International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning Methods Based on Storm Surge Disaster Loss in Computing Applications

M. Mohammed Thaha[1], Izwan Nizal Mohd Shaharanee[2], V. P. Murugan[3], T. P. S. Kumar Kusumanchi[4], D. V. Lokeswar Reddy[5]

Submitted: 16/09/2023 **Revised:** 23/10/2023 **Accepted:** 15/11/2023

Abstract: Storm surge, which impacts the entire coastline region, is China's most serious marine calamity. Storm surge disaster loss (SSDL) estimation is important for decision-making, sustainability, and disaster prevention. For early warning systems, disaster management, and disaster evaluation, an accurate storm surge water level forecast is essential. Comparing machine learning techniques to numerical simulation techniques, the former is more straightforward and efficient. Still, point predictions are the main focus of the majority of current machine learning-based research on storm surge prediction. In this paper, we explore the feasibility of employing the ConvLSTM model for spatial water level prediction. The ConvLSTM-based methodology is simpler, faster, and more accurate in predicting water levels without the need for boundary conditions or topography than standard numerical simulation methods. In addition, we take worst-case situations into account by employing the random forest model to anticipate the highest possible water increase. Based on our findings, the random forest model may prove to be a useful instrument in determining the highest water increase value linked to typhoon storm surges, which can help with efficient emergency reactions during natural disasters.

Keywords: *Storm surge, ConvLSTM, natural disasters, efficient, machine learning, water level*

1. Introduction

The unusual fluctuation in the level of saltwater brought on by a strong atmospheric disturbance is known as a storm surge, and it is a devastating natural occurrence. Typhoons and other disastrous weather phenomena, along with extratropical cyclones, are typically referred to as atmospheric disturbances. Storm surge catastrophes present a serious risk to coastal communities' property and people's safety as well as their infrastructure [1]. In addition, the storm surge manifests itself as a succession of disasters. In addition to storm surges, nearshore waves, astronomical tides, and the effect of coupling are also considered to be risk-formative components. Storm surges and the secondary risks they pose to coastal communities worldwide are estimated to cause injury to 45 million people annually. Thus, it is crucial to look into the SSDL estimate.

Regression and data mining have seen good results from ML algorithms, making it a viable solution. They are more controllable and extensible, allowing them to properly capture the nonlinear relationship. Natural catastrophe domains including landslides, floods, and forest fires have demonstrated the potential of artificial intelligence techniques. Research on SSDL estimates is, nonetheless, scarce. Utilizing the SVM model, SSDL was evaluated and predicted based on the minimal sample data. On the other hand, not much study has been done on SSDL estimates. This chapter evaluated and predicted SSDL using the SVM model based on small sample data. The number of systems with high fitting has grown with study, but choosing the right model remains hard. It is therefore necessary to continuously verify and evaluate the development and optimization of SSDL estimate models. For instance, this chapter used five models—BPNN, one-dimensional convolutional neural networks, decision trees (DT), Random Forests (RF), and eXtreme

[1]Assistant Professor (Sr.Grade), B.S.Abdur Rahman Crescent Institute of Science and Technology, GST Road, Vandalur, Chennai - 600 048, Tamilnadu, INDIA
Email: mohammedthaha@crescent.education
[2]School of Quantitative Sciences, Universiti Utara Malaysia, 06010UUM Sintok, Kedah, Malaysia.
Email: nizal@uum.edu.my
[3]Assistant Professor, Department of Mathematics, Panimalar Engineering College, No.391, Bangalore Trunk Road, Poonamallee, Varadarajapuram, Tamilnadu 600123.
Email: vpmurugan07@gmail.com
[4]Assistant professor, Department of IoT, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522302
Email: satishkumar8421@gmail.com
[5]Assistant Professor, Humanities and Social Sciences Department, JNTU College Of Engineering, Pulivendula-516390,
Kadapa (D), Andhra Pradesh, India

Gradient Boosting (XGBoost)—to fit the information and choose the most appropriate model for the association between debris flow incidents and triggering factors. There is potential for ML to estimate SSDL.

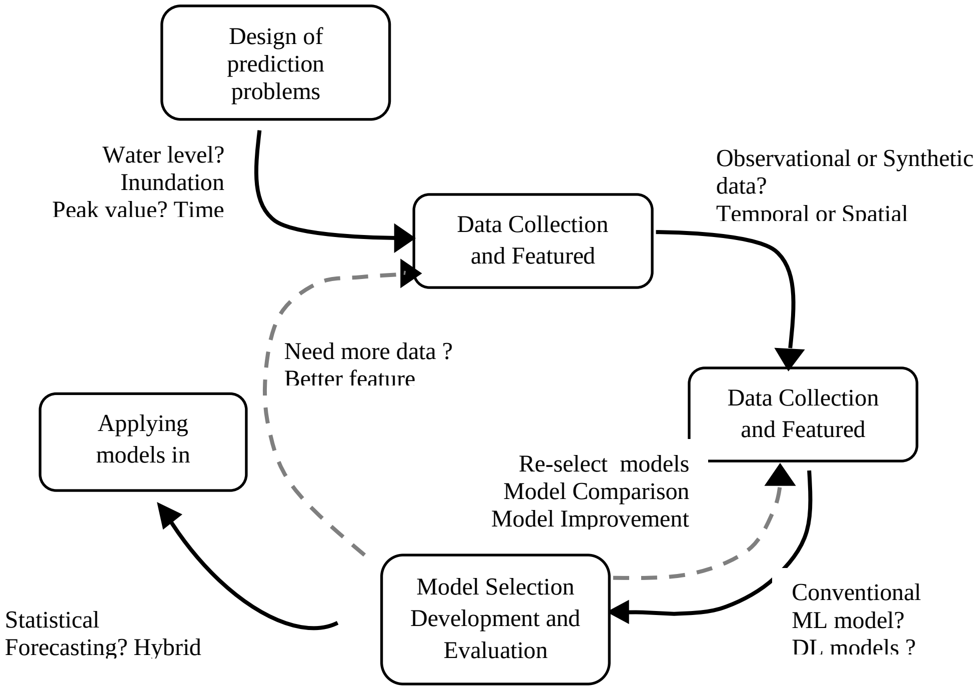

Fig. 1.1. An illustration of the process for using machine learning to create a storm surge forecast product

