Comparative Analysis of TEC Prediction Models During Geomagnetic Storm and Quiet Conditions Using Diffusion, Transformer, and SARIMA

Yuhuan Yuan[1] **Guozhen Xia**[1] **Xinmiao Zhang**[1]

[1]Department of Space Physics, Wuhan University, Wuhan, China

Key Points:

- This study introduces three innovative models designed to improve the prediction of ionospheric conditions, particularly during geomagnetic storms. These include the Diffusion model, the SARIMA model, and the Transformer model.
- The performance of Diffusion model is rigorously compared against the established C1PG model to evaluate their effectiveness and accuracy in forecasting geomagnetic conditions, with only one-year of training data.
- The findings suggest a significant step forward in the field of ionospheric forecasting, showcasing the potential of the generative model as a powerful tool for predicting geomagnetic disturbances with only one-year of training data.

Corresponding author: Xinmiao Zhang, zxmwhu@whu.edu.cn

Abstract

Total Electron Content (TEC) of the ionosphere under fluctuating geomagnetic conditions plays a pivotal role in space-based communication and navigation systems. Periods of geomagnetic storms have a significant effect on the ionospheric TEC, thereby impacting the precision of Global Navigation Satellite Systems (GNSS) as well as other communication systems. This study evaluates the performance of three predictive models: Diffusion, Transformer, and Seasonal Autoregressive Integrated Moving Average (SARIMA) by forecasting TEC during different geomagnetic storm conditions. Diffusion model predicts by simulating the random processes that govern ionospheric variability, and SARIMA model relies on statistical principles to emulate seasonality, trend, and autocorrelation structure within data. Our study uses TEC and geomagnetic data to differentiate between storm and quiet periods, evaluating each model's predictive accuracy in these distinct scenarios. Performance was benchmarked against the Center for Orbit Determination in Europe's (C1PG) 1-day ahead ionospheric forecasts. This study specifically concentrated on the performance of the prediction models within a singular grid cell located at the exact geographical coordinates of 114.3° E longitude and 30.5° N latitude. The study indicates that all three models exhibit exceptional forecasting abilities during both geomagnetic storm and quiet periods. Significantly, Diffusion Model surpasses the others, achieving an outstanding 85.64% of its predictions within the high-correlation interval ranging from 0.95 to 1.00 during quiet geomagnetic periods, while the C1PG model records 51.66%. Otherwise, during strong geomagnetic storm periods, Diffusion model operates at a 75.41% accuracy rate, while the C1PG achieves a 71.77% accuracy rate.

Keywords: TEC Prediction, Geomagnetic Storm, Quiet Conditions, Transformer Model, Diffusion Model, SARIMA.

Plain Language Summary

Predicting ionospheric conditions during complex geomagnetic storms presents a significant challenge due to the environment's inherent unpredictability. To enhance prediction accuracy with minimal training data, this study investigates three innovative models: the Diffusion model, SARIMA model, and Transformer model, and assesses their efficacy in comparison to the established C1PG model. Our comprehensive analysis, centered on prediction accuracy, reveals that the Diffusion model notably outshines the others, offering unparalleled performance in accurately forecasting conditions during both quiet and strong geomagnetic storm periods. Distinctively, the Diffusion model surpasses the other three models, achieving remarkable accuracy in the high-correlation range of 0.95 to 1.00. This achievement underscores its exceptional ability to provide reliable predictions in the face of geomagnetic variability.

1 Introduction

Many studies have focused on analyzing trends in the Total Electron Content (TEC) of the Earth's ionosphere, a subject of growing interest due to its implications for space-based technologies (Lean et al., 2011; Lastovicka et al., 2017). With the development of the rapid technological advancement, the precision and dependability of Global Navigation Satellite Systems (GNSS) have grown increasingly critical (Ratnam et al., 2018; Z. Li et al., 2019). The TEC of the ionosphere, a key parameter influencing radio wave propagation, significantly affects the performance of GNSS and other space-based communication systems (Béniguel, 2002). Condequently, Understanding and predicting TEC values is therefore crucial for mitigating the ionospheric impact on these technologies. However, the ionosphere's behavior is complex and influenced by various factors, including geomagnetic activity, which can vary from quiet to intense storm conditions (Reinisch & Galkin, 2011). This variability presents a challenge for accurate TEC prediction. The ionosphere's dynamics are complex and subject to a wide range of influences, including

geomagnetic activities that range from calm to severe storm conditions. This variability poses a significant challenge in reliably forecasting TEC values, particularly during periods of strong geomagnetic storms.

Improvements in time series forecasting and data analysis techniques have markedly increased the accuracy of TEC predictions, capturing both seasonal patterns and trends more effectively (Lim & Zohren, 2021). Recent studies, such as (Xia, Liu, et al., 2022) have employed Transformer models, which utilize self-attention mechanisms to effectively address the challenge of modeling long-term dependencies in global TEC variations. The Transformer model, with its advanced attention mechanisms, has shown great promise in capturing complex temporal dependencies. More recent transformer models adopt self-attention to mitigate the challenges of modeling long-term dependencies.

