) exhibits the lowest Root Mean Square Error (RMSE), indicating its suitability for representing daily SST values. Therefore, we select this specific observation hour to enhance the accuracy and reliability of our model's predictions, enabling a more comprehensive understanding of the temporal dynamics of SST within the study area.

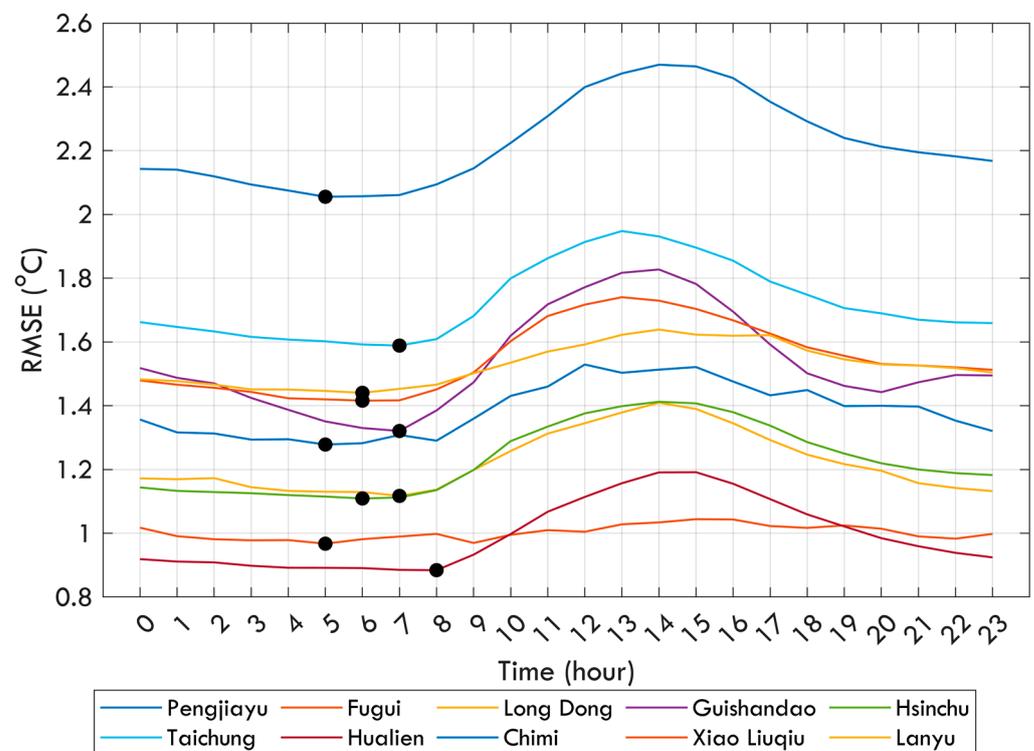


Figure 2. The comparison results of the RMSE values for buoy SST and Himawari SST, with black dots representing the minimum values at each observation location.

2.1.3. Typhoon Dataset

The typhoon dataset utilized in this research was sourced from the International Best Track Archive for Climate Stewardship (IBTrACS) [44]. This dataset provides comprehensive information on tropical cyclones, including their tracks, intensities, and other relevant attributes. The incorporation of typhoon data is crucial for distinguishing days with typhoons and pinpointing instances where the eye of a typhoon intersects with our research area. These records offer invaluable insights into atmospheric conditions, especially during periods of maximum cloud cover associated with typhoons, presenting a significant challenge in testing our proposed method. Addressing this challenge is integral to our study's focus on gap-filling in SST data.

2.2. Data Filling of Double U-Net Architecture

This study employs a Double U-Net architecture, as illustrated in Figure 3, to tackle the challenging task of SST gap filling and estimation. The architecture consists of two interconnected U-Net networks arranged in a cascaded manner. To train this architecture, we used the L4 Himawari SST daily data for the year 2022. This dataset, offering a rich source of information, enables the extraction of valuable physical knowledge about SST dynamics. The Double U-Net architecture is trained to discern intricate relationships and patterns within the SST data. The model inputs include three layers representing satellite images from three consecutive days, while the output predicts non-gap data for the first layer of the input. This approach, based on three days of observations, aims to minimize data gaps. As demonstrated in Figure 4, the SST patterns from the first and second days before the prediction day are relatively consistent. Consequently, combining observations from these three days allows for a more reliable representation of spatial patterns. Regarding the use of three consecutive days of SST data as the model input, this choice was not arbitrary but rather based on prior knowledge and assumptions about the temporal behavior of SST. We hypothesized that SST values do not significantly change over short periods, such as three consecutive days. This hypothesis is reasonable because SST generally exhibits gradual and continuous changes, particularly in the absence of major disturbances or weather events. By incorporating three consecutive days of SST data, we aimed to leverage the temporal correlation and continuity of SST patterns to fill gaps more effectively. It is important to note that we did not solely rely on this assumption. We thoroughly tested and evaluated the impact of using different numbers of consecutive days, including two and four days. Our experiments demonstrated that using three consecutive days (as shown in Figure 4) provided an acceptable balance between capturing sufficient temporal information and avoiding excessive smoothing or distortion of the SST patterns.

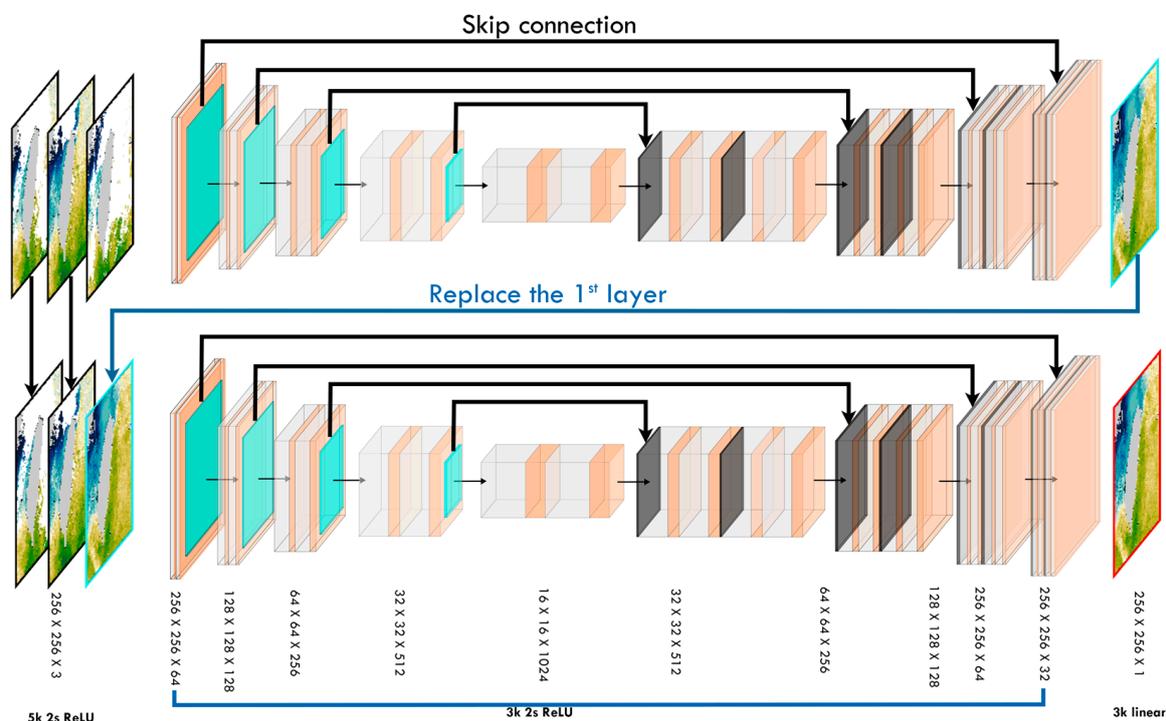


Figure 3. The Double U-Net Architecture. The term ‘k’ represents kernel size, and ‘s’ denotes stride length. For instance, ‘5 k 2 s’ indicates the layer uses a 5 × 5 kernel with a 2 × 2 stride. ‘ReLU’ and ‘linear’ refer to the activation functions used within this architecture. This architecture combines original U-Net features with max pooling and 2D Transpose layers, and L1 regularization is applied following each convolutional layer.