A flowchart of the process for using ML to create a storm surge forecast product may be found in Figure 1.1. "Formulation of prediction problems" is where the work begins, and it follows the filled-in arrows to advance step by step. Usually iterative, the process entails regular feedback and modifications to increase the product's performance and value [2]. When outcomes are not satisfactory, dot-pointed arrows indicate going back to the previous steps.

SVM can determine the global best solution by resolving a quadratic programming problem. On the other hand, more samples lead to a higher degree of complexity [3]. SVM offers significant benefits for addressing issues with low sample numbers, nonlinear sample data, and high sample dimensions. SVM provides good generalization capabilities in addition to great learning capabilities.

The rest of this work proceeds according to standard protocol, which involves building an ML model appropriate for a certain storm surge situation. Alternative methods for articulating problems related to coastal inundation and storm surge prediction are covered in Section 2. Section 3 describes the machine learning model's training data as well as certain requirements for producing a high-quality feature representation. In Section 4, the basic principles of machine learning are discussed along with a summary of frequently used methods and error measures for model evaluation. The article is concluded in Section 5 with some suggestions for further development.

2. Literature Review

Du, X., Li, X., et.al [4] Academics have investigated a wide range of loss estimating techniques for storm surges and maritime disasters. In general, there are two categories of estimating models: Models for machine learning (MLMs) and physical models. Physical models based on the inundation zone, submerged depth, flooded elements, and extent to which exposed portions produce direct tangible damage are used to calculate the economic losses of TSSDs. The model known as Sea, Lake, and Overland Surges from Hurricanes (SLOSH) for example, uses information from coastal sounding elements and typhoon variables to estimate storm surge heights.

Devaraj, J.,et.al [5] At the individual, corporate, and societal levels, among others, hurricane forecasting aids in preventing physical harm. Establishing precise measurements of tropical cyclone intensity could aid in the provision of early warning. To accurately predict the hurricane, the model's initial condition input must be of high quality. More information can be obtained by forecast users, suppliers, and policymakers by evaluating the trade-offs between various aspects. Here, we report an enhanced CNN model for precise tropical cyclone intensity and category estimation.

Fang, W., et.al [6] In order to anticipate water levels at several sites around the Pearl River Estuary in China, in this work, two-dimensional wind field features are directly extracted using a CNN and combined with tide level time data. This enables the study to leverage the strengths of BP and RNN in forecasting and CNN in TC feature extraction. The scientists also discovered that while performance was enhanced by including prediction and storm surge forecasts, this came at the expense of longer computation times.

Kolla, V. R. K. et.al [7] Weather map predicting techniques have been around for more than a century, ever before the invention of the telegraph. In order to analyse weather maps, the National Meteorological Centre can collect data in a timely manner. It appears like the low-pressure system is moving based on the weather chart. Our ability to study the weather system and use our understanding of it to estimate its future motion and intensity changes is the major basis of the weather map forecasting approach. To rapidly pinpoint the weather system causing variations in the local weather, we evaluate weather maps or other supplementary maps. The foundation of the weather map forecasting method is the accuracy of the weather map analysis.

Ian, V. K. et.al [8] Machine learning has the potential to enhance preparedness and response for extreme weather events like storm surges. The inability of current machine learning techniques to forecast storm surges accurately in the event of unexpected weather or abrupt changes in weather patterns stems from their reliance on past data for training. This leaves a research vacuum in the area of precisely predicting sea level anomalies during abrupt changes in weather patterns, like an increase in wind speed or a decrease in air pressure. In particular, when abrupt weather changes are detected, these models frequently generate erroneous tidal level estimates when tropical cyclones are present.

Chen, R., et.al [9] Depending on the type of learning task, machine learning algorithms, such as those for dimensionality reduction, feature selection, and prediction, can also be categorized. Only predictive algorithms will be discussed here because the main focus of this review is TC forecast modeling. Classification is the term used to describe a learning problem where the model's objective is to predict discrete values; regression is the term used to describe a learning job where the model's goal is to predict continuous values.

Snaiki, R., et.al [10] These studies offer a wealth of climate data that can be used as an input for estimating hurricane risk. Most climate models predict considerable variations in a number of environmental characteristics, including moisture content, warmth at the tropopause level, and environmental vertical wind shear. However, the results of these simulations can differ significantly due to intrinsic errors and model variances. Wind engineering academics predict that the maximum hurricane wind speeds will rise significantly as a result of global warming.

Yi, X., et.al [11] Indicators pertinent to eleven Chinese coastal provinces and cities spanning the years 2007–2016 were gathered for this research. Depending on the examination of significant positive and negative indicators, the entire storm surge disaster loss index is ascertained. While innovation in green marine science and technology acts as the adjustment variable, the rate of economic growth is the primary explanatory component. For a ten-year period, from 2007 to 2016, the aim of this study is to investigate the association between the amount of coastal economic development and storm surge disaster losses.