Otherwise, LSTM can also serve as an end-to-end model for TEC forecasting. (Xia, Zhang, et al., 2022) introduces an innovative ED-ConvLSTM network designed to accurately forecast TEC up to 7 days in advance, using data from the International GNSS Service (IGS) from 2005 to 2018. The model demonstrates superior performance over traditional methods like IRI2016 and the 1-day BUAA model (Wang et al., 2018), especially in predicting medium-term TEC variations, although it struggles with localized TEC enhancements and responses to geomagnetic storms. Furthermore, Deep learning approaches are increasingly prevalent in predicting Total Electron Content, (Fubin et al., 2021) introduces a global ionospheric TEC prediction model leveraging deep learning techniques, specifically employing an encoder-decoder architecture with a convolution-optimized Long Short-Term Memory network (ConvLSTM) for spatial and temporal forecasting of global TEC at a spatial resolution of 5° by 2.5° and hourly temporal accuracy. The model achieved a global root mean-square error of less than 1.5 TECU (1 TECU equals 1×10^{16} electrons per square meter) for one-day forecasts under quiet geomagnetic conditions, with a prediction error around 2.5 TECU during mild geomagnetic storms. Predicting TEC during geomagnetic disturbances remains a critical research challenge (Han et al., 2021). (Nath et al., 2023) introduces a hybrid method combining Ensemble Empirical Mode Decomposition (EEMD) and Long Short-Term Memory (LSTM) deep learning for predicting GPS-derived Total Electron Content (TEC) values over various GNSS stations in India, showcasing superior performance over standalone LSTM and EMD-LSTM models in capturing the dynamic ionospheric behavior influenced by multifaceted factors like seasonal, diurnal, and spatial variations, as well as solar geomagnetic conditions.

With the advancement of generative models in time series prediction, a novel theoretical approach to the TEC forecasting challenge has emerged. Diffusion networks excel in generative tasks through iterative refinement (Y. Li et al., 2022), while transformers leverage parallel processing and attention mechanisms to efficiently handle sequential data and capture contextual relationships (Vaswani et al., 2017). Similarly, the Diffusion model offers a novel approach by simulating the stochastic processes underlying ionospheric variations. On the other hand, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model provides a robust statistical baseline for time series forecasting, including the capability to model seasonal patterns effectively (Nobre et al., 2001). (Şentürk, 2020) highlights the superior predictive capabilities of deep learning techniques, specifically Long Short-Term Memory networks, over the ARIMA model in forecasting Total Electron Content at various latitudes using data from Global Ionosphere Maps. This discrepancy is attributed to the ARIMA model's limitation in simultaneously capturing seasonal and trend components effectively. (Yuan et al., 2023) develops an innovative model for accurate 1-day global Total Electron Content (TEC) forecasting by using generative data augmentation and an auto-correlation-based transformer model, significantly enhancing forecast accuracy over existing methods.

This paper aims to explore the capabilities and limitations of these three predictive models: Transformer, Diffusion, and SARIMA, in forecasting TEC values under dif-

ferent geomagnetic conditions. By employing a comparative analysis during periods of strong geomagnetic storms and quiet geomagnetic conditions, we seek to identify the most effective strategies for TEC prediction. Our study utilizes Global Ionosphere Maps (GIMs) provided by the International GNSS Service (IGS) and MIT's TEC data, transformed into a structured grid format for detailed analysis. In doing so, we contribute to the ongoing efforts to enhance the accuracy of TEC predictions, ultimately aiding in the optimization of GNSS and other affected technologies. The Earth's ionosphere has a significant impact on the propagation of radio waves, and during geomagnetic storms (Perreault & Akasofu, 1978; Lakhina & Tsurutani, 2016). Therefore, accurately predicting ionospheric conditions during geomagnetic storms is crucial for ensuring the reliability of these systems. By evaluating model performance in varying geomagnetic scenarios, this research not only highlights the potential of advanced predictive models in space weather forecasting but also provides valuable insights for further development of hybrid or new modeling approaches to tackle the ionosphere's unpredictability.

In this paper, we optimized the efficiency and reduced the training time of our prediction model by utilizing a smaller dataset. Instead of relying on extensive multi-year datasets for training, we demonstrated that using only 1-year and 2-year International GNSS Service (IGS) Total Electron Content (TEC) datasets can still yield robust predictions. Furthermore, acknowledging the significant impact of geomagnetic storms on ionospheric TEC, we conducted a comparative analysis of our model's performance during periods of strong storms versus quiet geomagnetic conditions. We make several key contributions to the field of Total Electron Content prediction:

1. First, we employ a generative probability model for predicting Total Electron Content (TEC) data, marking a significant departure from the predominantly CNN-based approaches in recent TEC prediction research. This innovative method enriches the landscape of TEC forecasting with a fresh perspective.
2. Secondly, we also utilize the SARIMA model which is specifically applied for its capabilities in understanding both seasonal patterns and trends.
3. Thirdly, we compared the prediction accuracy of four different models: Diffusion, SARIMA, Transformer, and C1PG models. The comprehensive comparisons between our models and the existing C1PG model demonstrate superior prediction performance, establishing our approach as a leading method in the field.
4. Lastly, implementing data normalization by using the MinMaxScaler before training the diffusion model significantly smooths the path to model convergence.

2 Methods

In this paper, we aims to explore the capabilities and limitations of these three predictive models: Diffusion, Transformer and SARIMA, in forecasting TEC values under different geomagnetic conditions. By employing a comparative analysis during periods of strong geomagnetic storms and quiet geomagnetic conditions, we seek to identify the most effective strategies for TEC prediction. Our study utilizes Global Ionosphere Maps (GIMs) provided by the International GNSS Service (IGS) and MIT's TEC data, transformed into a structured grid format for detailed analysis.

2.1 Generative Time Series Forecasting: Diffusion

In generative time series forecasting, $X = \{x_1, x_2, \cdots, x_n\}$ is the input TEC time series and $Y = \{y_{n+1}, y_{n+2}, \cdots, y_{n+24})$ is the output target time series, where Y can be generated from latent variables $Z \in \Omega_Z$ that can be drawn from the Gaussian distribution $Z \sim p(Z \mid X)$. The latent distribution can be further formulated as $p_\phi(Z \mid X) = g_\phi(X)$ where g_ϕ denotes a nonlinear function. Then, the data density of the tar-

get series is given by:

$$p_\theta(Y) = \int_{\Omega_Z} p_\phi(Z \mid X)\,(Y - f_\theta(Z))\,dZ,$$

where f_θ denotes a parameterized function. The target time series can be obtained directly by sampling from $p_\theta(Y)$. In Diffusion generative model, time series forecasting is to learn the representation Z that captures useful signals of X, and map the low dimensional X to the latent space with high expressiveness. By taking each input training data and to corrupt it using a multi-step noise process to transform it into a sample from a Gaussian distribution(Y. Li et al., 2022).

