Journal of Civil Engineering and Urban Planning (2025)
Clausius Scientific Press, Canada
DOI: 10.23977/jceup.2025.070109
ISSN 2616-3969 Vol. 7 Num. 1

Prediction and Analysis of Tropical Cyclone Genesis Based on Artificial Intelligence Technology and Its Application in Civil Engineering

Xue Wang[1,a], Xueting Qin[1,b,*]

[1]*Faculty of Construction, Guangdong Technology College, Zhaoqing, Guangdong, China*
[a]1257448273@qq.com, [b]1308490018@qq.com
[*]Corresponding author

Keywords: Tropical Cyclone, Artificial Intelligence, Variational Autoencoder, Wind Load, Building Structure

Abstract: Tropical cyclones (TCs) are among the most destructive natural disasters, posing significant threats to coastal regions, particularly in southeastern China, where rapid economic growth and urbanization have intensified the risks associated with TC-induced wind hazards. To enhance predictive capabilities and mitigate potential damage, this study leverages advanced artificial intelligence (AI) techniques, focusing on deep learning-based Variational Autoencoder (VAE) models, to analyze and forecast the genesis of tropical cyclones in the Northwest Pacific. By training the VAE on historical TC data, the model effectively captures the underlying patterns governing TC formation, enabling accurate simulations of both the frequency and spatial distribution of these events. The findings reveal that the VAE model performs robustly in replicating observed TC climatology, offering critical insights for risk assessment and disaster preparedness. Furthermore, the study highlights the practical applications of AI-driven TC predictions in civil engineering, particularly in improving wind load calculations and optimizing structural designs to enhance resilience against extreme wind events. This research underscores the potential of AI technologies in advancing meteorological forecasting and supporting sustainable infrastructure development in cyclone-prone regions.

1. Introduction

Tropical cyclones are among the most destructive natural disasters globally, particularly in the southeastern coastal regions of China, where rapid economic growth and urbanization have heightened the risks associated with TC-induced wind disasters [1,2]. Statistics show that approximately one-third of global tropical cyclones (around 30 annually) originate in the Northwest Pacific, with an average of seven TCs making landfall in southeastern China and the South China region each year, causing economic losses exceeding 20 billion yuan and hundreds of casualties [3,4]. In recent years, advancements in construction technology have led to the development of numerous high-rise buildings with novel structural systems in southeastern China. These modern, large-scale structures are characterized by their height and flexibility, making them highly sensitive to wind effects. The wind-induced responses (e.g., vibration) of these structures are highly correlated with wind speed, meaning even small changes in wind speed can lead to significant variations in wind load and wind-induced responses, thereby affecting the safety and comfort of the structures.

The impact of tropical cyclones on building structures is primarily manifested in wind load and wind-induced vibration responses. Wind load is a critical parameter in structural design, directly influencing the safety and economic viability of buildings. The strong winds associated with tropical cyclones can induce significant wind-induced vibrations, leading to structural fatigue, localized damage, or even total collapse. Therefore, accurate prediction of tropical cyclone genesis and trajectory is essential for wind load calculations and structural design.

2. Research Methodology

Deep learning, as a key predictive technology, plays a significant role in the field of machine learning. Compared to traditional machine learning methods, deep learning offers greater flexibility and adaptability. In recent years, with continuous technological advancements, new algorithms and models in deep learning have emerged, and their application demands have grown. Successful applications in fields such as healthcare, education, and finance have demonstrated the immense potential of deep learning. Consequently, an increasing number of researchers are exploring the use of deep learning techniques for predicting tropical cyclone activity.

This study employs a deep learning model known as the Variational Autoencoder (VAE) [5]. VAE is an unsupervised deep learning generative model capable of modeling the distribution of training data and generating diverse, complex data. In recent years, VAE has become a popular method for handling complex distribution problems, widely applied in areas such as facial recognition, handwritten digit recognition, image segmentation, scene modeling, and static image prediction [6]. However, its application in predicting key parameters of tropical cyclones remains limited.