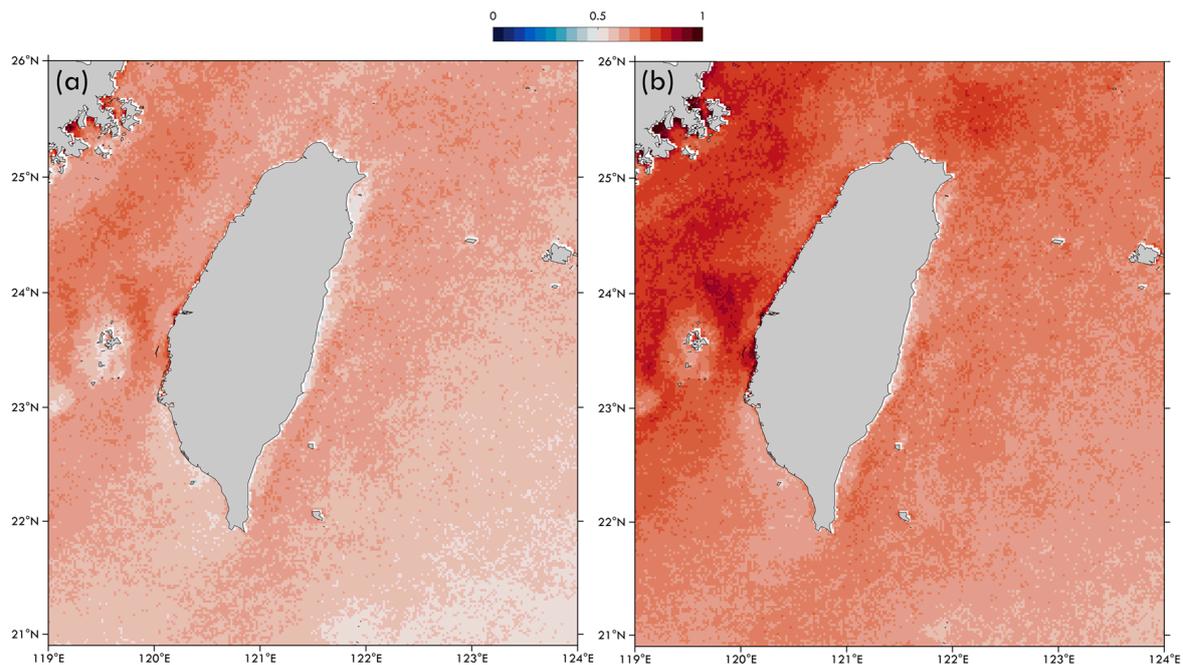


Figure 4. Absolute differences in daily mean Himawari L3 SST (°C): (a) represents the one-day difference, indicating the change in SST from one day to the next, calculated as the average absolute difference between successive days ($\sum \frac{1}{n} |dsst - d_{n-1}sst|$), and (b) represents the two-day difference, indicating the change over a two-day interval, calculated as the average absolute difference over two days ($\sum \frac{1}{n} |dsst - d_{n-2}sst|$).

Within the Double U-Net architecture, the output of the first U-Net serves as the first layer of input data for the second U-Net. This process enables the second U-Net to leverage the enhanced features learned by the first. The training involves optimizing the architecture's parameters using the Adam optimization algorithm, aiming to minimize the discrepancy between predicted SST values and the ground-truth data from the L4 Himawari dataset. Our methodology begins with training the Double U-Net architecture using the L4 Himawari SST daily data. We feed the model input patches of SST data along with corresponding ground-truth data to enable it to learn SST patterns effectively. This step optimizes the model's parameters to minimize the discrepancy between predicted and actual SST values. The training process utilizes the Adam optimization algorithm, starting with an initial learning rate of 10^{-3} , which is progressively decreased to 10^{-7} if no improvement in loss is observed over five consecutive epochs. After the initial training phase, the model undergoes fine-tuning using the official JAXA Himawari daily SST nighttime L3 data product, which covers the period from July 2015 to December 2022. This dataset is particularly valuable as it provides night-time observations, which exhibit lower variability influenced by solar factors. The fine-tuning process begins with a learning rate of 10^{-5} , also decreasing to 10^{-7} if there is no loss improvement over five consecutive epochs.

However, employing the U-Net, a supervised model that requires a target label, presents a challenge when used with the Himawari L3 SST data, which contains gaps, leaving the actual values for those regions unknown. To address this issue, a preprocessing step is conducted prior to the fine-tuning process. Specifically, a subset of valid pixel points is randomly selected to serve as target values for prediction and contribute to the loss function, aiding in optimizing the U-Net model. Approximately 5% of random points from the available data, equating to almost 3000 points in the absence of cloud cover, are chosen for this purpose, as illustrated in Figure 5. Notably, the random selection of pixel locations is performed independently for each satellite image, ensuring the creation of a diverse and representative subset that facilitates effective model optimization.

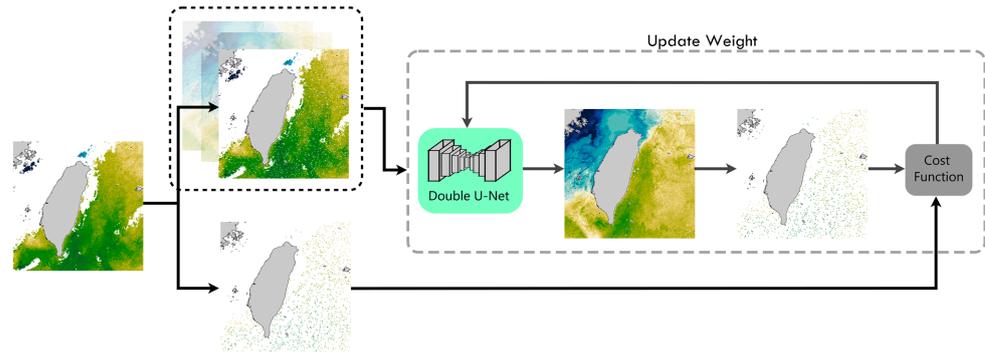


Figure 5. The fine-tuning process using Himawari L3 SST data. The arrows show the direction of data flow, whereas grey arrows indicate the internal processes within the Double U-Net architecture. The gray line moving in the opposite direction represents the process of updating weights.

By incorporating this subset of valid target values, the fine-tuning process can proceed effectively despite the data gaps, enabling the U-Net model to learn from and adapt to the Himawari night-time SST data. The entire training process is illustrated in Figure 6. To evaluate our proposed pre-trained U-Net model’s performance, we conducted a comparative analysis with the widely used DINEOFs method. Proven to be an effective technique for filling gaps in oceanographic datasets, DINEOFs has been applied and tested in multiple studies, facilitating the interpolation of single or multiple related marine variables [43,44]. To validate the accuracy of the gap-filled SST data, we use in situ observations from 10 sites around Taiwan, provided by the CWA Taiwan, as ground-truth data. The model’s performance is evaluated by comparing its predictions to both these in situ observations and the results obtained from the DINEOFs method. Metrics such as Root Mean Square Error (RMSE), Mean Error (ME), and R are utilized to quantify the accuracy of the model’s predictions in filling the gaps and estimating SST values. Additionally, the CWA dataset is highly valuable for assessing the performance of the models, particularly when applied to specific scenarios like typhoon occurrences. Such cases pose a considerable challenge, especially during periods of maximum cloud cover associated with typhoons, thus providing a critical test for our proposed method. Post-typhoon, there tends to be a significant observational gap, leading to a lack of SST data for several days. The identification of typhoon days is based on the IBTrACS dataset, which specifically pinpoints instances when typhoons are within our study area.

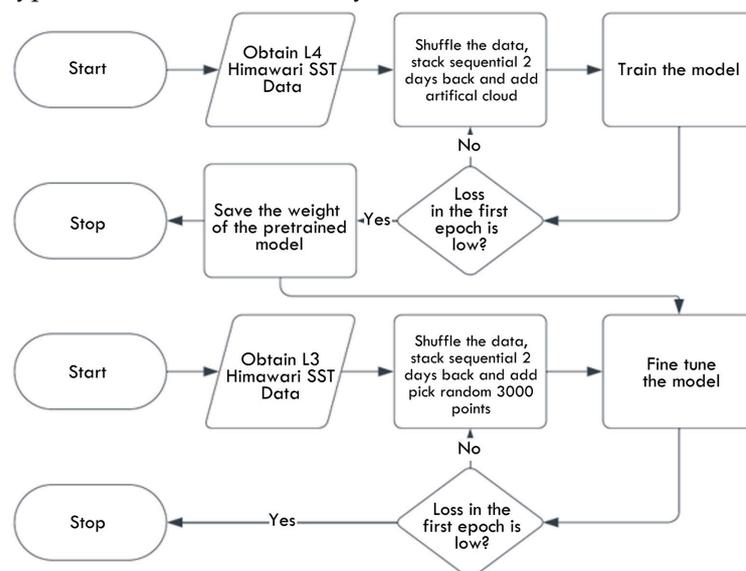


Figure 6. A flow chart that illustrates the interconnected processes involved in producing a pre-trained model and implementing the transfer learning process.

3. Results

3.1. Evaluation of Filling Capabilities of Double U-Net and DINEOF Methods

3.1.1. Comparison with Buoy Data

We introduce a U-Net model designed to fill data gaps in SST observations, and we compare its performance with that of the widely used DINEOFs method. Figure 7 illustrates an example of gap-filled SST data for 29 January 2020, using both our proposed U-Net model (shown in Figure 7d) and the DINEOF method (shown in Figure 7e). Using three consecutive days of Himawari satellite observation data, both approaches successfully generate complete SST maps by filling the gaps in SST data. Due to the KCt being closer to Taiwan during winter, the position at 123° E longitude is no longer within the KC region. From a physical oceanography perspective, the high SST gradient presented in the DINEOFs results appears to be unusually far to the east. However, our U-Net model presents several significant advantages over DINEOFs. Notably, it captures a wider range of SST dynamics, offering a more detailed spatial pattern representation. This advantage is particularly apparent in modeling the KC, where our method provides a sharper and more distinct representation compared to the more interpolation-like appearance produced by DINEOFs.