The diffusion process is gradually adding Gaussian noise to the input data with the Markov chain. In Markov chains, the probabilities of transitioning from one state to another are fixed and can be represented by a matrix known as the transition matrix(Dhariwal & Nichol, 2021). The process of adding noise is defined through a sequence of conditional distributions, following a Markov chain. The noise addition at each step t is described by:

$$q(X^{(t)} \mid X^{(t-1)}) = \mathcal{N}\left(X^{(t)}; \sqrt{1-\beta_t}X^{(t-1)}, \beta_t I\right) \tag{1}$$

Equation 1 represents a step in the forward diffusion process, specifically the distribution of the data at time t given the data at last time t-1. This notation specifies the conditional distribution of the data at time t given the data at the preceding timestep t-1. It's a probabilistic model of how the data evolves from one step to the next in the diffusion process. where β_t is the variance of the Gaussian noise added at step t, and I is the identity matrix. The sequence of $\boldsymbol{\beta} = \{\beta_1, \cdots, \beta_T \mid \beta_t \in [0,1)\}$ values is chosen to gradually increase the amount of noise over time.

The process starts with the original data $X = X^{(0)} \sim q\left(X^{(0)}\right)$, where $q\left(X^{(0)}\right)$ represents the distribution of the original data. Given the original data $X^{(0)}$, the distribution of the noised data at any step t can be directly obtained as:

$$q(X^{(t)} \mid X^{(0)}) = \mathcal{N}\left(X^{(t)}; \sqrt{\bar{\alpha}_t}X^{(0)}, (1-\bar{\alpha}_t)\,I\right) \tag{2}$$

where $\bar{\alpha}_t = \prod_{s=1}^{t}(1-\beta_s)$ represents the cumulative product of noise reduction factors up to step t.

Through a predetermined Markov chain, Gaussian noise is progressively added to the data in steps. This is governed by an increasing variance schedule $\boldsymbol{\beta} = \{\beta_1, \cdots, \beta_T \mid \beta_t \in [0,1)\}$, which controls the level of noise added at each step. α_t and $\bar{\alpha}_t$ are defined to manage the cumulative effect of the noise.

$$X^{(t)} = \sqrt{\bar{\alpha}_t}X^{(0)} + (1-\bar{\alpha}_t)\,\delta_X \tag{3}$$

At each step t, the original data $X^{(0)}$ is transformed into $X^{(t)}$ by adding noise, represented as denotes standard Gaussian noise. This step illustrates how original data is combined with noise to generate new, diffused data. Here, δ_X represents standard Gaussian noise, and this equation shows how the data $X(t)$ is a mixture of the scaled original data and added noise.

Similarly, for the output series Y, the diffusion process applies, with possibly a different scale parameter ω to adjust the noise levels:

$$Y^{(t)} = \sqrt{\overline{\alpha'}}Y^{(0)} + \sqrt{1-\overline{\alpha'}_t}\delta_Y \tag{4}$$

where $\overline{\alpha'}_t$ is adjusted by ω for the output series.

The figure 1 showcases the generative diffusion model's workflow, with the left side representing the training phase, where the model learns from data, and the right side

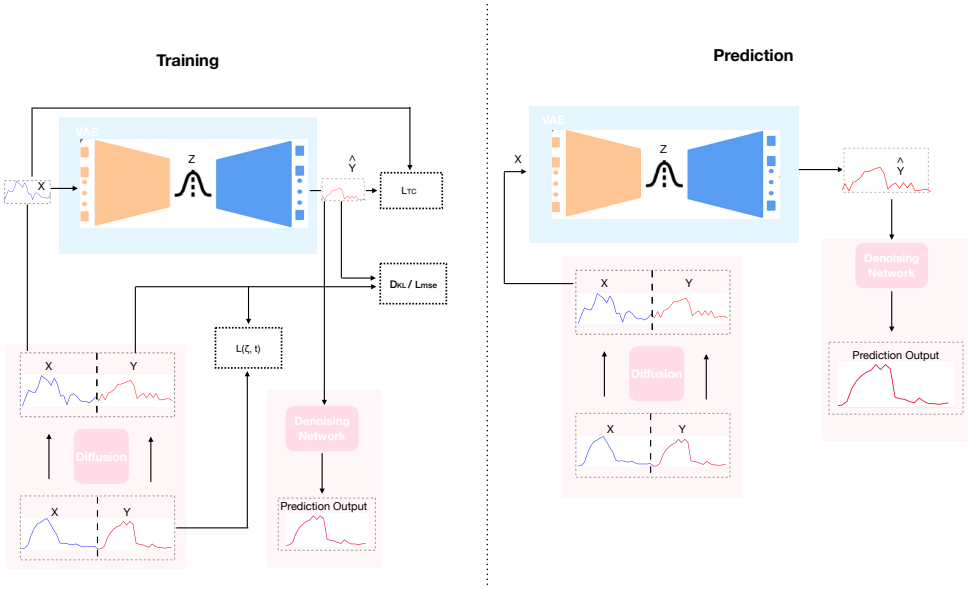


Figure 1. Training and Prediction Pipelines of the Diffusion Model

illustrating the prediction phase, where the model applies what it's learned to make forecasts.