3. Methods and Materials

The choice of data representation (or features) has a major effect on how well machine learning models work because different representations might entangle and obscure the underlying explanatory variables that underlie the data [12]. When algorithms are just getting close to reaching their limit, data, and features have the greatest impact on machine learning projects and define the upper bound on job performance.

Data preparations

Samples for an ML job are usually separated into three categories: test set, validation set, as well as training set. The model is trained using the training set; the parameters that yield the "best" model performance are found using the validation set; and the test set is utilized to confirm that the optimal model that has been chosen performs as expected. Of the samples, 25% are kept as a test set. The validation and train sets are included in the CV set.

This document compiles 132 typhoon storm surge catastrophe incidents. The incidents, which cover 21 years from 2000 to 2020 and include 50 indicators, are spread out among 11 coastal province administrative zones of China. To perform data comparison and machine learning scientifically, preprocessing is required due to the broad spatiotemporal range of this data set. The next sections have the data preparations presented.

Modification of economic indicators

The primary focus of this section is the problem of the suggested data set's broad temporal span. Comparing the same economic data from different years in order to determine how much inflation has affected things is not very interesting. For the "Real economy indicator" to properly represent the state of the economy, inflation must be eliminated. The GDP adjustment hypothesis modifies the economic measures. The conversion formula is given as follows:

$$Real\ GDP = Nominal\ GDP / GDP\ Deflator \tag{1}$$

The nominal economic indicators used in this article are the basic economic indicators. To get the real economic indicators that are used for comparison, the primary economic indicators are modified by Equation (2).

$$Real\ economic\ indicators = Nominal\ economic\ indicators / GDP\ Deflator \tag{2}$$

The GDP Deflator refers to the Trading Economics and the National Bureau of Statistics' International Statistical Yearbook.

Normalization

Due to the differences in their nature, indicators in a multi-indicator system typically have numerous orders of magnitude and units of measurement. The significance of the indicators with higher values would be exaggerated and the importance of the indicators with lower values would be over weakened if there was a significant variation in the order of scales among the various indicators [13]. Normalizing the original data is therefore required to ensure the accuracy of the findings. Equation (3) normalizes each indication of each storm surge occurrence.

$$s_i' = (s_i - s_{min}) / (s_{max} - s_{min}) \tag{3}$$

If s is the supplied indicator, s_i is the given indication of the data set's i storm surge event, and s_i' is the indicator following normalization. The terms "s_{max}" and "s_{min}" denote the lowest and greatest values of the indicator between all hurricane-related events, respectively. Following normalization, there is numerical comparability between the indicators of various dimensions.

Oversampling

Class equilibrium dispersion and equivalent error cost serve as the foundation for the implementation of classification reduction techniques on tasks involving classification. When a certain class in the information set has a large proportion, it can negatively affect the classifier's performance (class imbalance). The IV grade's events number in the main data set is significantly greater than that of the other grades, which has an important effect on the performance of calculated models. The Synthetic Minority Oversampling Technique, or SMOTE, was created to address issues with class imbalance. Therefore, the above-mentioned difficulty is addressed using the SMOTE.

The primary goal of the SMOTE technique is to synthesise more samples artificially in order to boost the minority sample set's data volume. By choosing at random a location on the line that joins a sample with its closest neighbours, new samples are produced. In Eq (4), the SMOTE method is displayed.

$$s_{new} = s_i + rand(0,1) \times (s_{j-} s_i) \tag{4}$$

Where s_i and s_j stand for the same indicator's two different values respectively, and s_{new} is the synthesized sample. The initial sample gathering is expanded to include all s_{new}, freshly generated samples.

The CV set is subjected to oversampling using the SMOTE method.

Comprehensive indicator system construction

The development of a rational and scientific indicator system for SSDL calculations is critical. Up until recently, China has lacked a centralized method for selecting SSDL indicators. Yet, Studies have concentrated on various periods and fields of study. Similar indicators may be discovered in the indicator system. The research presented here develops a complete indicator system, thoroughly examines the storm surge manufacturing procedure and data availability, and collects a large amount of open-source data. Three views are used in the construction of the complete indicator system: resilience, the risk of the disaster-bearing body, & the danger of disaster-causing elements. Typhoons, tidal levels, high winds or gusts, precipitation, sea level, topography, economics, society, population, catastrophe scenario, healthcare, a few of the factors that are considered are land use type, geographic features, public propaganda, and education.

Transmission for the distinctive indicator system

When overfitting occurs when there are too many factors taken into account relative to the amount of dataset, learning algorithms may become difficult. It seems sensible to normalize the loss function or use fewer variables for this reason.

Whenever possible, the necessary number of indicators will be included at the start of the modelling procedure in order to prevent the missing indicator's impact on the model deviation. Therefore, the comprehensive indicator technique starts with 50 indications input, even if noise and indicator interaction will always exist. This process of identifying which indicators within the dataset have an impact on the result is known as "indicator selection." The goal is to minimise the high-dimensional data size while keeping or improving accuracy. A strong and adaptable technique that can extract lower-dimensional representation from higher-dimensional information while preserving the most original information is principal component analysis (PCA). The classifier's success is gauged using the recursion feature elimination (RFE) method, which eliminates attributes one by one. An explanation of PCA and RFE is given below.