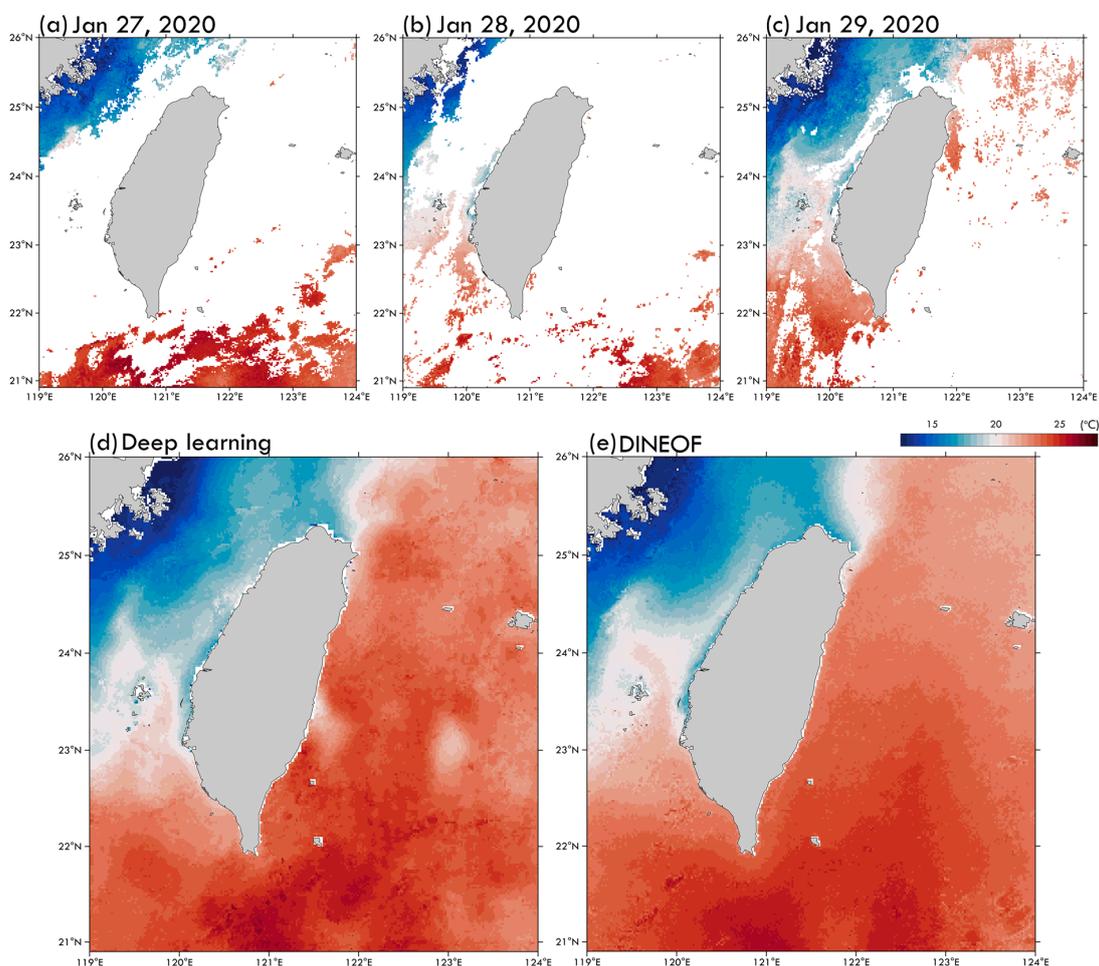


Figure 7. (a–c) The input images (covering the period from 27 January to 29 January 2020) and the output images (for 29 January 2020) generated by (d) the proposed machine learning Double U-Net model and (e) DINEOFs.

The comparison results of the gap-filled method with buoy observations for several periods including our Double U-Net method and DINEOFs method are in Figure 8 and Table 2. It is noteworthy that DINEOFs consistently undervalues SST to a greater extent than our proposed model. Our study's evaluation indicates that the Double U-Net ar-

chitecture outperforms the DINEOFs method in SST estimation for most study areas. In a comparison of 10 buoy observations, the Double U-Net approach for filling SST data showed high correlation results ($R > 0.6$) in nine cases, with an average correlation coefficient of 0.84, except for Lanyu, where the correlation coefficient was only 0.37. In contrast, the DINEOFs method demonstrated high correlation at only four stations, with an average correlation coefficient across all stations of 0.54. Regarding the RMSE results for the gap-filled SST compared to buoy observations, the Double U-Net method yielded an outcome of 2.5 °C, while the DINEOFs method’s results were as high as 5.0 °C. According to previous research [45], there is an inherent RMSE of 1.12 °C between the Himawari satellite and these buoy observations. The reasons for such discrepancies are complex, due to the difference between satellite-observed skin SST and buoy-measured bulk SST. These two types of SST can vary due to the intensity of solar insolation and wind speed, leading to the formation of diurnal warm layers and cool skin layers [45]. For the sake of data consistency, only 7 a.m. local time data were used for daily measurements; hence, it is normal to expect more errors. This error was also reflected in the mean error (ME) comparison; in the early morning when solar radiation is still weak, the average buoy data tend to be significantly warmer than the satellite skin SST. Therefore, the SST filled by both methods turned out to be lower on average compared to the bulk SST of buoy, with the Double U-Net method at -1.44 °C and the DINEOF method at -2.07 °C.

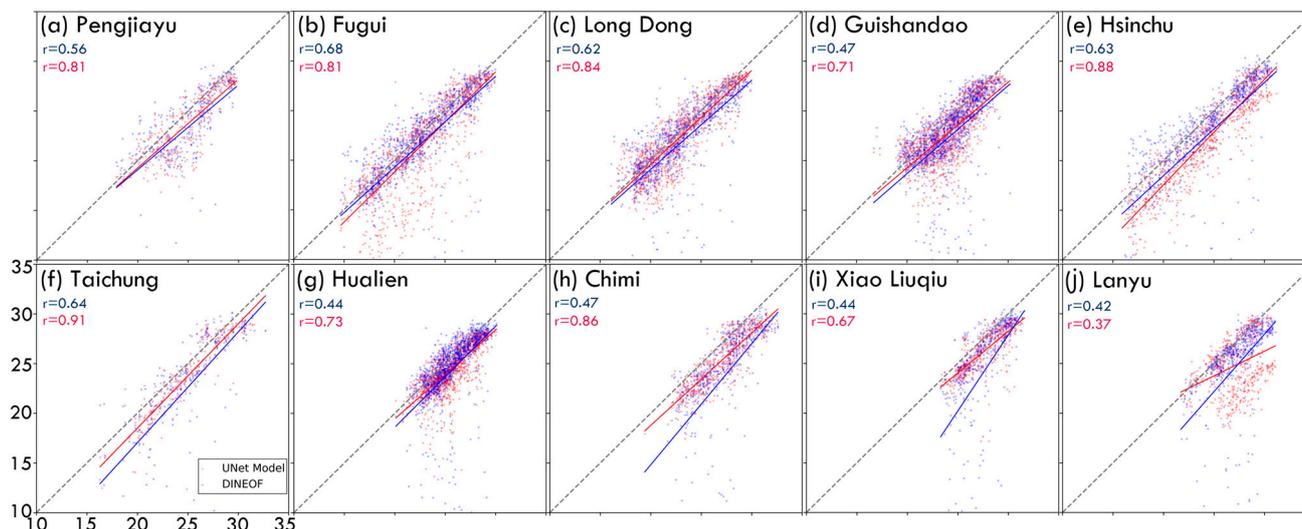


Figure 8. A comparison of SST data, as calculated by the Double U-Net (red dots) and DINEOFs (blue dots) methods (represented on the vertical axis), against buoy observation data (represented on the horizontal axis). (a) Pengjiayu, (b) Fugui, (c) Lond Dong, (d) Guishandao, (e) Hsinchu, (f) Taichung, (g) Hualien, (h) Chimi, (i) Xiao Liuqiu, and (j) Lanyu.

Table 2. The comparison results of SST calculated by the Double U-Net and DINEOFs methods with buoy observations at each station (see Figure 8) include the correlation coefficient (R), Root Mean Square Error (RMSE), and Mean Error (ME).

Buoy	Number of Samples	R		RMSE (°C)		ME (°C)	
		DINEOF	U-Net	DINEOF	U-Net	DINEOF	U-Net
Pengjiayu	318	0.56	0.81	4.03	2.24	-1.61	-1.24
Fugui	1039	0.68	0.81	3.77	2.94	-0.96	-1.19
Long Dong	1065	0.62	0.84	3.79	1.84	-1.23	-0.58
Guishandao	1328	0.47	0.71	4.49	2.48	-1.80	-1.33
Hsinchu	804	0.63	0.88	4.90	3.19	-1.80	-2.21
Taichung	257	0.64	0.91	5.37	2.14	-2.40	-1.26
Hualien	1324	0.44	0.73	4.16	1.98	-1.31	-1.14
Chimi	549	0.47	0.86	6.83	2.21	-3.62	-1.73
Xiao Liuqiu	481	0.44	0.67	7.17	2.31	-3.44	-1.37
Lanyu	562	0.42	0.37	5.72	3.63	-2.50	-2.37

3.1.2. The Challenge of Filling in Data When a Typhoon Passes through

Typhoons bring extensive cloud cover, rendering satellites almost incapable of observing SST. Additionally, typhoons induce vertical movements in the ocean, presenting a significant challenge for filling spatial data gaps. Figure 9 visually illustrates the distribution of typhoon eyes within our study area from July 2015 to December 2022, providing additional details such as typhoon names and the dates they passed through. Notably, while most typhoons traverse our region within a single day, some have a prolonged impact, affecting the area for two days and thus adding complexity to the SST dynamics. Figure 10 showcases the application of the gap-filling algorithm in our study area during the occurrence of several typhoons, specifically Nepartak, Malakas, Megi, and Barijat. The first row of Figure 10 displays the original observational data from Himawari satellite observations, establishing the baseline for SST conditions. The second and third rows display the results of DINEOFs and our U-Net model, respectively, in filling SST data gaps during typhoon events, allowing for a visual assessment of their comparative efficacy.