The Loss function is componented by 4 parts. The $\mathcal{D}_{\mathrm{KL}}$ encourages the distribution generated by the model to be as close as possible to the true target distribution. $L(\zeta, t)$ aims to "clean" the generated target series by aligning them more closely with the uncorrupted original data. The Total Correlation loss L_{TC} from the latent variables of VAE model. the Mean Square Error (MSE) loss L_mse between $\bar{Y}_t$ and Y_t. Thus, The Loss function for training this generative diffusion prediction model is following:

$$\mathcal{L} = \psi \cdot \mathcal{D}_{\mathrm{KL}}\left(q\left(Y^{(t)}\right) \| p_\theta\left(\widehat{Y}^{(t)}\right)\right) + \lambda \cdot L(\zeta, t) + \gamma \cdot L_{\mathrm{TC}} + L_{\mathrm{mse}}\left(\widehat{Y}^{(t)}, Y^{(t)}\right) \tag{5}$$

$\psi, \lambda, \cdot$ are weight parameters that adjust the importance of this term. $\mathcal{D}_{\mathrm{KL}}$ means the Kullback-Leibler divergence (DKL)(Pérez-Cruz, 2008), which measures the difference between the actual target series distribution $q\left(Y^{(t)}\right)$ and the predicted target series distribution by the model $p_\theta\left(\widehat{Y}^{(t)}\right)$ at time t. L_{TC} stands for the Total Correlation loss of the latent variables, used for disentangling the latent variables. Minimizing the total correlation among latent variables encourages the model to learn more independent feature representations, which helps improve the quality and interpretability of the generation. L_mse calculates the mean square error (MSE) between the model's predicted series $\bar{Y}_t$ and the actual target series Y_t.

2.2 SARIMA

Seasonal AutoRegressive Integrated Moving Average (SARIMA) is widely used for forecasting and understanding seasonal time series data, where observations show regular patterns at fixed intervals, such as daily, monthly, or quarterly seasonality (Valipour, 2015). The SARIMA model combines three core components: AutoRegressive (AR), Differencing, and Moving Average (MA), addressing both seasonal and non-seasonal dynamics within the dataset. These components work together to adjust the model based on past values (AR), the difference between observations (D), and past forecast errors (MA), thereby accounting for complexity in both seasonal and non-seasonal fluctuations. The

SARIMA model is mathematically represented as follows, where: y_t is the time series data at time t, B is the lag operator, ϕ_i and Φ_i are the coefficients of the non-seasonal and seasonal AR terms, θ_i and Θ_i are the coefficients of the non-seasonal and seasonal MA terms, ε_t is the error terms at time t, assumed to be white noise.

$$\left(1-\sum_{i=1}^{p}\phi_i B^i\right)\left(1-\sum_{i=1}^{P}\Phi_i B^{is}\right)(1-B)^d\left(1-B^s\right)^D y_t=\left(1+\sum_{i=1}^{q}\theta_i B^i\right)\left(1+\sum_{i=1}^{Q}\Theta_i B^{is}\right)\varepsilon_t \tag{6}$$

AR Component: The autoregressive non-seasonal component represented by $\sum_{i=1}^{p}\phi_i B^i$, and seasonal component respresentes by $\sum_{i=1}^{P}\Phi_i B^{is}$. AR component captures the relationship between the current observation and a certain number of lagged observations (previous values in the time series). The B term represents the backshift operator is commonly used in time series analysis $\left(B^i y_t = y_{t-i}\right)$. It represents the lag operator, which shifts the time series backward by a certain number of time period. The parameters p, P denote the unseasonal and seasonal autoregressive orders, s is the seasonal period.

Differencing Component: SARIMA model need the Time series data is stable. Differencing is a method which is used for make a time series more stationary. The parameters d, D indicate the number of non-seasonal and seasonal differences.

MA Component: $\sum_{i=1}^{q}\theta_i B^i, \sum_{i=1}^{Q}\Theta_i B^{is}$ denote the seasonal moving average part and non-seasonal part, which models the relationship between an observation and its seasonal lagged forecast errors. The parameters q, Q denote the order of the moving average part, which captures the relationship between an observation and a specified number of lagged forecast errors.

In summary, SARIMA model parameters include orders for the autoregressive (AR), differencing (I), and moving average (MA) components, both on a non-seasonal and seasonal basis, along with the length of the seasonal cycle. These parameters collectively characterize the structure and dynamics of the time series, facilitating its modeling and forecasting.

2.3 Transformer

Transformer is composed of encoder and decoder and Positional Encoding. The encoder is composed of alternating layers that feature multi-headed self-attention (MSA) mechanisms and multilayer perceptron (MLP) blocks. Prior to each block, layer normalization (LN) is applied, enhancing the model's efficiency and stability during training. Furthermore, residual connections are introduced following each block to facilitate the flow of gradients through the network, significantly aiding in the prevention of the vanishing gradient problem often encountered in deep networks. The decoder is composed of dimension expansion, three CNN blocks and an up sampling layer. In the CNN bocks, batch normalization (BN) is used before every CNN.

Multi-Headed Self-Attention (MSA) allows the model to jointly attend to information from different representation subspaces at different positions (Vaswani et al., 2017). A single attention head inhibits this capability. The basic structure is the same as self-attention. However, MSA is divided into multiple heads, and the self-attention calculation is performed in parallel. The output vectors from the different heads are then concatenated together, enabling the model to jointly attend to information from different representation subspaces at different positions. This allows different heads to learn different levels of knowledge. This part is used to inject some information about the relative or absolute position tokens from the TEC sequence. To this end, add positional encodings at the bottoms of the encoder and decoder stacks. The positional encodings have the same dimension dmodel as the embeddings so that the two can be summed. here, the pos is the position and i is the dimension. That is, each dimension of the positional

encoding corresponds to a sinusoid.

$$\text{MSA}\ (Q, K, V) = \ \text{Concat}\ (\ \text{head}\ _1, \ldots,\ \text{head}\ _h)\, W^O \tag{7}$$

where head = Attention $\left(QW_i^Q, KW_i^K, VW_i^V\right)$, and the projections are parameter matrices $W_i^Q \in R^{d_{\text{model}} \times d_k}, W_i^K \in R^{d_{\text{model}} \times d_k}, W_i^V \in R^{d_{\text{model}} \times d_v}, W^O \in R^{hd_v \times d_{\text{model}}}$, $d_{\text{model}} = 512, d_v, d_k = 64$, and $h = 8$.