Recursive feature elimination

Permutation significance measure is used as a rating criterion by RFE to recursively reduces the indicators. Following the computation of the significance rankings for each indication the indicator with the lowest relevance value is eliminated from the classifier and the remainder Subsets will be built with indicators to reconstruct the SSDL calculation model. Every subset is used to retrain the classifier model, and the model's performance is also computed.

Depending on the classifier model being used, the important ranking scores can be computed using either the Gini index or the information gain method. To summarize, the process of calculating ranking importance involves the subsequent steps: (1) Determine the correctness of a simple decision tree (DT); in particular, the out-of-bag or Gini index is frequently used as an assessment indices to assess accuracy. (2) Based on the calculated accuracy, permute the indicators. (3) Recalculate the DT's accuracy by utilizing the permuted indicator to remove the covariate's information content. (4) Find the error in the accuracy between the recalculated (from (3)) accuracy and the initial (from

(1)) accuracy. (5) Continue with steps (1) through (4) for each DT; the total essential score is derived from the DTs' average accuracy. There are several advantages to the significance ranking evaluation mentioned above. It takes into account the effect of each predictor separately as well as the impact of multi-indicator relationships on other input indicators. It is objective and extensively applicable. The "feature_importances" toolset, which the tree-based models give, is used in this research to build the importance ranking.

The distinctive indicator system will be selected in this investigation using RFE and PCA, and the outcomes will be examined and contrasted.

Factors causing spikes in tropical storms

As Figures 3.1 illustrate, a storm surge is the result of several factors. The cyclonic winds are the main force behind tropical storm surges. Another significant factor that might affect the surge level is wave set-up. In contrast, the increase in water level caused by the low air pressure is negligible. A more thorough description of each force is provided below (For more thorough explanations of storm surge forcings, refer to the following descriptions, which are condensed and leave out much of the intricacy of storm surge dynamics).

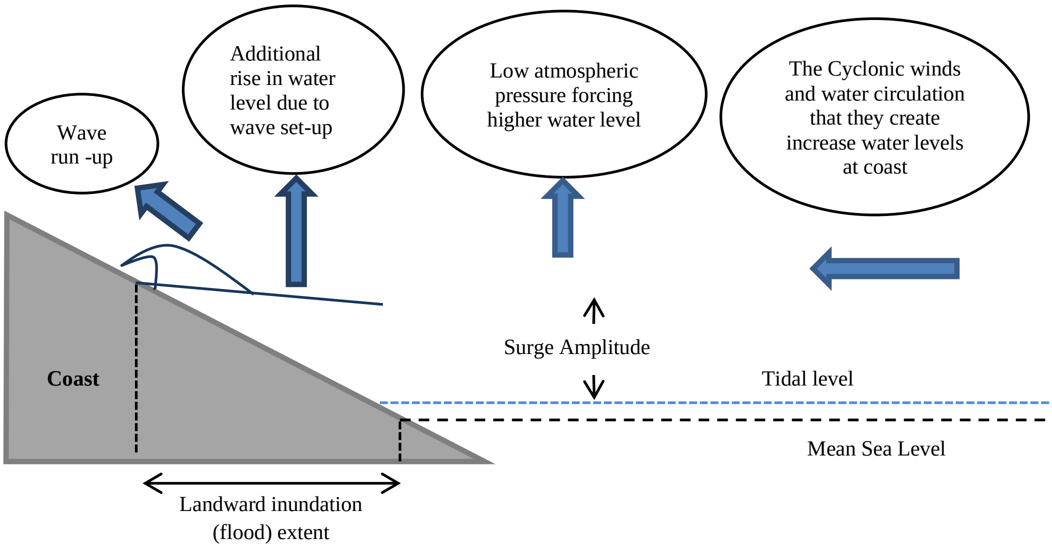

Fig. 3.1. Schematic graphic illustrating the many factors involved in the creation of tropical storm surges

Machine learning methods

We will quickly describe four more machine-learning techniques that we evaluated in our experiments in this section. These comprise Support Vector Regression (SVR), k-nearest Neighbours, and the Multi-Layer Perceptron. Because they may capture the mapping between variables that are inputs and outputs (forecast issues) without directly investigating the natural laws that control storm surge dynamics, these models are sometimes referred to as data-driven models. These models are entirely dependent on knowledge gleaned from the information collected. In Figure 3.2 Diagrammatic representations of the hazards-affected objects and disaster-causing causes in ML are illustrated.