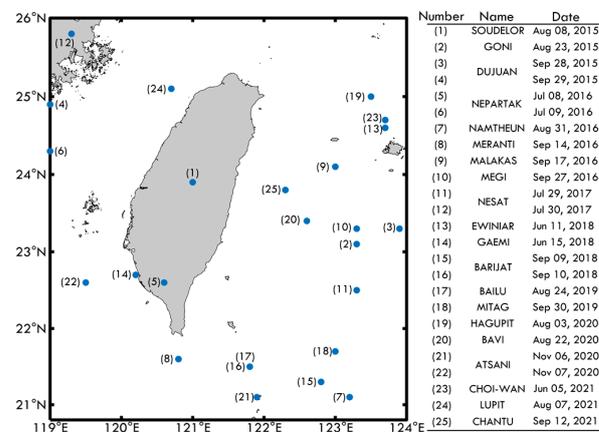


Figure 9. The geographical positions of typhoon eyes and the corresponding dates for typhoons passing through the seas near Taiwan.

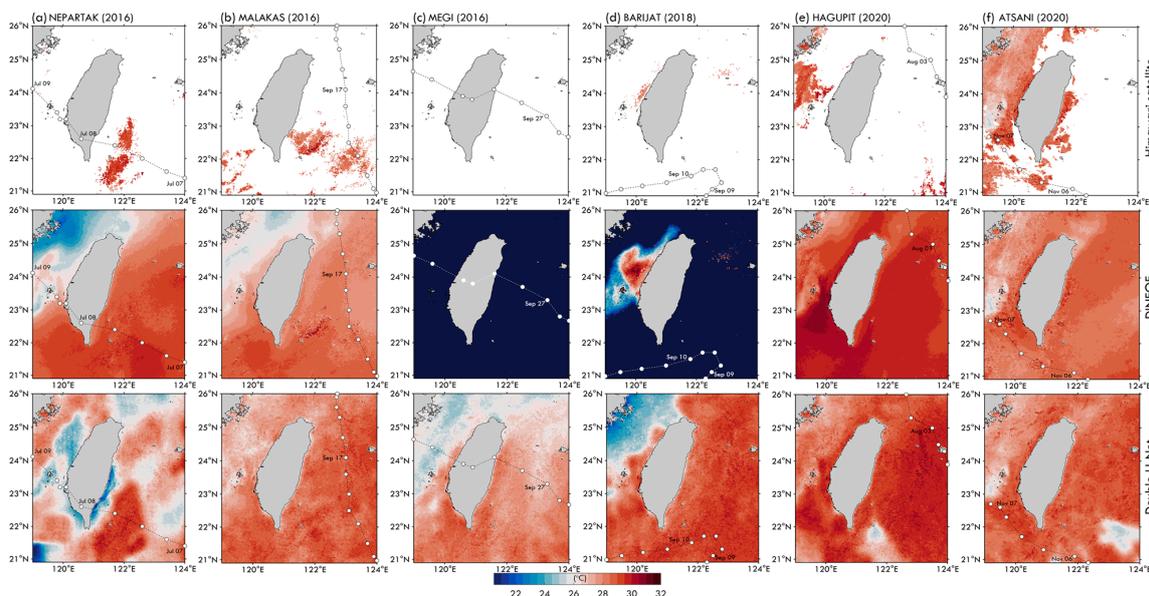


Figure 10. Examples of the spatial patterns of SST from the Himawari satellite, with gaps filled by DINEOFs and Double U-Net during typhoon occurrences. (a) 8 July 2016, NEPARTAK (2016), (b) 17 September 2016, MALAKAS (2016), (c) 27 September 2016, MEGI (2016), (d) 9 September 2018, BARIJAT (2018), (e) 3 August 2020, HAGUPIT (2020), and (f) 6 November 2020, ATSANI (2020). White dots connected by black dashed lines represent the tracks of the typhoons.

We evaluate our model's capability to capture and reconstruct SST patterns under challenging meteorological conditions. During the passage of Typhoon Megi through our study area, an evident lack of observations suggests extensive cloud cover, which obstructs satellite observations. A similar observation shortfall is noted during Typhoon Barijat, particularly in mid-western Taiwan. This situation contrasts with Typhoons Nepartak and Malakas, for which significant observations are retained, albeit with sparse SST data. The DINEOFs method proves effective for gap-filling during Nepartak and Malakas. However, it appears less adept for Typhoons Megi and Barijat, leading to underestimations and incomplete gap filling. This is particularly true for Megi, where noticeable underestimation occurs in all images, and for Barijat, where only areas with minimal observations receive adequate filling. Our U-Net model, on the other hand, excels in addressing the gaps left by Himawari SST observations, especially during the occurrences of Megi and Barijat where DINEOFs underperforms. While our model successfully fills gaps for Nepartak, the resultant values appear anomalous, indicating significant SST discrepancies in areas adjacent to the filled gaps. Moreover, the gap-filled values for Nepartak exhibit considerable differences between our U-Net model and DINEOFs. Conversely, for Malakas, the variance in filled values between DINEOFs and our U-Net model is minor. Figure 10e,f reveal a significant aspect of the algorithm, highlighting its potential to misinterpret ocean dynamics. This is demonstrated in two distinct typhoon cases: Hagupit on 3 August 2020 (Figure 10e), and Atsani on 7 November 2020 (Figure 10f). Despite over 50% cloud cover caused by typhoons, both algorithms effectively fill gaps. However, the reconstructed images, juxtaposed with the plotted locations of typhoon eyes, reveal intriguing patterns. For Hagupit, our U-Net algorithm exhibits more anomalous patterns than DINEOFs. Specifically, during Hagupit's passage through the western-central Taiwan, U-Net generates a pattern with low SST values, and an even broader, lower pattern is observed on the eastern side. The lack of sensors at these locations to validate the accuracy of these patterns leaves their representation of actual values in question. These cannot be conclusively identified as the cold wake caused by the typhoon, particularly given their distance from the typhoon's eye. During Atsani's passage, both algorithms produce cold patterns, with U-Net creating two and DINEOFs one. The low SST pattern to the west appears valid, as nearby sea regions with minimal cloud cover permit satellite observation of SST. Thus, both DINEOFs and U-Net successfully reconstruct the dataset with gaps, resulting in a continuous, low-temperature pattern. However, the eastern cold pattern generated exclusively by U-Net is questionable, situated in the open ocean without observational validation. Additionally, a small cold wake pattern near the typhoon eye on 6 November 2020, is visible in the U-Net-generated data, likely because the model used data from two days before the prediction day. This nuanced analysis underscores the algorithm's tendency to introduce unique patterns, some of which may not have empirical support, emphasizing the need for careful interpretation and validation in certain scenarios.

A critical assessment is illustrated through a scatter plot (Figure 11), where the horizontal axis represents in situ observations, and the vertical axis denotes values generated by models. Each of the 10 plots corresponds to ground-truth data from buoys, with regression lines included to evaluate the data observed by DINEOFs and our U-Net model. The results commonly show an underestimation in predicting high SST values across most plots, a phenomenon discussed in the previous section. Despite this trend, predictions from the U-Net model are notably close to the reference values, indicating a higher accuracy in capturing SST dynamics. This observation is corroborated by numerical evaluations presented in Table 3, where the R, ME, and RMSE values highlight the superior performance of the U-Net model. The ME and RMSE metrics for U-Net surpass those of DINEOFs across all locations, showing significant differences in accuracy. However, DINEOFs' predictions tend to be systematically biased toward lower SST values, indicating an underestimation. Pengjiayou (Figure 11a) serves as an exception, where DINEOFs demonstrates good estimation accuracy. Nevertheless, our U-Net model not only exceeds DINEOFs in Pengjiayou but also achieves a closer alignment of the regression line with the identity line, emphasizing

its superior accuracy and reliability in capturing SST variations. Conversely, the scatter plot for Xiaoliuqiu (Figure 11i) depicts a less favorable scenario. DINEOFs consistently underestimates SST values, leading to a regression line with a divergent slope, indicating that lower predicted SST values often correlate with higher actual values. Although our U-Net model also shows reduced accuracy in Xiaoliuqiu, it does not exhibit the same level of underestimation as DINEOFs. During typhoon periods, the Double U-Net method exhibited high correlation in SST at 7 out of 10 stations, while the DINEOFs method only did so at 2 stations (Table 3). The calculated RMSE and ME results both indicate that the Double U-Net method performs better than DINEOFs. Specifically, during typhoons, the Double U-Net method achieved an average RMSE of 1.91 °C and an ME of -1.04 °C, whereas the DINEOFs method had a significantly higher RMSE of 7.42 °C and an ME of -3.94 °C.