Layer Normalization (LN) works by normalizing the inputs across the features for a single training example(Xiong et al., 2020; Ba et al., 2016). It addresses the limitations of applying batch normalization (BN) to recurrent neural networks (RNNs), which stems from BN's dependence on batch-wise statistics that are not straightforward to compute in RNNs due to their sequential nature (Bjorck et al., 2018; Luo et al., 2018). LN circumvents this issue by normalizing the inputs across each neuron within a layer based on the mean and variance computed from a single training instance, as opposed to an entire batch. This normalization strategy ensures that LN can be seamlessly integrated into RNNs by calculating normalized statistics at each timestep, thereby harmonizing the internal state distributions throughout the network's depth. Specifically, LN achieves this by aggregating the statistics across all hidden neurons within the same layer, enhancing the training stability and speed of neural architectures, especially those prone to internal covariate shifts.

Then, position embeddings are added to the image embeddings to retain positional information, which is given as follows, where pos is the position and i is the dimension.

$$\boldsymbol{PE}_{(pos,2i)} = \sin\left(pos/10,000^{2i/d_{\text{model}}}\right) \tag{8}$$

$$\boldsymbol{PE}_{(pos,2i+1)} = \cos\left(pos/10,000^{2i/d_{\text{model}}}\right) \tag{9}$$

3 Forecasting Performance Evaluation and Results Analysis

3.1 Data Source and Processing

In this research, the Global Ionosphere Maps (GIMs) provided by the International GNSS Service (IGS) were utilized. Each model leveraged a comprehensive dataset spanning one year as the training data. Specifically, the training data was derived from MIT's TEC datasets, which were subsequently transformed into a grid format consisting of meshes sized 71 by 73. For the purpose of our TEC prediction analysis, we focused on a single grid cell corresponding precisely to the geographical coordinates of 114.3°E longitude and 30.5°N latitude. This selected grid cell, representative of the target region, measures 5 degrees in longitude and 2.5 degrees in latitude, mirroring the dimensions of one standard grid cell in our dataset.

Min-Max Normalization, also known as feature scaling, is a technique used to normalize the range of features in data(Singh & Singh, 2020; Sola & Sevilla, 1997). This method scales the range of the feature to be between a specified minimum and maximum value, often 0 and 1 . Here's how it is implemented in data processing. The formula for Min-Max Normalization is as follows:

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{10}$$

Where, X_{norm} is the normalized value. X is the original value. $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the feature across the dataset, respectively. First, identify the minimum and maximum values for each column in the dataset. This step involves determining the smallest and largest values that each feature assumes across all data points. Next, apply the Min-Max normalization formula to each data point, scaling the feature

Table 1. Classification of Geomagnetic Storms Based on Dst Index

Storm Level	Dst Index Range
Moderate Storm	$\mathrm{Dst} < -50\,\mathrm{nT}$
Intense Storm	$-100\,\mathrm{nT} \leq \mathrm{Dst} < -50\,\mathrm{nT}$
Extreme Storm	$-200\,\mathrm{nT} \leq \mathrm{Dst} < -100\,\mathrm{nT}$

values to fall within a specified range, typically between 0 and 1. Finally, repeat these steps for all columns in the dataset to ensure that each feature is scaled similarly, promoting uniformity in data scale across the dataset. This systematic approach facilitates model training and contributes to more reliable performance.

The Dst (Disturbance Storm Time) index is an index of magnetic activity derived. Geomagnetic storms are identified by a significant decrease in the Dst index. The classification of storms into moderate, intense, or extreme categories is based on how low the Dst index drops, as outlined in the Table 3.1. Specifically, a geomagnetic storm is generally considered to occur when the Dst index falls below -50 nT.

We will analyze the prediction accuracy during periods of strong geomagnetic activity and quiet geomagnetic conditions. For instances of strong geomagnetic storms, our prediction analysis will focus on the period from 2015 to 2017 and evaluated the prediction data which is in strong geomagnetic storm periods during 2015-2017. We train our models (Diffusion, ARIMA and Transformer) on data from the year before the one we're predicting. So, we use 2014 data to predict 2015, 2015 data for 2016, and 2016 data for 2017. For each year we're predicting, we only use data from the immediate previous year to train our models. We split this yearly data into two parts: 20% for validation and 80% for training. This approach allows us to assess the models' performance in forecasting TEC values under varying geomagnetic influences, with a particular emphasis on their accuracy during significant geomagnetic disturbances.

For the quiet geomagnetic condition, there were only four days in 2019 when the Dst index exceeded -50 nT, meaning that for most of the year, 2019 had relatively low levels of geomagnetic activity and was therefore chosen as a study object for quiet geomagnetic conditions.

Figure 2 contains two subfigures depicting the change curve of the loss function for the diffusion model and Transformer model as the number of epochs increases during the training process. We observed a clear trend of decreasing training loss as the number of training epochs increased. The two models converged to a stable loss value, demonstrating the success of our training and validation process. In the legends, 'diffusion_1_train', 'diffusion_2_train', 'diffusion_3_train' refer to the training loss of diffusion models which is trained with 2014,2015,2016 data, respectively, while 'diffusion_1_vali', 'diffusion_2_vali', 'diffusion_3_vali' denote their corresponding validation loss. For quiet geomagnetic condition, the legends 'diffusion_4_train' and 'diffusion_4_vali' are used for recording the training loss and validation loss of the model trained by 2017-2018. By implementing the Min-Max Normalization technique for data processing, the Diffusion model reaches convergence within 10 epochs, markedly fewer than the Transformer model, which necessitates a larger number of epochs when normalization is not applied.