This study utilizes a deep learning simulation approach, based on the tropical cyclone best-track dataset released by the China Meteorological Administration, to construct a tropical cyclone genesis model. The model employs a VAE structure, capable of handling data with latent variables and generating new data not present in the input data [6]. VAE trains two main functions through neural networks: q(z/x) and p(x/z), corresponding to the encoder network and decoder network (also known as the inference network and generative network), respectively. As shown in Figure 1, the encoder network, parameterized by θ, maps the input x to the latent variable z, while the decoder network, parameterized by φ, reconstructs the output x from the latent variable z. This approach allows VAE to effectively simulate and predict tropical cyclone behavior.

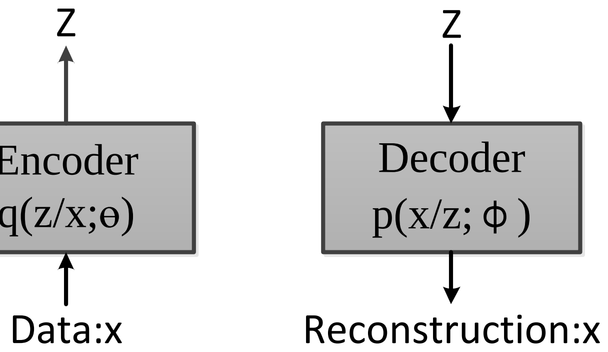

Fig. 1. Basic Structure of VAE

In this study, the VAE generative model consists of an encoder network and a decoder network, as shown in Figure 1. The encoder network includes a connection layer, mean layer, variance logarithmic layer, and output layer, while the decoder network comprises one sampling layer and four connection layers. The model takes the monthly tropical cyclone genesis count in the Northwest Pacific region (2.5°×2.5° grid) from January to December over a 40-year period (1979–2018) as input. The model outputs two components: (1) the predicted monthly tropical cyclone genesis count for the next 100 years, and (2) the spatial distribution of monthly tropical cyclone genesis for the next 100 years. This design enables the VAE model to effectively simulate the genesis patterns of tropical cyclones and provide reliable data for long-term predictions.

This study compares the VAE generative model with traditional probabilistic models to evaluate the predictive accuracy of the AI model. Rumpf et al. [7,8], Emanuel et al. [9], and Hall and Jewson [10] have used historical genesis points to estimate the genesis probability at each location in the ocean using Gaussian kernel density functions. Therefore, this study adopts the biased cross-validation method proposed by Sain et al. [11] to calculate the optimal bandwidth for longitude, latitude, and time dimensions, as shown in Equation (1). Since the genesis location of tropical cyclones is determined by three dimensions (longitude, latitude, and time, i.e., d=3), Equation (1) can be further simplified to Equation (3) for calculation. Based on this, the kernel density function (Equation (1)) is used to estimate the probability density of tropical cyclone genesis count, the probability density of longitude and latitude distribution, and the probability density of genesis time [12]. This comparative analysis provides a basis for validating the superiority of the VAE model.

$$f(x)=\frac{1}{nh}\sum_{i=1}^{n}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(x-x_i)^2}{2h^2}\right) \tag{1}$$

x—parameter to be estimated;
x_i—sample parameter;
h—optimal bandwidth;
n—sample size.

$$BCV(h_1,\dots,h_d)=\frac{1}{(2\sqrt{\pi})^d nh_1h_2h_d}+\frac{1}{4n(n-1)h_1h_2h_d}\times$$
$$\sum_{i=1}^{n}\sum_{j\neq i}\left[\left(\sum_{k=1}^{d}\Delta_{ijk}{}^2\right)^2-(2d+4)\left(\sum_{k=1}^{d}\Delta_{ijk}{}^2\right)+\left(d^2+2d\right)\right]\prod_{k=1}^{d}\phi\Delta_{ijk} \tag{2}$$