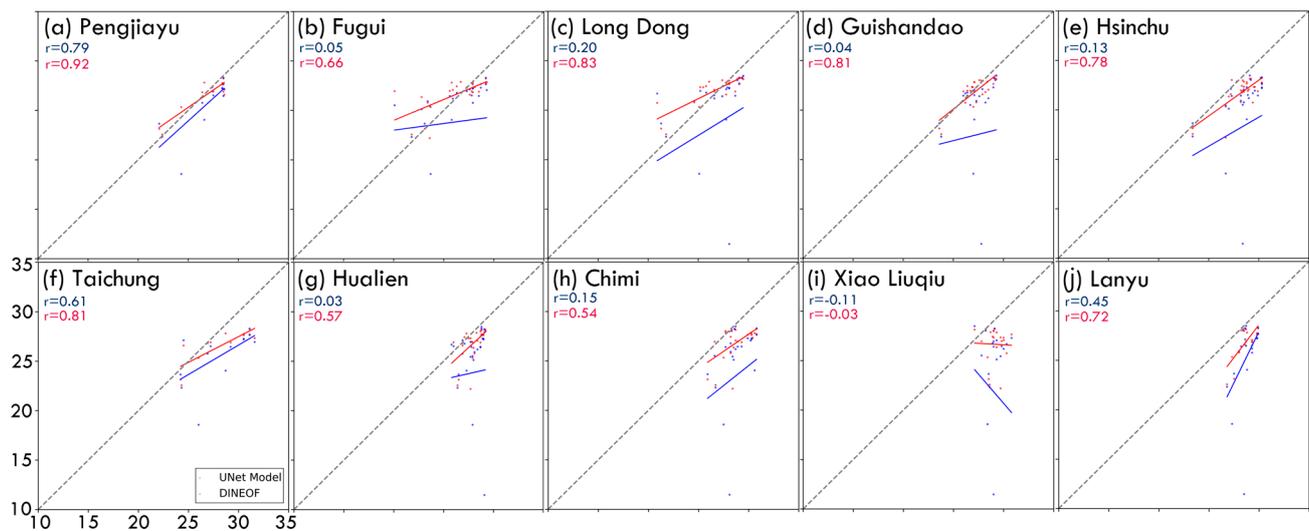


Figure 11. A comparison of SST data during the passage of a typhoon through the seas near Taiwan (see Figure 9), as calculated by the Double U-Net (red dots) and DINEOFs (blue dots) methods (shown on the vertical axis), versus buoy observation data (shown on the horizontal axis). (a) Pengjiayu, (b) Fugui, (c) Lond Dong, (d) Guishandao, (e) Hsinchu, (f) Taichung, (g) Hualien, (h) Chimi, (i) Xiao Liuqiu, and (j) Lanyu.

Table 3. The comparison results of SST calculated by the Double U-Net and DINEOFs methods with buoy observations at each station (see Figure 11) include the correlation coefficient (R), Root Mean Square Error (RMSE), and Mean Error (ME).

Buoy	Number of Samples	R		RMSE (°C)		ME (°C)	
		DINEOF	U-Net	DINEOF	U-Net	DINEOF	U-Net
Pengjiayu	13	0.79	0.92	2.02	0.98	−1.21	−0.22
Fugui	24	0.05	0.66	7.57	1.97	−2.35	0.42
Long Dong	24	0.20	0.83	8.29	1.51	−3.18	0.33
Guishandao	25	0.04	0.81	9.67	0.94	−4.75	−0.34
Hsinchu	31	0.13	0.78	9.22	1.90	−5.08	−1.56
Taichung	13	0.61	0.81	3.51	2.30	−2.61	−1.62
Hualien	29	0.03	0.57	8.23	1.76	−4.10	−1.16
Chimi	25	0.15	0.54	8.88	2.44	−5.24	−1.93
Xiao Liuqiu	22	−0.11	−0.03	11.68	3.18	−7.37	−2.45
Lanyu	19	0.45	0.72	5.08	2.15	−3.46	−1.82

3.2. Spatial Patterns of Gap-Filled SST from Daily to Monthly Scale

Due to the excessive number of missing values in the original satellite observations on the oceanic spatial grid, the key to a successful gap-filled SST lies in its ability to capture the temporal and spatial variations from daily timescales to monthly averages. Five critical

locations in the seas near Taiwan were selected to determine whether the SST calculated by the Double U-Net method accurately reflects daily SST variations. These locations include the Northern Taiwan Strait (120.3° E, 25.6° N), Northern Taiwan (121.5° E, 25.6° N), Northeastern Taiwan (122.3° E, 25.4° N), Penghu Channel (119.9° E, 23.6° N), and the KC region (122° E, 23.5° N). The Northern Taiwan Strait site, located at the southern end of the East China Sea (ECS), is dominated by the warm Taiwan Strait Current (TSC) in summer and the cold China Coastal Current and the relatively warmer Northern Taiwan Coastal Current (NTCC) in winter [41], resulting in dramatic SST changes throughout the year, with seasonal temperature variations reaching up to 15 °C. The Northern Taiwan site, along the northern coast of Taiwan, experiences daily tidal current shifts, with extensions of the northeastward TSC in summer [46] and westward inflows from the KC branch in winter [41], leading to seasonal temperature differences of 10–12 °C and significant oceanic frontal formations. The Northeastern Taiwan site is located at the boundary between the ECS shelf and the KC, experiencing cold domes due to seabed topography effects in summer [47] and KC branch intrusions into the shelf in winter [48], with seasonal temperature variations of only 5–6 °C. The Penghu Channel site, between the western coast of Taiwan and the Penghu Islands, sees SST influenced by the northwestward KC branch in winter, the northward South China Sea monsoon flow in summer, and the southward NTCC in winter [44], with significant wind direction effects, resulting in high temperatures in summer and SSTs between KC and ECS waters in winter. The KC site, located along the main axis of the KC east of Taiwan, sees temperatures of about 20–22 °C in winter and 28–30 °C in summer [41]. Figure 12 presents the daily scale SST time series filled by the Double U-Net method, calculated with an 11-day moving average. The SST time series variation characteristics align well with the results mentioned in previous studies, allowing for further geospatial information analysis to understand SST variation characteristics.

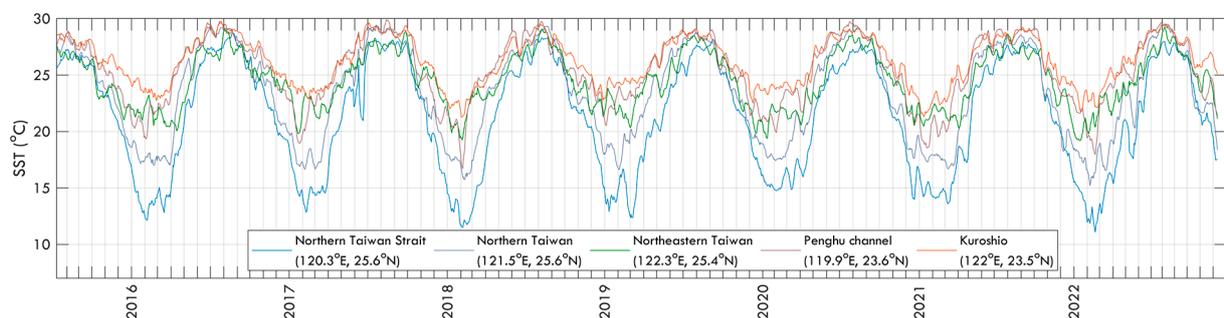


Figure 12. The daily timescale SST time series for five critical locations in the seas near Taiwan was calculated using the Double U-Net method, with the SST processed using an 11-day running mean.

Figure 13 shows the monthly average variations at the aforementioned five characteristic locations, with the largest temperature difference occurring in the northern Taiwan Strait region, where SST ranges from 13 to 27 °C, and the smallest difference in the KC region, with SST ranging from 23 to 28 °C. The SSTs of other sea areas fluctuate within the temperature ranges of these two regions. Figure 14 presents the spatial features of SST data filled by the Double U-Net method, spanning from July 2015 to December 2022. Monthly SST exhibits excellent temperature dynamics, including oceanic fronts at the boundary between the KC and the ECS shelf, strong mesoscale oceanic fronts caused by currents off the northern coast of Taiwan, summer upwelling in the Taiwan Bank due to topographic effects, the boundary between I-Lan Bay and the western edge of the KC, the path of the KC branch intrusion into the Penghu Channel in winter, mesoscale cold domes off the northeastern corner of Taiwan in summer, and even the sub-mesoscale cold wake waters of Green Island are visible. The results combined in Figures 12 and 14 underscore that using machine learning's Double U-Net to fill missing spatial grid SST data can effectively analyze daily timescale variations and trends in monthly average spatial patterns, and these results have high spatial resolution.