We analysis the prediction ability by analyzing the root-mean-square error (RMSE) and Daily Pearson correlation coefficient of the predicted values by each model each month under these relatively geomagnetic conditions The formula for the Daily Pearson correlation coefficient cc, which measures the linear relationship between two continuous ob-

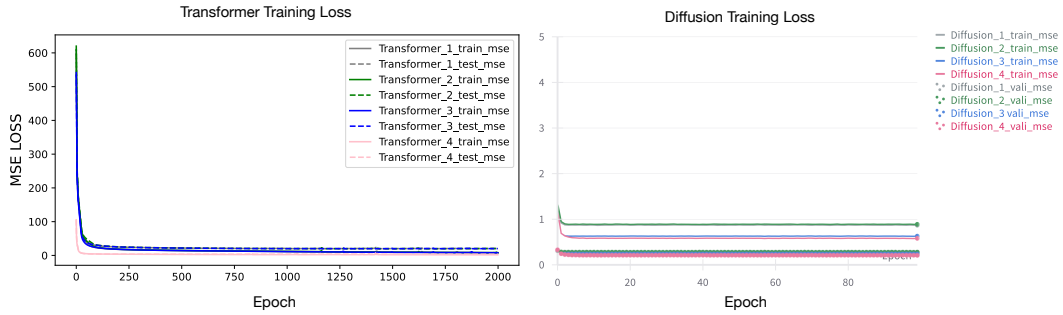


Figure 2. Training and Validation Losses of Diffusion Models and Transformer Models

served TEC variables X and forecasting Y, is given by:

$$cc_{xy} = \frac{\mathrm{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} \tag{11}$$

Here, cc_{xy} denotes the Pearson correlation coefficient between X and Y, $\mathrm{cov}(X, Y)$ represents the covariance of X and Y, σ_X is the standard deviation of variable X, σ_Y is the standard deviation of variable Y. The covariance is computed as:

$$\mathrm{cov}(X, Y) = E\left[(X - \mu_X)\,(Y - \mu_Y)\right] \tag{12}$$

where, E is the expected value operator, μ_X is the mean of X, μ_Y is the mean of Y.

Root-mean-square error is also used to estimate the forecasting performance of the model. The lower the RMSE value, the better the model's accuracy in prediction and the formula for calculating RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(TEC_{\mathrm{ori}} - TEC_{\mathrm{pred}}\right)^2} \tag{13}$$

where N is the total number of data samples, TEC_{ori} and TEC_{pred} are the observed value and forecasting value, respectively.

3.2 Prediction Analysis

This Figure 3 and Figure 4 assemble a set of five subplots for time series prediction analysis, comparing and evaluating the prediction performance of four different models (Diffusion, SARIMA, Transformer and C1PG) against the actual observed values during quiet geomagnetic conditions and strong geomagnetic conditions. For each figure, the first subplot displays the predictions made by the Diffusion model (depicted as a red line) versus the actual observed values (in blue), both measured in TECU units. The RMSE between these predictions and observations is calculated and displayed above the plot within an orange bounding box. The subsequent second, third, and fourth subplots are dedicated to showcasing the predictions from the SARIMA model, the Transformer model, and the C1PG model, respectively. Each model's predictions are again represented by a red line and placed alongside the actual observed values, with the corresponding RMSE calculation for each model provided for comparative analysis. The final, fifth subplot diverges to exclusively represent the variations of the Dst index throughout the designated prediction period, offering insight into the geomagnetic conditions prevailing at the time.

Figure 3 presents a comparative visualization of each model's ability of forecast accurately in quiet geomagnetic storm periods from March 20 to March 27, 2019, and from

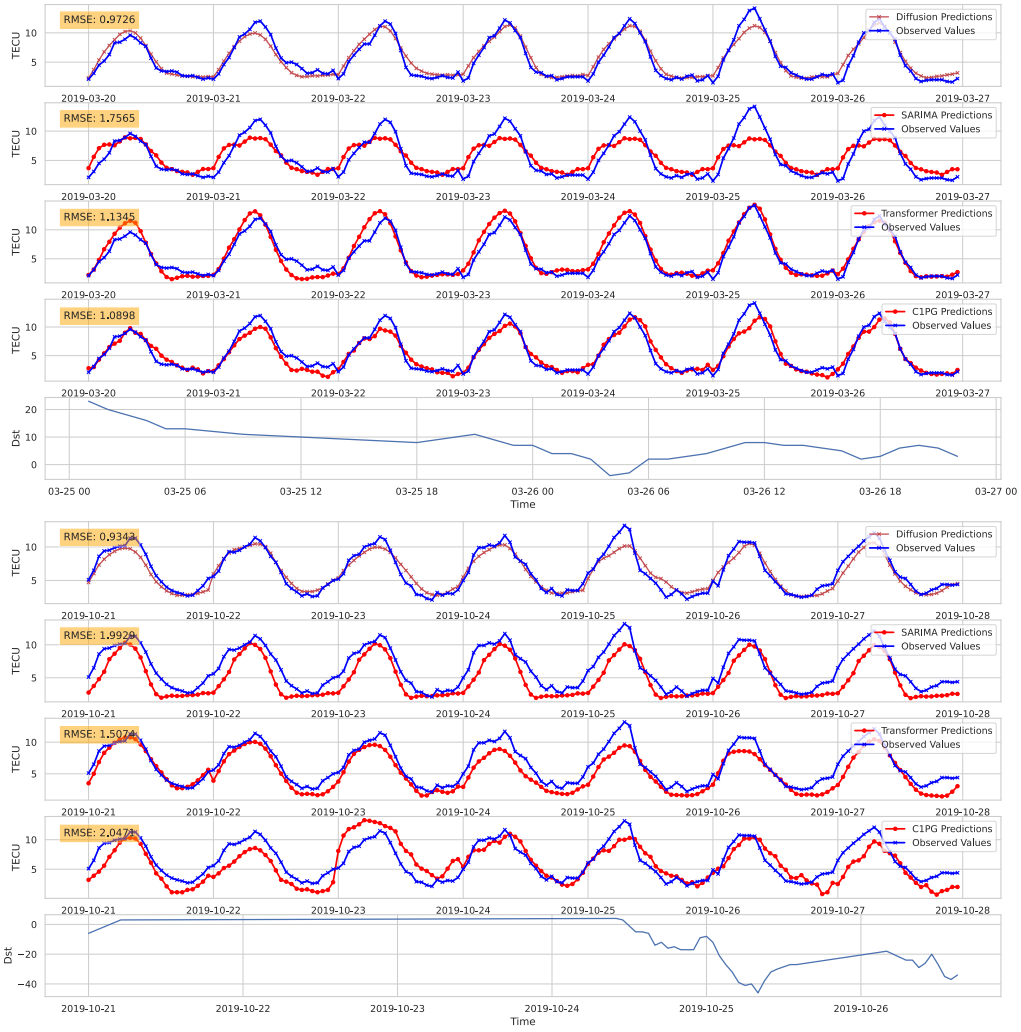


Figure 3. Comparative Forecasting Performance of Models During Quiet Geomagnetic Periods in 2019