$$BCV(h_1,h_2,h_3)=\frac{1}{\left(2\sqrt{\pi}\right)^3 nh_1h_2h_3}+\frac{1}{4n(n-1)h_1h_2h_3}\times$$
$$\sum_{i=1}^{n}\sum_{j\neq i}\left[\left(\Delta^2{}_{ij1}+\Delta^2{}_{ij2}+\Delta^2{}_{ij3}\right)^2-10\left(\Delta^2{}_{ij1}+\Delta^2{}_{ij2}+\Delta^2{}_{ij3}\right)+15\right]\times$$
$$\left[\frac{1}{2\pi\sqrt{2\pi}}\exp\left(-\frac{\Delta^2{}_{ij1}+\Delta^2{}_{ij2}+\Delta^2{}_{ij3}}{2}\right)\right] \tag{3}$$

h1, h2.....hd—optimal bandwidth for different dimensions;

$\Delta_{ijk}=\frac{x_i-x_j}{h}$ —cross-validation sample for the k-th dimension.

3. Results and Analysis

3.1 Kernel Probability Density Model

The data for this study are derived from the tropical cyclone best-track dataset provided by the China Meteorological Administration [13,14]. Using the kernel density estimation method, the annual genesis count, occurrence time, and spatial location (longitude and latitude) of tropical cyclones in the Northwest Pacific from 1979 to 2018 were quantitatively assessed. The study area was divided into 2.5°×2.5° grid cells, and the kernel density calculation results were used to quantify the genesis probability of tropical cyclones in each grid.

The three-dimensional characteristics (longitude, latitude, time) of tropical cyclone genesis locations are shown in Figure 2. Using Equation (1), the annual genesis count was estimated, revealing that the annual genesis count of tropical cyclones is concentrated in the range of 25–30 (Figure 3). Notably, due to partial missing historical data and differences in statistical standards among meteorological agencies, the historical distribution exhibits non-uniform characteristics. To validate the simulation results, the Kolmogorov-Smirnov (K-S) test and rank-sum test were used to analyze the consistency between the simulated data and historical data distributions. The results showed no significant difference at the 0.05 significance level (p=0), confirming the reliability of the model.

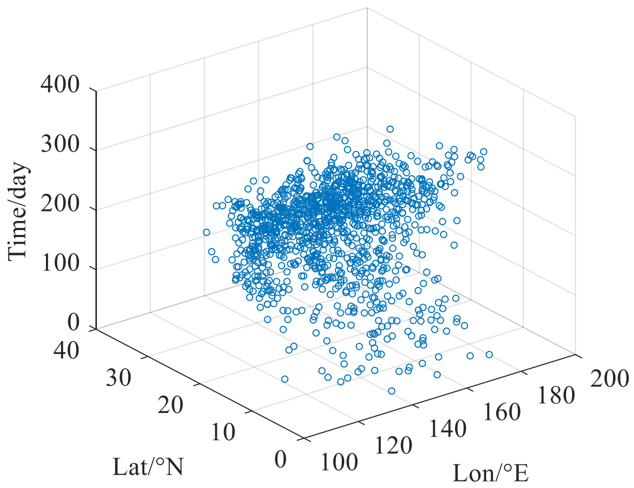

Fig. 2. Spatial Distribution of Tropical Cyclones

Further analysis of the spatiotemporal distribution patterns of tropical cyclones in the Northwest Pacific was conducted using the kernel density estimation method. The temporal dimension analysis (Figure 4) indicates that the peak period for tropical cyclone genesis is between July and October, with the highest frequency occurring in August. The simulation results show that the error for April–October and November is below 20%, with errors for April, May, August, September, and October further reduced to below 10%. In the spatial dimension analysis, the longitude simulation exhibited higher errors in the 100 °E–110 °E range, while errors in other regions were below 6%. In the active tropical cyclone region of 110 °E–160 °E, the error was further reduced to below 3%. Specifically, in the 130 °E–140 °E region (where the genesis probability is highest), the simulation error was only 0.02%. In the latitude simulation, tropical cyclones primarily originated in the 5 °N–20 °N region, with errors generally below 5%. In the 10 °N–15 °N region (where the probability density peaks), the error was only 0.87%. Overall, the simulation accuracy in the spatial dimension was significantly higher than in the temporal dimension, and increasing the data volume significantly reduced errors.