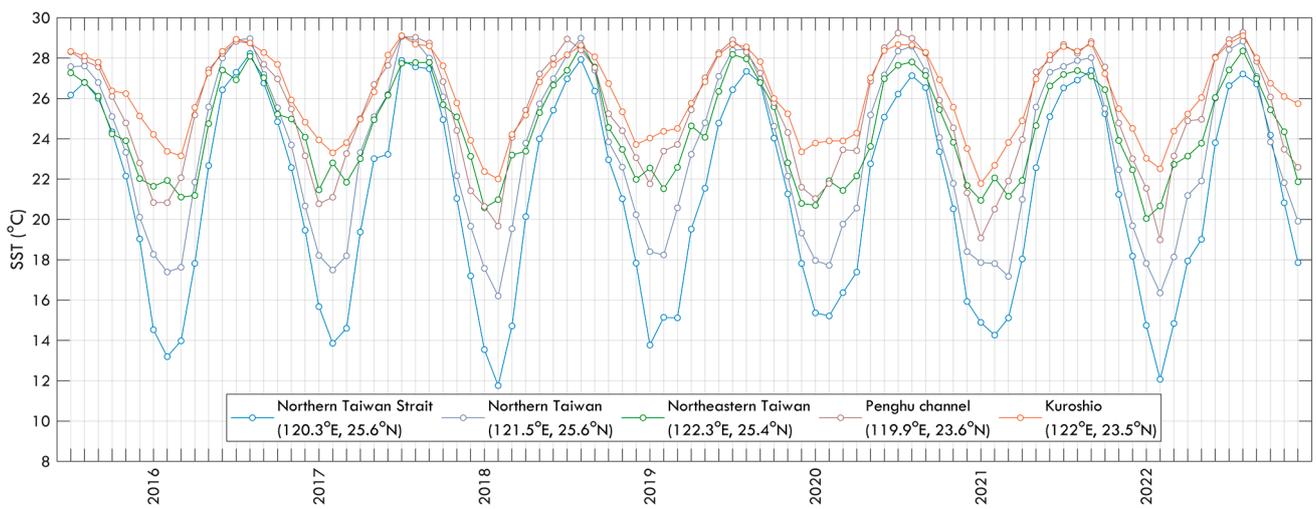


Figure 13. The monthly timescale SST time series for five critical locations in the seas near Taiwan was calculated using the Double U-Net method.

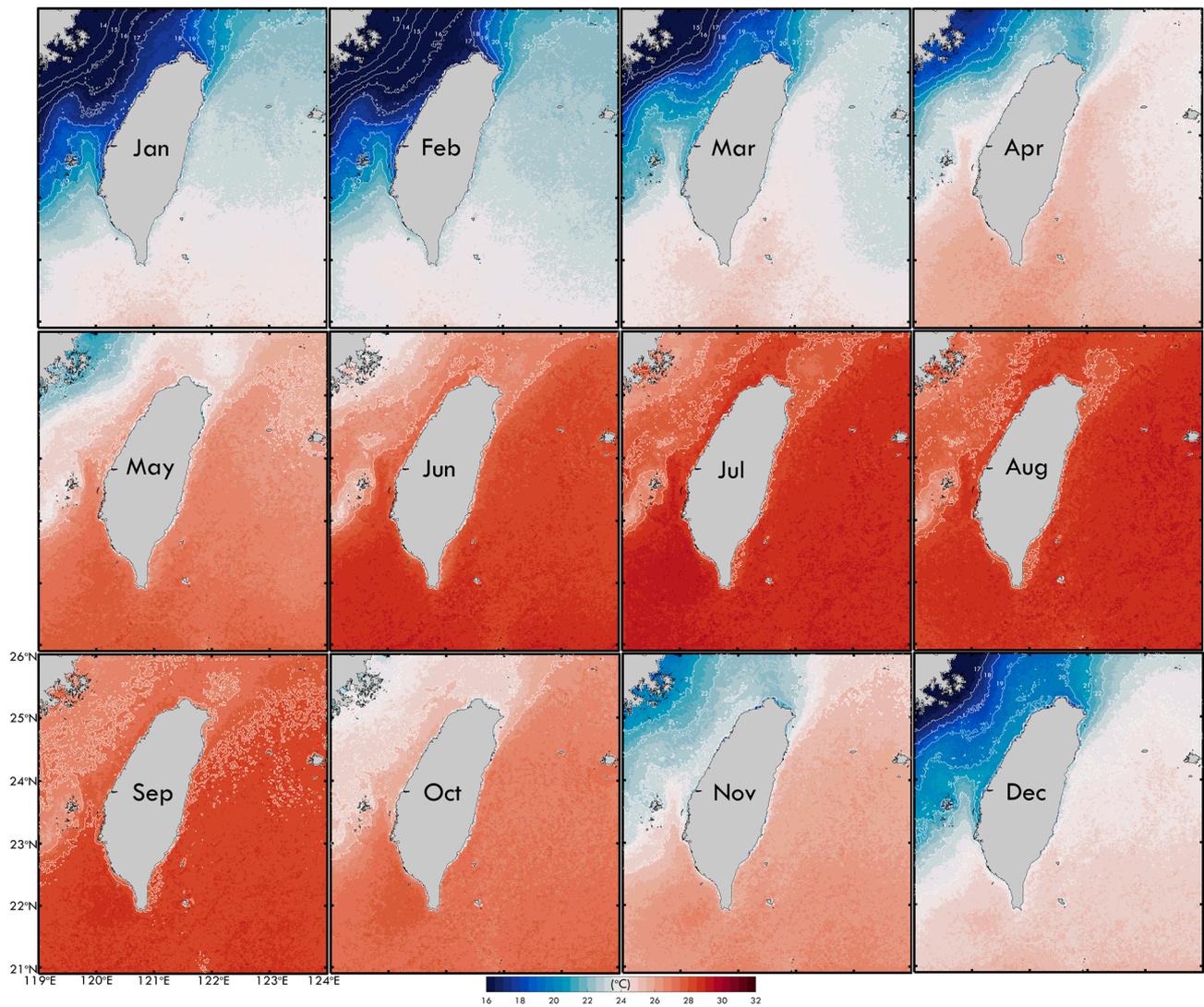


Figure 14. Spatial distribution features of SST calculated by the Double U-Net, covering the period from July 2015 to December 2022. White solid lines represent isotherms.

4. Discussion

The development of advanced gap-filling techniques is crucial for creating complete SST datasets that accurately represent the intricate ocean dynamics across temporal and spatial scales. Our U-Net model has shown several advantages over prevailing methods like DINEOFs in reconstructing missing satellite SST data. A key advantage of our model is its ability to capture finer-scale SST features, outperforming DINEOFs in this aspect. As illustrated in Figure 7, the U-Net prediction reveals a distinctly sharper KC front, in contrast to the more diffused representation offered by DINEOFs. This enhanced capability is likely due to the support from the L4 Himawari dataset, contributing to the U-Net's superior performance. The accurate depiction of major ocean currents such as the KC is pivotal for a comprehensive analysis of their impact on both regional and global climates. The KC is known for its high spatiotemporal SST variability, driven by meandering and eddy formations that result from complex bathymetric interactions and instabilities. Our U-Net model's flexible nonlinear framework is effectively able to learn these nuanced dynamics from the training data, unlike DINEOFs' linear empirical orthogonal function approach, which tends to smooth over intricate signals induced by currents. Resolving the sharp KC front and associated mesoscale features is also critical for applications in ocean modeling and ship navigation that rely on the accurate resolution of strong SST gradients.

The U-Net model, while demonstrating considerable advantages, also produced potentially spurious lower SSTs in data-sparse areas near 23° N and 123° E, which lack in situ validation. This issue underscores the risk of overfitting that deep neural networks face when exploiting wider contextual patterns in regions with limited constraints. Achieving a balance between higher resolution reconstruction and avoiding unphysical artifacts remains a significant research challenge. Hybrid techniques, which combine physics-based regularization with data-driven modeling, could offer a solution to this trade-off. Despite the U-Net model's advantages, DINEOFs displayed superior accuracy in open ocean regions, particularly in areas like Lanyu, as evidenced by correlation metrics. This suggests that empirical orthogonal functions may be more adept at capturing mesoscale SST signatures, especially in environments influenced by less complex bathymetric interactions and without internal tides. The data-driven U-Net model relies heavily on representative training data that cover diverse oceanic conditions. A lack of sufficient samples from specific areas, such as Lanyu, may limit the model's ability to generalize. Expanding in situ sensor coverage to enrich the training dataset, along with comprehensive validation, is vital for ensuring robust global reconstructions by the U-Net model. Incorporating additional satellite measurements could also help mitigate observational gaps caused by factors like cloud cover, which intermittently limit Himawari's visibility. Integrating microwave and infrared SST data using the U-Net model has the potential to create complete high-resolution maps by synergistically blending diverse datasets. No single technique universally outperforms others across all oceanic conditions. Hybrid approaches that combine multivariate optimization, empirical orthogonal functions, and deep neural networks could optimally harness their respective strengths for SST reconstruction. For example, integrating physics-based constraints can regularize the U-Net model to avoid unphysical artifacts, while its nonlinear flexibility enhances representations of major currents and eddies that DINEOFs might miss. Developing integrated frameworks represents an important ongoing area of research.