October 21 to October 28, 2019. The Diffusion model leads in performance, demonstrating the highest precision with the lowest RMSE of 0.9726 TECU, closely followed by the Transformer model at an RMSE of 1.08 TECU.

Figure 4 illustrates the comparative forecasting performance of four models during periods of intense geomagnetic storms, specifically from September 1st to 7th, 2016, and from August 28th to September 5th, 2017. Among the models, the Diffusion model stands out for its superior accuracy, achieving the lowest RMSE of 0.9343 TECU. In contrast, the C1PG model exhibits the least accurate forecasts with an RMSE of 2.047 TECU.

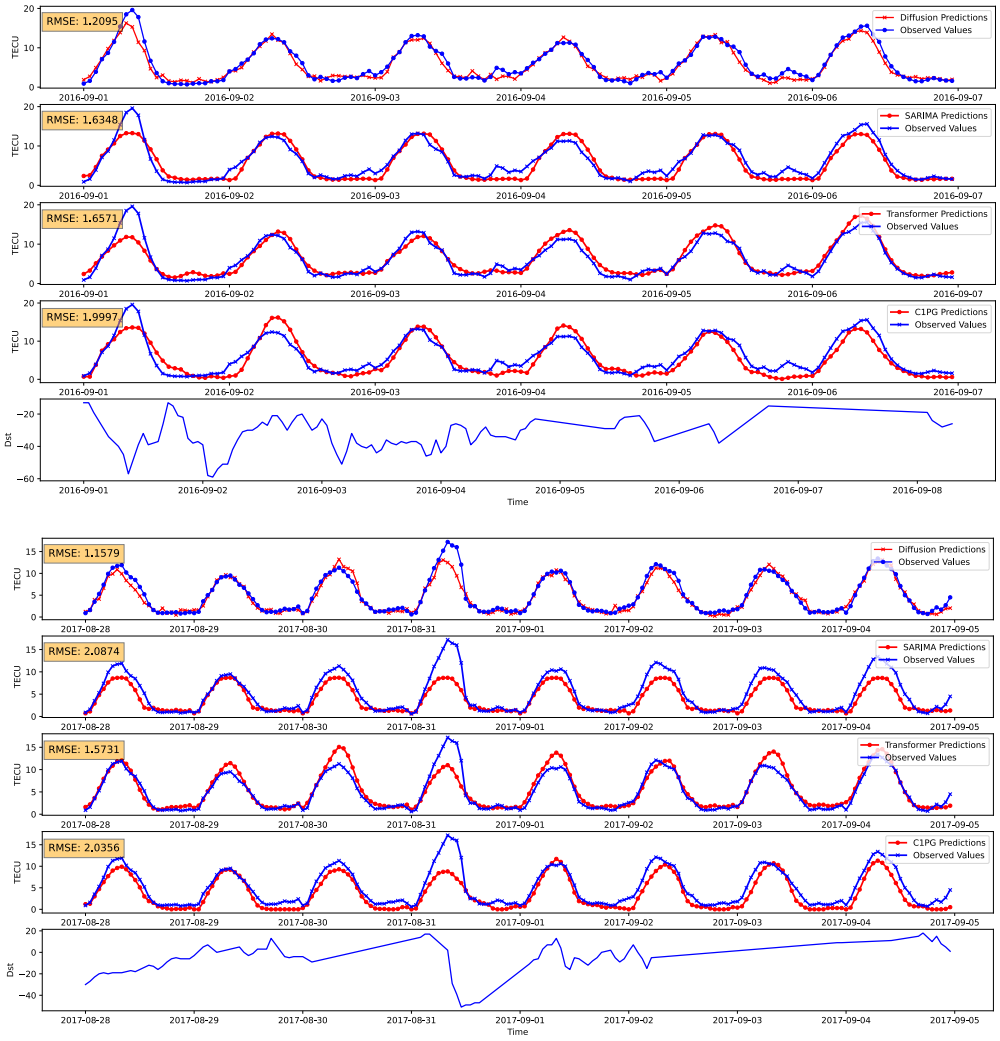


Figure 4. Strong geomagnetic storms

Figure 5 present the calculated monthly RMSE values of different models throughout the year 2019. This visualization allows for a direct comparison of each model's performance in terms of prediction accuracy on a month-by-month basis. This line chart visualizes the RMSE values for four different models: Diffusion, C1PG, Transformer, and SARIMA, over 2019 year. Each line represents the fluctuation in RMSE values for a specific model from January to December. The y-axis indicates the RMSE value, which mea-

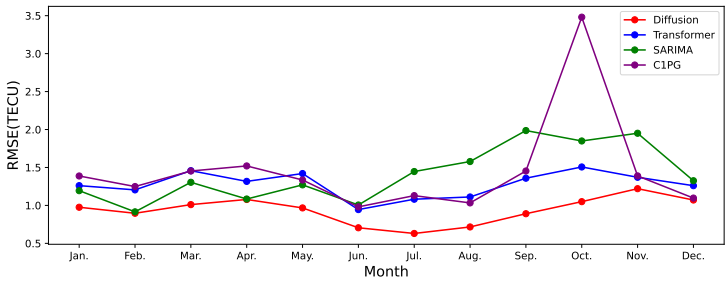


Figure 5. Comparison of Monthly RMSE Values from different four Models in 2019 during quiet geomagnetic periods.