Based on the kernel density estimation results, this study generated a probability density distribution map of tropical cyclone genesis in the Northwest Pacific (Figure 5). A comparison between historical and simulated data revealed that both datasets showed the highest genesis probability in the 130°E–135°E and 10°N–20°N regions (historical value: 0.0851 vs. simulated value: 0.0878), with a difference of only 4. In most regions, the difference approached zero, indicating that the kernel density model performs exceptionally well in simulating scenarios without climate change.

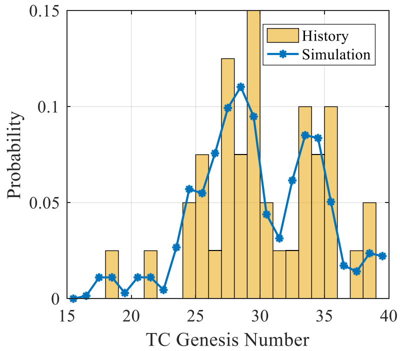

Fig. 3. Probability Distribution of Annual Tropical Cyclone Genesis Count

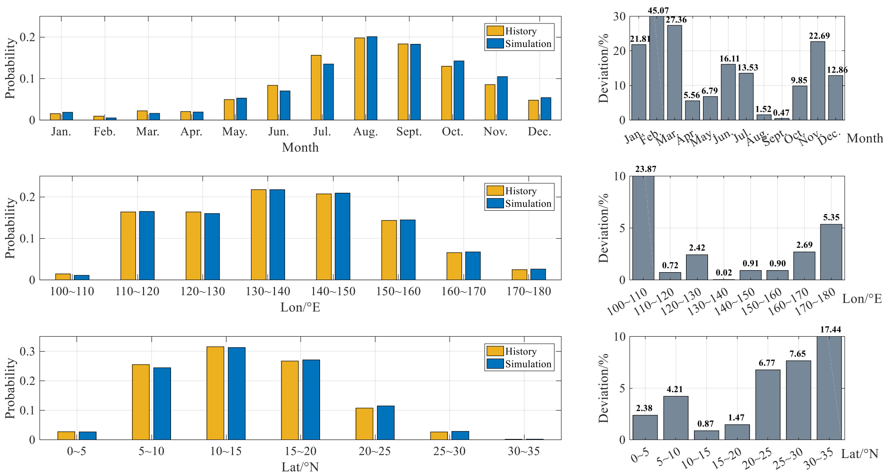

Fig. 4. Probability Distribution and Error Comparison of Tropical Cyclones in Different Dimensions

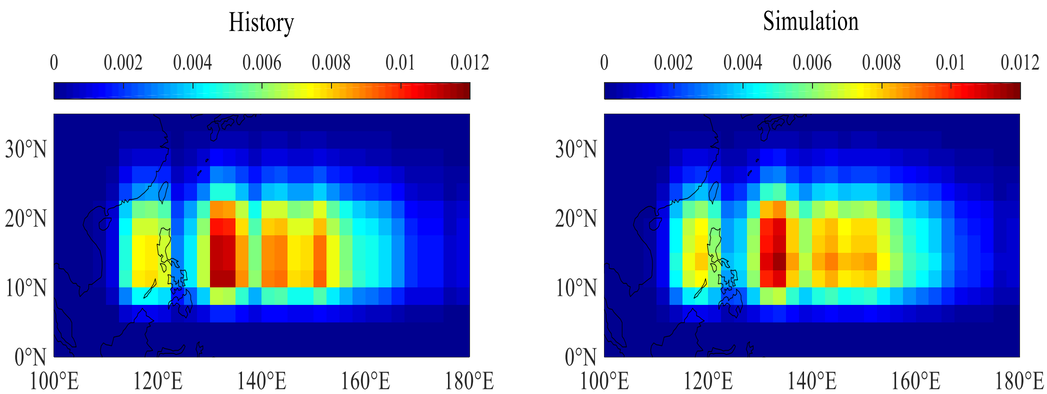

Fig. 5. Historical and Simulated Probability Density Distribution of Tropical Cyclone Genesis Locations