The U-Net model boasts a significant advantage with its robust performance in filling extensive SST data gaps during typhoons, which are characterized by heavy cloud cover constraints. As illustrated in Figure 10, DINEOFs encountered difficulties in reconstructing SSTs during Typhoons Megi and Barijat, whereas the U-Net generated complete maps that aligned well with available observations. Accurately capturing typhoon-induced ocean dynamics is crucial for analyzing their impacts on vertical mixing, upper ocean heat content, and the sea state. The U-Net's data-driven framework facilitates the modeling of complex physical interactions between storms and the ocean. However, the U-Net's response to typhoons revealed some inconsistencies. Specifically, during Typhoon Nepar-

tak, it produced anomalous SST discontinuities between neighboring regions. Moreover, the predictions from U-Net and DINEOFs diverged significantly for Nepartak compared to other storms, underscoring the challenges of extrapolating beyond observations under certain extreme conditions. Introducing physical constraints and regularization methods would aid in preventing unphysical artifacts in the reconstructed SST fields. Additionally, both models produced questionable remote cold wakes without observational validation, as observed in Figure 10. Distinguishing true cold wakes from spurious artifacts remains a critical area of ongoing research. Implementing statistical flagging of potential model errors, informed by storm positions, could help mitigate this issue. Overall, the U-Net model demonstrates considerable potential in reconstructing SSTs under the meteorologically challenging conditions presented by typhoons. With further refinement, guided by oceanographic principles, it can offer a robust gap-filling technique for enabling more reliable SST analysis on a global scale.

5. Conclusions

This study introduces a Double U-Net model specifically designed for filling data gaps in SST observations and conducts a comparative performance analysis with the widely used DINEOFs method. Utilizing three consecutive days of Himawari satellite observation data, both methods successfully generated complete SST maps by effectively filling in the gaps. However, the Double U-Net model demonstrated several key advantages, notably its ability to capture a broader range of SST dynamics and provide more detailed spatial patterns. This advantage was particularly evident in its representation of the KC, where the U-Net model produced sharper and more distinct images compared to the more interpolated appearance of DINEOFs. Comparative analysis with buoy observations revealed the Double U-Net model's superior correlation results and lower RMSE and ME values, underscoring its enhanced accuracy over DINEOFs in most study areas. Despite the inherent challenges posed by the differences between satellite-observed skin SST and buoy-measured bulk SST, the U-Net model's performance marks a significant advancement in the methodology for filling SST data gaps, contributing to a more accurate and comprehensive understanding of ocean temperature dynamics.

Through detailed comparisons and visual analyses of typhoon events, the Double U-Net model has shown superior performance in capturing complex SST variations and accurately filling data gaps more effectively than DINEOFs. It highlights the U-Net model's enhanced ability to fill SST data gaps under challenging conditions, such as during typhoons, outperforming the DINEOFs method. However, it also points out the U-Net model's potential for generating anomalous patterns or underestimations, emphasizing the importance of careful validation and interpretation of the reconstructed SST data.

The study demonstrates the effectiveness of the Double U-Net method in accurately filling data gaps in SST observations across critical locations near Taiwan, effectively capturing both daily and monthly SST variations. By analyzing SST dynamics at five key locations, including the Northern Taiwan Strait, Northern Taiwan, Northeastern Taiwan, Penghu Channel, and the KC region, the research highlights the method's ability to reflect complex oceanographic features and temperature dynamics, such as oceanic fronts, mesoscale variations, and upwelling effects. The findings, illustrated through detailed SST time series and spatial feature analyses, affirm the Double U-Net's potential for high-resolution SST data reconstruction, offering valuable insights into understanding and analyzing SST variation characteristics with enhanced accuracy and spatial resolution. Nonetheless, further investigations and refinements are needed to address the model's limitations and improve its performance in specific areas.

Author Contributions: Conceptualization, Dimas Pradana Putra; methodology, Dimas Pradana Putra; software, Dimas Pradana Putra; validation, Dimas Pradana Putra and Po-Chun Hsu; formal analysis, Dimas Pradana Putra and Po-Chun Hsu; investigation, Dimas Pradana Putra and Po-Chun Hsu; resources, Dimas Pradana Putra and Po-Chun Hsu; data curation, Dimas Pradana Putra and Po-Chun Hsu; writing—original draft preparation, Dimas Pradana Putra; writing—review and editing, Po-Chun Hsu; visualization, Dimas Pradana Putra and Po-Chun Hsu; project administration, Po-Chun Hsu; funding acquisition, Po-Chun Hsu. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science and Technology Council, grant number 111-2611-M-008-007 and 112-2621-M-008-002.

Data Availability Statement: The Himawari satellite-derived SST data were obtained from the JAXA p-tree system (<https://www.eorc.jaxa.jp/ptree/index.html>, accessed on 20 May 2023). The typhoon dataset was obtained from the IBTrACS (<http://www.ncdc.noaa.gov/ibtracs/>, accessed on 20 May 2023). The CWA buoy dataset was obtained from the Safe Ocean (<https://ocean.cwb.gov.tw/V2/>, accessed on 20 May 2023). The code used in this research is available at <https://github.com/Arcturion/UNET-gap-filling/> (accessed on 5 March 2024).

Acknowledgments: The authors appreciate all the data used, provided from open databases. The authors thank anonymous reviewers and academic editors for their comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Merchant, C.J.; Block, T.; Corlett, G.K.; Embury, O.; Mittaz, J.P.; Mollard, J.D. Harmonization of space-borne infra-red sensors measuring sea surface temperature. *Remote Sens.* **2020**, *12*, 1048. [CrossRef]
2. Minnett, P.J.; Alvera-Azcárate, A.; Chin, T.M.; Corlett, G.K.; Gentemann, C.L.; Karagali, I.; Li, X.; Marsouin, A.; Marullo, S.; Maturi, E.; et al. Half a century of satellite remote sensing of sea-surface temperature. *Remote Sens. Environ.* **2019**, *233*, 111366. [CrossRef]
3. Jia, X.; Ji, Q.; Han, L.; Liu, Y.; Han, G.; Lin, X. Prediction of sea surface temperature in the East China Sea based on LSTM neural network. *Remote Sens.* **2022**, *14*, 3300. [CrossRef]
4. Zhang, R. On the persistence and coherence of subpolar sea surface temperature and salinity anomalies associated with the Atlantic multidecadal variability. *Geophys. Res. Lett.* **2017**, *44*, 7865–7875. [CrossRef]
5. Long, A.P.; Haberlin, D.; Lyashevskaya, O.; Brophy, D.; O’Hea, B.; O’Donnell, C.; Scarrott, R.G.; Lawton, C.; Doyle, T.K. Interannual variability of gelatinous mesozooplankton in a temperate shelf sea: Greater abundance coincides with cooler sea surface temperatures. *ICES J. Mar. Sci.* **2021**, *78*, 1372–1385. [CrossRef]
6. Callejas, I.A.; Osborn, K.; Lee, C.; Mishra, D.R.; Auil Gomez, N.; Carrias, A.; Cherrington, E.A.; Griffin, R.; Rosado, A.; Rosado, S.; et al. A GEE toolkit for water quality monitoring from 2002 to 2022 in support of SDG 14 and coral health in marine protected areas in Belize. *Front. Remote Sens.* **2022**, *3*, 1020184. [CrossRef]
7. Quiring, K.; Carroll, G.; Champion, C.; Heymann, E.W.; Harcourt, R. The diet of greater crested terns off southeast Australia varies with breeding stage and sea surface temperature. *Mar. Biol.* **2021**, *168*, 143. [CrossRef]
8. De, K.; Nanajkar, M.; Arora, M.; Nithyanandan, M.; Mote, S.; Ingole, B. Application of remotely sensed sea surface temperature for assessment of recurrent coral bleaching (2014–2019) impact on a marginal coral ecosystem. *Geocarto Int.* **2022**, *37*, 4483–4508. [CrossRef]
9. Quartly, G.D.; Srokosz, M.A. SST observations of the Agulhas and East Madagascar retroreflections by the TRMM Microwave Imager. *J. Phys. Oceanogr.* **2002**, *32*, 1585–1592. [CrossRef]
10. Schumann, G.J.P. Grand challenges in microwave remote sensing. *Front. Remote Sens.* **2020**, *1*, 603650. [CrossRef]
11. Mavrovic, A.; Sonnentag, O.; Lemmetyinen, J.; Baltzer, J.L.; Kinnard, C.; Roy, A. Reviews and syntheses: Recent advances in microwave remote sensing in support of terrestrial carbon cycle science in Arctic–boreal regions. *Biogeosciences* **2023**, *20*, 2941–2970. [CrossRef]
12. Wick, G.A.; Bates, J.J.; Scott, D.J. Satellite and skin-layer effects on the accuracy of sea surface temperature measurements from the GOES satellites. *J. Atmos. Ocean. Technol.* **2002**, *19*, 1834–1848. [CrossRef]
13. Merchant, C.J.; Embury, O.; Bulgin, C.E.; Block, T.; Corlett, G.K.; Fiedler, E.; Good, S.A.; Mittaz, J.; Rayner, N.A.; Berry, D.; et al. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data* **2019**, *6*, 223. [CrossRef] [PubMed]
14. Cao, M.; Mao, K.; Yan, Y.; Shi, J.; Wang, H.; Xu, T.; Fang, S.; Yuan, Z. A new global gridded sea surface temperature data product based on multisource data. *Earth Syst. Sci. Data* **2021**, *13*, 2111–2134. [CrossRef]
15. Jääskeläinen, E.; Manninen, T.; Hakkarainen, J.; Tamminen, J. Filling gaps of black-sky surface albedo of the Arctic sea ice using gradient boosting and brightness temperature data. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102701. [CrossRef]
16. Jung, S.; Yoo, C.; Im, J. High-resolution seamless daily sea surface temperature based on satellite data fusion and machine learning over Kuroshio Extension. *Remote Sens.* **2022**, *14*, 575. [CrossRef]