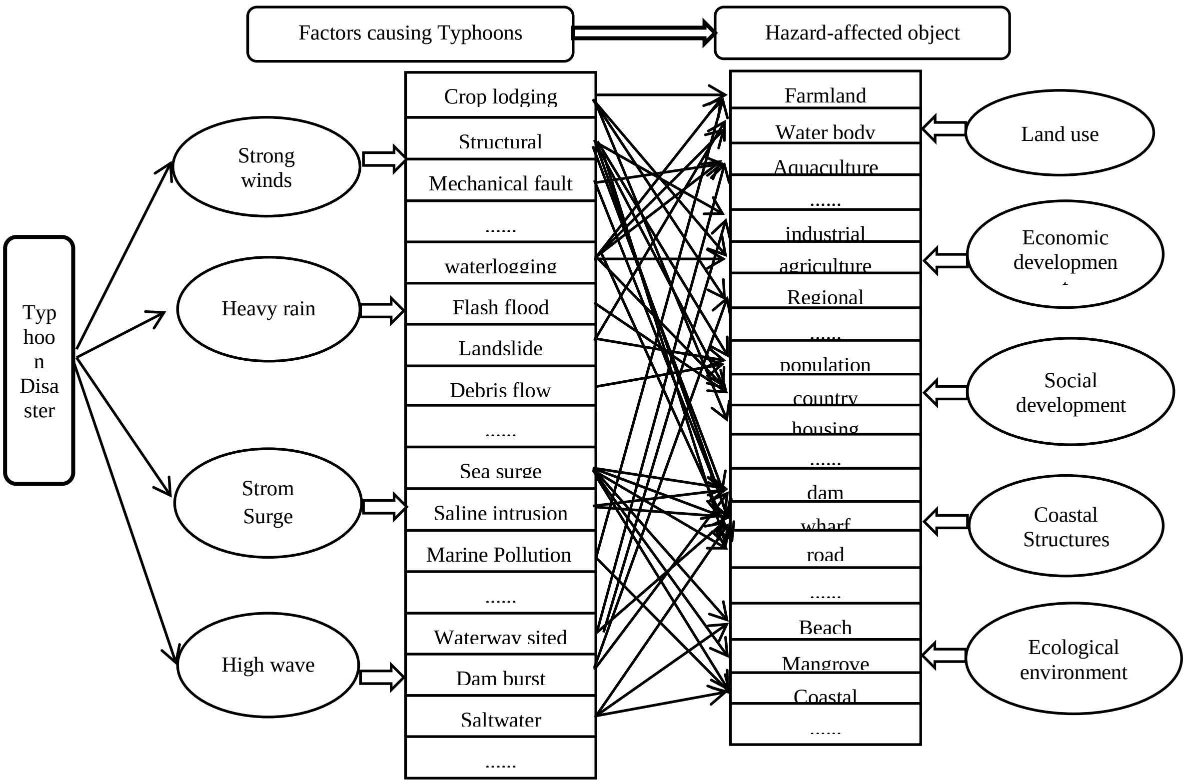

Fig. 3.2. Diagrammatic representations of the hazards-affected objects and disaster-causing causes in ML

K-nearest neighbor

A non-parametric statistical machine learning technique called k-NN makes predictions based on the target that the k closest neighbors of a given query point produce in Figure 3.3. Specifical, we compute the Euclidean measure value between a particular data point and every other point in our training set. Next, we choose the neighborhood's k closest neighbors. The average of the goal output values generated by these k closest neighbors is then used to set the forecast value. The prediction s_t is specifically specified as:

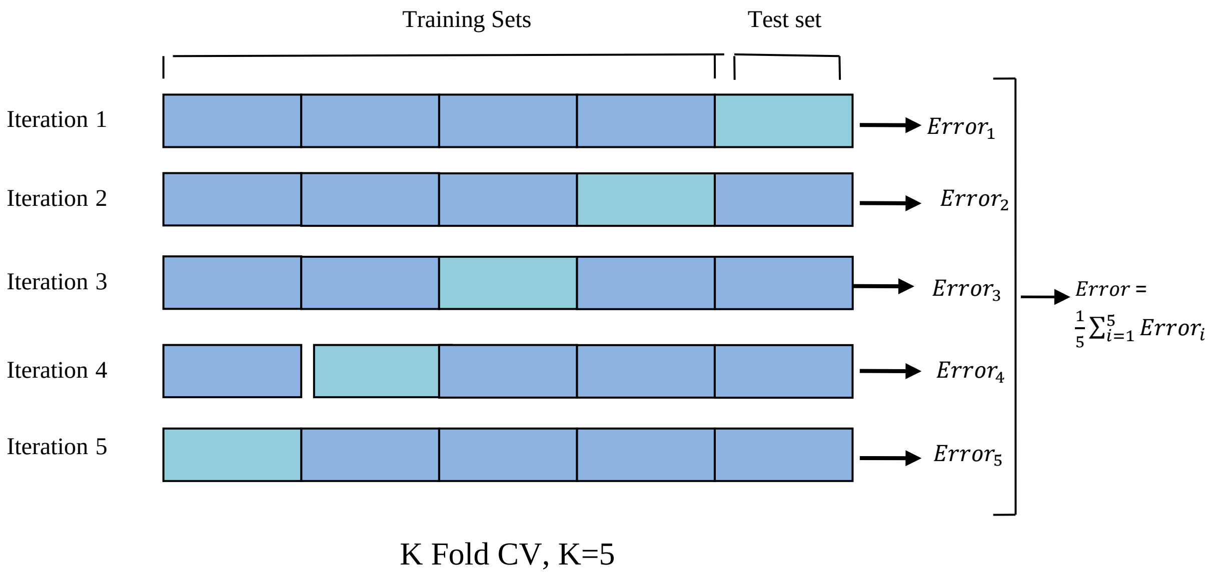

Fig. 3.3. K-fold Cross-Checking

$$x_t^E = \frac{1}{k}\sum_{t\in R(X,n)} x_t \qquad (5)$$

Where the k closest neighbor index assigned to the X (n) attribute vector is expressed by $\mathrm{R(X, n)}$. It makes sense that the sample average of the output surge level of the k closest neighbors to X (n) is what the prediction x_t^E in Equation (5) represents.

Support vector machines

Regression and classification issues may both be handled with support vector machines. Support vector regression (SVR) is the term used when it is used to a regression problem. The forecast is generated by taking into account a linear model as an instance.

$$e(x) = w^T x + p, \qquad (6)$$

Where the weighted vector, a bias, & the input vector are represented, respectively, by the symbols w, p, and x. For each n = 1... N, let x_n represent the n-th training input vector and y_n represent the desired output.