3.2 VAE Model

This section employs the Variational Autoencoder (VAE) to construct a tropical cyclone genesis model, thereby expanding the research methodology. The input data consist of the monthly genesis count in the Northwest Pacific (2.5°×2.5° grid) from 1979 to 2018. After training the model, the output includes the monthly genesis count and spatial distribution of tropical cyclones for the next 100 years. This study adopted the average of five independent outputs from three VAE models as the final prediction to enhance the robustness of the results.

The model performance analysis (Figure 6) shows that the historical annual genesis count was 29.73, while the predicted means for the three VAE models were 30.88 (Model 1), 30.54 (Model 2), and 30.24 (Model 3), closely aligning with the historical data. The probability density distribution indicates that Model 1 has a broader prediction range (20–36), with the probability peak distributed between 34–35. Model 2 (23–34) and Model 3 (25–33) exhibit more concentrated unimodal distributions (peaks at 32–33 and 30–31, respectively), suggesting that model complexity significantly impacts output stability.

The spatial distribution comparison (Figure 7) shows that the simulation results closely match the historical data in the 110°E–150°E and 5°N–20°N regions. The genesis hotspots are divided into two regions, with Manila as the boundary: 110°E–120°E (12.5°N–20°N) and 127.5°E–145°E (7.5°N–17.5°N). The differences between models may stem from insufficient data quality or limitations in generalization ability, but overall, the results validate the applicability of VAE for long-term tropical cyclone prediction.