17. Davis, L.L.; Thompson, D.W.; Kennedy, J.J.; Kent, E.C. The importance of unresolved biases in twentieth-century sea surface temperature observations. *Bull. Am. Meteorol. Soc.* **2019**, *100*, 621–629. [[CrossRef](#)]
18. O'carroll, A.G.; Armstrong, E.M.; Beggs, H.M.; Bouali, M.; Casey, K.S.; Corlett, G.K.; Dash, P.; Donlon, C.J.; Gentemann, C.L.; Høyer, J.L.; et al. Observational needs of sea surface temperature. *Front. Mar. Sci.* **2019**, *6*, 420. [[CrossRef](#)]
19. Parker, D.E.; Folland, C.K.; Jackson, M. Marine surface temperature: Observed variations and data requirements. *Clim. Chang.* **1995**, *31*, 559–600. [[CrossRef](#)]
20. Sukhotin, A.; Berger, V. Long-term monitoring studies as a powerful tool in marine ecosystem research. *Hydrobiologia* **2013**, *706*, 1–9. [[CrossRef](#)]
21. Taylor, S.P.; Haywood, A.M.; Valdes, P.J.; Sellwood, B.W. An evaluation of two spatial interpolation techniques in global sea-surface temperature reconstructions: Last Glacial Maximum and Pliocene case studies. *Quat. Sci. Rev.* **2004**, *23*, 1041–1051. [[CrossRef](#)]
22. Kusuma, D.W.; Murdimanto, A.; Sukresno, B.; Jatisworo, D. Comparison of interpolation methods for sea surface temperature data. *J. Fish. Mar. Res.* **2018**, *2*, 103–115. [[CrossRef](#)]
23. Zhang, Y.; Feng, M.; Zhang, W.; Wang, H.; Wang, P. A Gaussian process regression-based sea surface temperature interpolation algorithm. *J. Oceanol. Limnol.* **2021**, *39*, 1211–1221. [[CrossRef](#)]
24. Urquhart, E.A.; Hoffman, M.J.; Murphy, R.R.; Zaitchik, B.F. Geospatial interpolation of MODIS-derived salinity and temperature in the Chesapeake Bay. *Remote Sens. Environ.* **2013**, *135*, 167–177. [[CrossRef](#)]
25. Schneider, T. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Clim.* **2001**, *14*, 853–871. [[CrossRef](#)]
26. Beckers, J.M.; Rixen, M. EOF calculations and data filling from incomplete oceanographic datasets. *J. Atmos. Ocean. Technol.* **2003**, *20*, 1839–1856. [[CrossRef](#)]
27. Beckers, J.M.; Barth, A.; Alvera-Azcárate, A. DINEOF reconstruction of clouded images including error maps—application to the Sea-Surface Temperature around Corsican Island. *Ocean. Sci.* **2006**, *2*, 183–199. [[CrossRef](#)]
28. Alvera-Azcárate, A.; Barth, A.; Sirjacobs, D.; Beckers, J.M. Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF. *Ocean. Sci.* **2009**, *5*, 475–485. [[CrossRef](#)]
29. Jouini, M.; Lévy, M.; Crépon, M.; Thiria, S. Reconstruction of satellite chlorophyll images under heavy cloud coverage using a neural classification method. *Remote Sens. Environ.* **2013**, *131*, 232–246. [[CrossRef](#)]
30. Chen, S.; Hu, C.; Barnes, B.B.; Xie, Y.; Lin, G.; Qiu, Z. Improving ocean color data coverage through machine learning. *Remote Sens. Environ.* **2019**, *222*, 286–302. [[CrossRef](#)]
31. Stock, A.; Subramaniam, A.; Van Dijken, G.L.; Wedding, L.M.; Arrigo, K.R.; Mills, M.M.; Cameron, M.A.; Micheli, F. Comparison of cloud-filling algorithms for marine satellite data. *Remote Sens.* **2020**, *12*, 3313. [[CrossRef](#)]
32. Krasnopolsky, V.; Nadiga, S.; Mehra, A.; Bayler, E.; Behringer, D. Neural networks technique for filling gaps in satellite measurements: Application to ocean color observations. *Comput. Intell. Neurosci.* **2016**, *2016*, 6156513. [[CrossRef](#)] [[PubMed](#)]
33. Konik, M.; Kowalewski, M.; Bradtke, K.; Darecki, M. The operational method of filling information gaps in satellite imagery using numerical models. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *75*, 68–82. [[CrossRef](#)]
34. Buttler, J.V.; Zscheischler, J.; Mahecha, M.D. An extended approach for spatiotemporal gapfilling: Dealing with large and systematic gaps in geoscientific datasets. *Nonlinear Process. Geophys.* **2014**, *21*, 203–215. [[CrossRef](#)]
35. Yang, J.; Huo, J.; He, J.; Xiao, T.; Chen, D.; Li, Y. A DBULSTM-Adaboost Model for Sea Surface Temperature Prediction. *PeerJ Comput. Sci.* **2022**, *8*, e1095. [[CrossRef](#)]
36. Chan, D. Combining Statistical, Physical, and Historical Evidence to Improve Historical Sea-Surface Temperature Records. *Harv. Data Sci. Rev.* **2021**, *3*. [[CrossRef](#)]
37. Barth, A.; Alvera-Azcárate, A.; Licer, M.; Beckers, J.M. DINCAE 1.0: A convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations. *Geosci. Model Dev.* **2020**, *13*, 1609–1622. [[CrossRef](#)]
38. Patil, K.R.; Iiyama, M. Deep learning models to predict Sea surface temperature in tohoku region. *IEEE Access* **2022**, *10*, 40410–40418. [[CrossRef](#)]
39. Baker, S.; Huang, Z.; Philippa, B. Lightweight Neural Network For Spatiotemporal Filling of Data Gaps in Sea Surface Temperature Images. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4204310. [[CrossRef](#)]
40. Chai, X.; Gu, H.; Li, F.; Duan, H.; Hu, X.; Lin, K. Deep learning for irregularly and regularly missing data reconstruction. *Sci. Rep.* **2020**, *10*, 3302. [[CrossRef](#)]
41. Hsu, P.C.; Centurioni, L.; Shao, H.J.; Zheng, Q.; Lu, C.Y.; Hsu, T.W.; Tseng, R.S. Surface current variations and oceanic fronts in the southern East China Sea: Drifter experiments, coastal radar applications, and satellite observations. *J. Geophys. Res. Ocean.* **2021**, *126*, e2021JC017373. [[CrossRef](#)]
42. Hsu, P.C.; Lee, H.J.; Lu, C.Y. Impacts of the Kuroshio and Tidal Currents on the Hydrological Characteristics of Yilan Bay, Northeastern Taiwan. *Remote Sens.* **2021**, *13*, 4340. [[CrossRef](#)]
43. Hsu, P.C. Surface Current Variations and Hydrological Characteristics of the Penghu Channel in the Southeastern Taiwan Strait. *Remote Sens.* **2022**, *14*, 1816. [[CrossRef](#)]
44. Knapp, K.R.; Kruk, M.C.; Levinson, D.H.; Diamond, H.J.; Neumann, C.J. The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. *Bull. Am. Meteorol. Soc.* **2010**, *91*, 363–376. [[CrossRef](#)]

45. Hsu, P.C. Evaluation of Wind and Solar Insolation Influence on Ocean Near-Surface Temperature from In Situ Observations and the Geostationary Himawari-8 Satellite. *Remote Sens.* **2022**, *14*, 4975. [[CrossRef](#)]
46. Lu, C.Y.; Hsu, P.C.; Zheng, Q.; Ho, C.R. Variations in flow patterns in the northern Taiwan Strait observed by satellite-tracked drifters. *Remote Sens.* **2022**, *14*, 2154. [[CrossRef](#)]
47. Hsu, P.C.; Zheng, Q.; Lu, C.Y.; Cheng, K.H.; Lee, H.J.; Ho, C.R. Interaction of coastal countercurrent in I-Lan Bay with the Kuroshio northeast of Taiwan. *Cont. Shelf Res.* **2018**, *171*, 30–41. [[CrossRef](#)]
48. Hsu, P.C.; Macagga RA, T.; Lu, C.Y.; Lo DY, J. Investigation of the Kuroshio-Coastal Current Interaction and Marine Heatwave Trends in the Coral Habitats of Northeastern Taiwan. *Reg. Stud. Mar. Sci.* **2024**, *71*, 103431. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.