The error function is calculated using the formula below:

$$J = \frac{1}{2}\left|\left|w\right|\right|^2 + V\sum_{i=1}^{m}|y_n - e(x_n)|_\varepsilon, \qquad (7)$$

The first term in this mistake function is used to punish the model's complexity. The last term is commonly referred to as the "ε-insensitive loss function," meaning that errors below epsilon are not penalized [13]. The learned function for the linear example might be found by using the following minimizing method to Eq. (8).

$$e(x) = \sum_{n=1}^{N}(\alpha_n^* - \alpha_n)x_n^T x + p \qquad (8)$$

Where the Lagrange multipliers are α_n and α_n^*. A fundamental idea in the theory of SVR is the term "support vector," which refers to training vectors that produce nonzero Lagrange multipliers. The amount of support vectors indicates the degree of complexity of the model, whereas non-support vectors indirectly help in the solution. By using kernelization and the addition of the kernel's function κ in a replication of kernel Hilbert space (RKHS), this model may be enlarged to the non-linear scenario.

$$e(x) = \sum_{n=1}^{N} (\alpha_n^* - \alpha_n)\kappa(x_n^T x) + p \quad (9)$$

In our SVR tests, we use the commonly recognized Gaussian kernel. The Gaussian function's standard deviation is expressed by its width, δκ.

Multi-layer perceptron networks

The most often used type of perception network is the multi-layer one. The ANN (Artificial Neural Networks) model learns the network configuration using the error backward propagation approach. One hidden layer artificial neural networks (ANNs) are usually employed in hydrologic modeling because those networks are thought to offer enough complexity to effectively represent the nonlinear features of the hydrologic cycle. The ANN model for forecasting is defined as:

$$x_t^E = e(\mathrm{X_t}, \mathrm{w}, \theta, \mathrm{n}, \mathrm{h}) = \theta_0 + \sum_{i=1}^{h} w_i^{out} \varphi\left(\sum_{j=1}^{n} w_{ij} x_t + \theta_i\right), \quad (10)$$

When the transfer function is indicated by φ; The weight of the connection between the j-th node of the input layer and the i-th node of the hidden layer is denoted by w_{ij}; θ_i represents the bias associated with the hidden layer's i-th node; The weight of the connection between the i-th component of the layer that is hidden and the output layer node is denoted by w_i^{out}; and the output node's bias is indicated by θ_0. To use Equation (6) for surge level forecasting, an appropriate training technique is required to determine the ideal values of w and θ.

4. Implementation and Experimental results

4.1 ConvLSTM Model-Based Spatiotemporal Water Level Prediction

Once the aforementioned operations were finished, we had two models that were educated on different training sets. We assessed the accuracy and stability of these models using Super-Typhoon Usagi's storm surge water gain mechanism as a test case. Super-Typhoon Usagi, one of the strongest tropical cyclones in the western North Pacific that year, made landfall in 2013 off the southern shore of Shanwei Town, Guangdong Province. Importantly, our ConvLSTM models were not trained using the data associated with this typhoon. Furthermore, the model's predictions were verified using actual data from water level measurements made during Typhoon Usagi at Henglan Island Station (114.182◦ E, 22.110◦ N). The real sea level measurement data at the tide check station came from the integrated lake level datasets at the National Marine Information Centre.

4.2 Outcomes of a Single-Step Prediction

We tested the two constructed models using the previously described methodology, and both models performed exceptionally well on the test data. We matched these network models' predictions with the actual data over the following hour, using measures such as mean absolute errors (MAE) and mean squared error (MSE) to gauge how accurate they were.

We computed the MAE by contrasting the ConvLSTM model's predictions with the actual data to evaluate the two models' predictive abilities, displays the geographic spread of the MAE for the model educated on the Water Level Dataset, whereas the MAE distribution for the models trained on the Water Level Variation Dataset. Notably, both models' forecasts often show higher absolute error levels in places close to the shore.