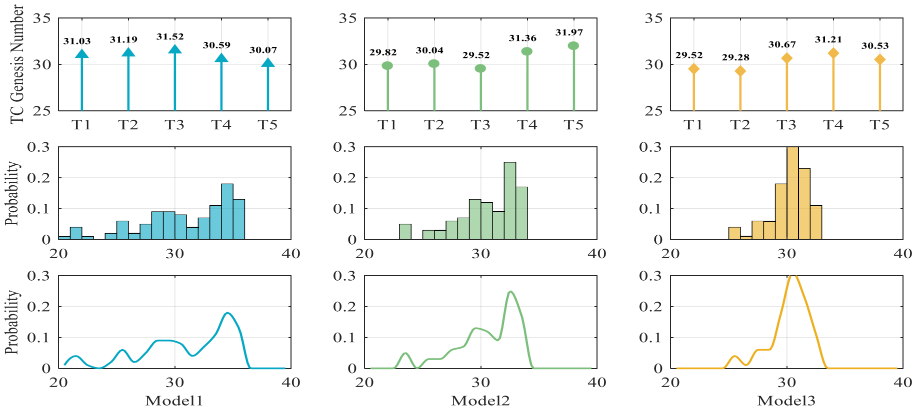

Fig. 6.Estimated Annual Tropical Cyclone Genesis Count by Different Models

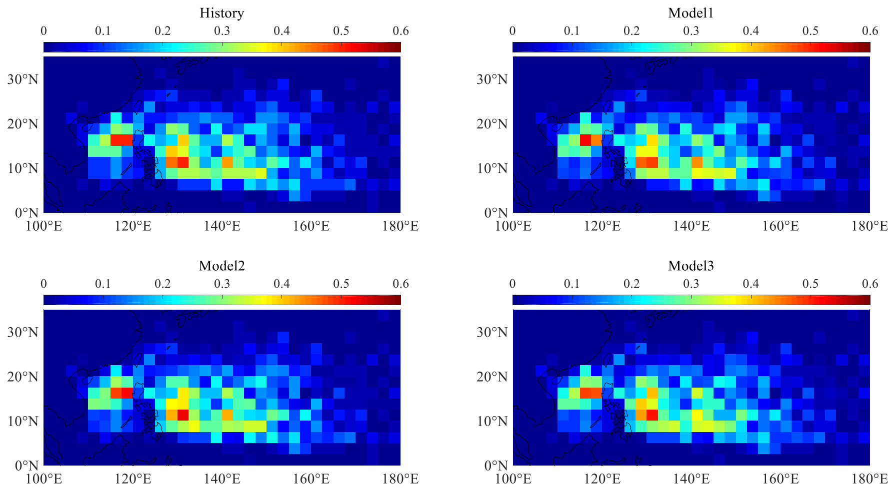

Fig. 7. Historical and Future Distribution of Tropical Cyclone Genesis Count

4. Conclusion

This study demonstrates the effectiveness of using artificial intelligence, particularly the Variational Autoencoder (VAE) model, in predicting tropical cyclone generation and distribution. The VAE model provides accurate simulations of tropical cyclone behavior, offering valuable insights for wind load calculations and structural design in civil engineering. The results highlight the potential of AI-driven approaches in mitigating the risks associated with tropical cyclones, particularly in vulnerable regions like southeastern China.

References

[1] Fischer T, Su B, Wen S. Spatio-temporal analysis of economic losses from tropical cyclones in affected provinces of China for the last 30 years (1984–2013)[J]. Natural Hazards Review, 2015, 16(4): 04015010.

[2] Tong B, Wang X, Fu J Y, et al. Short-term prediction of the intensity and track of tropical cyclone via ConvLSTM model[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2022, 226: 105026.

[3] Chan J C L. Interannual and interdecadal variations of tropical cyclone activity over the western North Pacific[J]. Meteorology and Atmospheric Physics, 2005, 89(1): 143-152.

[4] Xiao F, Xiao Z. Characteristics of tropical cyclones in China and their impacts analysis[J]. Natural Hazards, 2010, 54(3): 827-837.

[5] Kingma D P, Welling M. Auto-encoding variational bayes[J]. arXiv preprint arXiv:1312.6114, 2013.

[6] Doersch C. Tutorial on variational autoencoders[J]. arXiv preprint arXiv:1606.05908, 2016.

[7] Rumpf J, Weindl H, Höppe P, et al. Stochastic modelling of tropical cyclone tracks[J]. Mathematical Methods of Operations Research, 2007, 66(3): 475-490.

[8] Rumpf J, Weindl H, Höppe P, et al. Tropical cyclone hazard assessment using model-based track simulation[J]. Natural Hazards, 2009, 48(3): 383-398.

[9] Emanuel K, Ravela S, Vivant E, et al. A statistical deterministic approach to hurricane risk assessment[J]. Bulletin of the American Meteorological Society, 2006, 87(3): 299-314.

[10] Hall T M, Jewson S. Statistical modelling of North Atlantic tropical cyclone tracks[J]. Tellus A, 2007, 59(4): 486-498.

[11] Sain S R, Baggerly K A, Scott D W. Cross-validation of multivariate densities[J]. Journal of the American Statistical Association, 1994, 89(427): 807-817.

[12] Chen Y, Duan Z, Yang J, et al. Typhoons of western North Pacific basin under warming climate and implications for future wind hazard of East Asia[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2021, 208: 104415.

[13] Ying M, Zhang W, Yu H, et al. An overview of the China Meteorological Administration tropical cyclone database[J]. Journal of Atmospheric and Oceanic Technology, 2014, 31(2): 287-301.

[14] Lu X, Yu H, Ying M, et al. Western North Pacific tropical cyclone database created by the China Meteorological Administration[J]. Advances in Atmospheric Sciences, 2021, 38(4): 690-699.