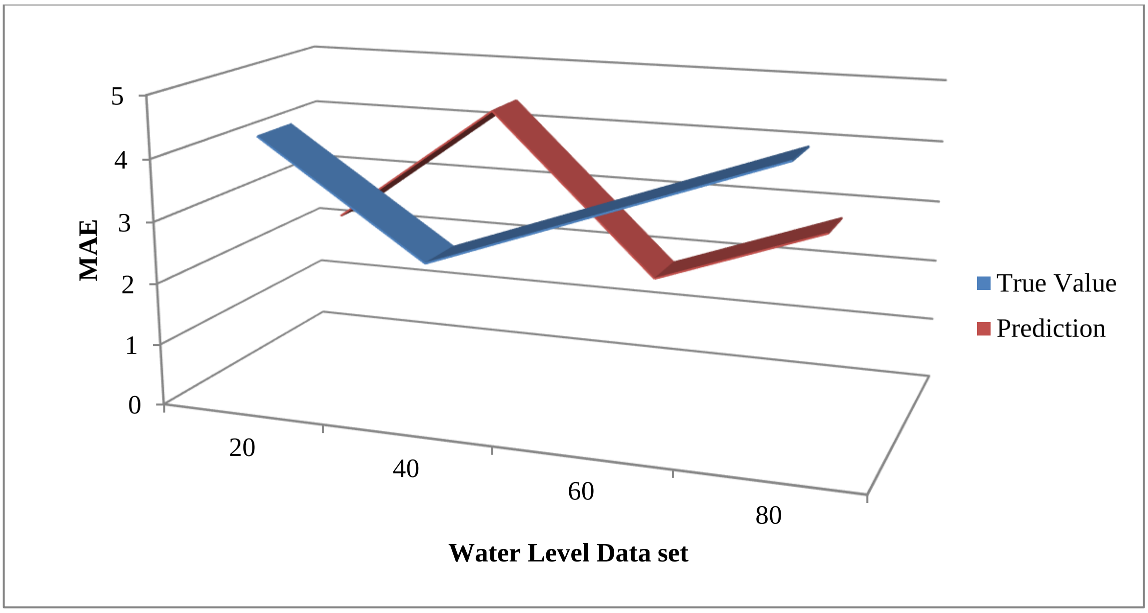

Fig 4.1. The model's prediction outcome after training on the MAE and Water Level Change Dataset

The results of the research area's global water level data projection are displayed above. We randomly chose a point and extracted its prediction outcomes from the global prediction outcomes in order to thoroughly assess the dependability of the model. The model trained on this Water Level Dataset has higher dispersion and volatility in its residuals, as seen by the two distinct violin charts [14, 15]. Above are the global findings of the research area's water level data prediction. We chose a location at random and took its prediction results out of the global prediction results in order to thoroughly assess the reliability of the model. The two distinct violin charts show that the residuals of the model trained on this Water Level Dataset are more dispersed and volatile, MAE and MSE. Table 1 provides an instant of the findings.

Table 1. Evaluation table of MAE and MSE of the two models

	MAE (m)	MSE (m^2)
The Model trained on the Water Level Dataset	0.028	0.0039
The Model trained on the Water Level Change Dataset	0.015	0.0008

It is clear from looking at the figures in the table that indirect water level prediction, which forecasts changes in water level, provides better outcomes than direct water level forecasting. It is beneficial to utilize the model of neural networks for prediction because its activation function is always constrained; demonstrating that compared to the water level value, the numerical range inside the region of water level variation is narrower and more concentrated. Furthermore, we conducted experiments to evaluate the MAE & MSE of model projections on the learning set using the model that we had learnt on datasets of varying sizes. As seen in Figure 4.1, both models had declining MAE & MSE values as the amount of the dataset increased. When the size of the data set rose, the model trained on the Waters Level Dataset showed a more pronounced fluctuation in its forecasts than the model trained on the Water Level Change Dataset. It seems that the ConvLSTM model can generate reliable predictions with a smaller dataset when trained on the Water's Level Change Dataset.

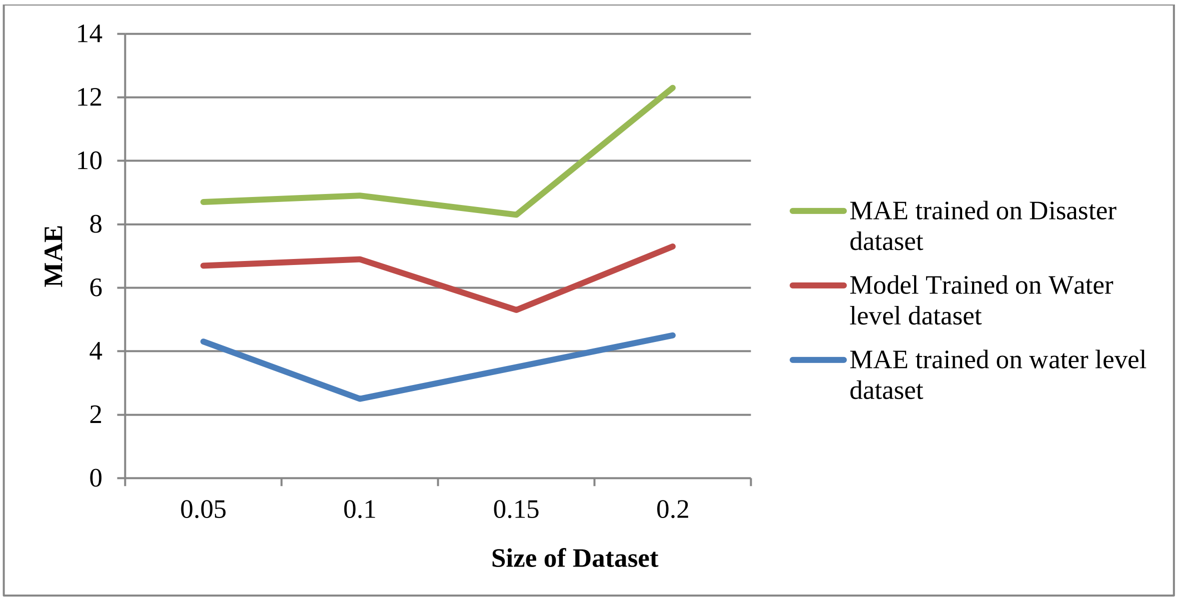

Fig. 4.2. multi-step forecasts made by the two models of how dataset size affects changes in MAE and MSE

4.3 Results of Multi-Step Prediction

The water level or the water level change information for the next hour is predicted using the hourly data of the present time and the previous eight hours, as shown in the above findings. We now use a recursive multi-step forecasting technique to anticipate the water level or the water level change data for the next 12 hours using the hourly data on the water level or change for the last 8 hours and the current time. We use the MAE statistical measure to assess the long-term prediction performance of both models. We research the predictive power of the models across prolonged periods, starting from 10:00 a.m. on September 20, 2013.

As Figure 4.2 illustrates, the model trained on the Waters Level Variation Dataset consistently outperforms the Waters Level Dataset model when it comes to longer-term predictions. The model's reliability for forecasting the water's level or change in the next six hours based on data from the previous eight hours and the current time is rather high. But when the prediction time increases, the outcomes usually go worse.

5. Conclusion

The existence of ML models in SSDL estimation study is confirmed, and this work obtains the optional SSDL estimation model and a high-accuracy estimation approach. With their rapid development and continued concentration of population along the coast, coastal cities are increasingly vulnerable to marine dynamics disasters such as storm surges. This poses a significant barrier to sustainable development. This study's outcome is groundbreaking as it is the first comprehensive assessment of coastal China's vulnerability to storm surges from the perspective of prefecture-level cities. The vulnerability index, which offers a thorough assessment of the multidimensional structure of vulnerability, was created by integrating the three components of vulnerability. Moreover, we assessed the influence of environmental factors on the susceptibility of disaster-prone regions to storm waves at the patch scale.

However, it is challenging to assess the model's performance's relevance and to validate it with any level of reliability due to the lack of data. In order to improve our dataset—particularly for use with random forest models—this thesis needs to keep investigating the use of artificial typhoon data generating methods. Apart from ConvLSTM, we also intend to investigate and test other spatiotemporal prediction models to see how well they work and whether they are appropriate for our particular needs. With regard to predicting water levels in relation to typhoons, these endeavours are expected to enhance our comprehension of prediction methodologies and maybe augment the accuracy and robustness of our predictive models.

References

[1] Hien, N. T., Tran, C. T., Nguyen, X. H., Kim, S., Phai, V. D., Thuy, N. B., & Van Manh, N.

(2020). Genetic programming for storm surge forecasting. *Ocean Engineering*, *215*, 107812.

[2] Qin, Y., Su, C., Chu, D., Zhang, J., & Song, J. (2023). A Review of Application of Machine Learning in Storm Surge Problems. *Journal of Marine Science and Engineering*, *11*(9), 1729.

[3] Mu, J., Wu, F., & Zhang, A. (2014, August). Housing value forecasting based on machine learning methods. In *Abstract and Applied Analysis* (Vol. 2014). Hindawi.

[4] Du, X., Li, X., Zhang, S., Zhao, T., Hou, Q., Jin, X., & Zhang, J. (2022). High-accuracy estimation method of typhoon storm surge disaster loss under small sample conditions by information diffusion model coupled with machine learning models. *International Journal of Disaster Risk Reduction*, *82*, 103307.

[5] Devaraj, J., Ganesan, S., Elavarasan, R. M., & Subramaniam, U. (2021). A novel deep learning based model for tropical intensity estimation and post-disaster management of hurricanes. *Applied Sciences*, *11*(9), 4129.

[6] Fang, W., Xue, Q., Shen, L., & Sheng, V. S. (2021). Survey on the application of deep learning in extreme weather prediction. *Atmosphere*, *12*(6), 661.

[7] Kolla, V. R. K. (2018). Forecasting the Future: A Deep Learning Approach for Accurate Weather Prediction. *International Journal in IT & Engineering (IJITE)*.

[8] Ian, V. K., Tse, R., Tang, S. K., & Pau, G. (2023). Bridging the Gap: Enhancing Storm Surge Prediction and Decision Support with Bidirectional Attention-Based LSTM. *Atmosphere*, *14*(7), 1082.

[9] Chen, R., Zhang, W., & Wang, X. (2020). Machine learning in tropical cyclone forecast modeling: A review. *Atmosphere*, *11*(7), 676.

[10] Snaiki, R., Wu, T., Whittaker, A. S., & Atkinson, J. F. (2020). Hurricane wind and storm surge effects on coastal bridges under a changing climate. *Transportation research record*, *2674*(6), 23-32.

[11] Yi, X., Sheng, K., Wang, Y., & Wang, S. (2021). Can economic development alleviate storm surge disaster losses in coastal areas of China?. *Marine Policy*, *129*, 104531.

[12] McIvor, A. L., Spencer, T., Möller, I., & Spalding, M. (2012). *Storm surge reduction by mangroves*. The Nature Conservancy and Wetlands International.

[13] Liu, Y., Zhao, Q., Hu, C., & Luo, N. (2023). Prediction of Storm Surge Water Level Based on Machine Learning Methods. *Atmosphere*, *14*(10), 1568.

[14] Zhang, S., Zhang, J., Li, X., Du, X., Zhao, T., Hou, Q., & Jin, X. (2022). Estimating the grade of storm surge disaster loss in coastal areas of China via machine learning algorithms. *Ecological Indicators*, *136*, 108533.

[15] Gu, J., Zhu, M., & Jiang, L. (2011). Housing price forecasting based on genetic algorithm and support vector machine. Expert Systems with Applications, 38(4), 3383-3386.

[16] Dhablia, A. (2021). Integrated Sentimental Analysis with Machine Learning Model to Evaluate the Review of Viewers. Machine Learning Applications in Engineering Education and Management, 1(2), 07–12. Retrieved from http://yashikajournals.com/index.php/mlaeem/article/view/12

[17] Soundararajan, R., Stanislaus, P.M., Ramasamy, S.G., Dhabliya, D., Deshpande, V., Sehar, S., Bavirisetti, D.P. Multi-Channel Assessment Policies for Energy-Efficient Data Transmission in Wireless Underground Sensor Networks (2023) Energies, 16 (5), art. no. 2285, .