sures the models' prediction errors, with lower values indicating more accurate predictions. The x-axis represents the months of the year. The Diffusion model consistently demonstrates the lowest RMSE values across most months, indicating it generally offers the most accurate predictions among the four models. This suggests that the Diffusion model is particularly effective in handling the complexity and variability of the data. All models exhibit some level of seasonal variability in their prediction accuracy. For instance, the RMSE values tend to be lower during the summer months (June, July, August) for the Diffusion and C1PG models, suggesting these models perform better in predicting during this period. Conversely, the SARIMA model's accuracy dips significantly in the latter half of the year, especially from August to November.The C1PG model shows the highest variability in RMSE values, with a significant spike in October. This suggests that the model might struggle with certain types of data or events occurring in that month. The Transformer and SARIMA models exhibit moderate RMSE values, with the Transformer model showing a slight increase in prediction error towards the end of the year. Overall, the line chart effectively communicates the comparative performance of the four models over time, highlighting their strengths and weaknesses in forecasting accuracy across different times of the year. It provides valuable insights into the models' behavior under varying conditions, which can inform decisions on model selection or improvement strategies.

Table 2 presents the distribution of the Cross-Correlation Coefficient (CC) range for four models (Diffusion, Transformer, C1PG and SARIMA) during both Quiet (Dst¿-50nT) and Strong (Dst¡-50nT) geomagnetic storm periods. The CC ranges are divided into four categories: CC below 0.6, CC between 0.6 and 0.9, CC between 0.9 and 0.95, and CC between 0.95 and 1. The table quantifies the percentage of instances falling into each of these ranges for the respective models.

Table 2. CC Range Distribution for Different Models in Quiet and Strong Periods

Geomagnetic Strom Periods	CC Range	Diffusion	Transformer	SARIMA	C1PG
Quiet periods	$CC < 0.6$	0.00%	0.00%	0.00%	0.28%
	$0.6 \leq CC < 0.9$	1.10%	2.49%	3.59%	14.36%
	$0.9 \leq CC < 0.95$	13.26%	21.55%	20.99%	33.70%
	$0.95 \leq CC \leq 1$	**85.64%**	75.97%	75.41%	51.66%
Strong periods	$CC < 0.6$	0.00%	0.81%	1.62%	1.61%
	$0.6 \leq CC < 0.9$	6.05%	11.70%	11.34%	10.08%
	$0.9 \leq CC < 0.95$	18.55%	13.31%	19.83%	16.53%
	$0.95 \leq CC \leq 1$	**75.41%**	74.20%	67.21%	71.77%

From the table, it's observed that none of the models have instances with a CC lower 0.6 during magnetically quiet periods, indicating a generally strong performance across all models in terms of correlation with the observed data. For quiet periods, the table shows that a significant majority of the predictions made by the Diffusion model fall into the highest accuracy range ($0.95 < CC < 1$), with 85.64% of its predictions exhibiting nearly perfect correlation with the observed data. The Transformer, SARIMA, and C1PG models also show a strong preference for the higher accuracy ranges, but with a lesser proportion of their predictions achieving the highest CC range compared to the Diffusion model. During strong geomagnetic periods, the proportion of predictions falling into the highest CC range slightly decreases for all models, indicating a generally lower prediction accuracy under these more challenging conditions. However, a substantial majority of predictions from each model still achieve a CC of 0.95 or higher, demonstrating their robustness. The Diffusion model continues to lead in prediction accuracy, with 75.41% of its predictions in the highest CC range, although the differences in performance among the models are somewhat reduced compared to the quiet period scenario. The majority of instances for all four models fall within the highest CC range of 0.95 to 1, with the Diffusion model showing the highest percentage at 85.64% during magnetically quiet periods and 75.41% during magnetically strong periods, followed by the Transformer model at 75.97% and 74.20%, and the SARIMA model at 75.41% and 67.21%. This demonstrates that while all models perform well, the Diffusion model tends to have a slightly higher proportion of instances with very strong to almost perfect correlation with the observed data during magnetically strong periods. For the range of 0.6 to 0.9, the Diffusion model shows the lowest percentage of instances at 1.10%, followed by the Transformer model at 2.49%, and the SARIMA model has the highest at 3.59%. This suggests that the SARIMA model has slightly more instances with moderate correlation compared to the other models. In the CC range of 0.9 to 0.95 during magnetically quiet periods, the Diffusion model has 18.55% of its instances, which is less than the Transformer model at 21.55% and the SARIMA model at 20.99%. This indicates that both the Transformer and SARIMA models have a higher proportion of their correlations in the high range, but not the highest, compared to the Diffusion model. Overall, the table highlights the strong performance of these models in correlating with observed data during periods of magnetic quiet, with diffusion model exhibiting best.

In order to analysis the prediction ability of models in strong geomagnetic storm periods, we identified the days between 2015 to 2017 when the Dst index fell below -50nT, and calculate the Daily Pearson Correlation Coefficient of prediction data and Observation data. In Figure 6, we calculated the daily Pearson CC of the strong geomagnetic storm periods happened during 2015-2017 year and presents a scatter plot. Figure 7 presents the calculated Daily CC values of different four models throughout the year 2019. This visualization allows for a direct comparison of each model's performance in terms of prediction accuracy on a month-by-month basis.

4 Conclusions

In conclusion, this study presents a significant advancement in the domain of ionospheric forecasting by introducing and evaluating three innovative models: the Diffusion model, the SARIMA model, and the Transformer model. Focused on enhancing the accuracy of predictions during geomagnetic storms, these models offer a fresh perspective on forecasting ionospheric conditions. The comparative analysis between the performance of generative Diffusion model and the established C1PG model, utilizing just one year of training data, underscores the efficacy and potential of the Diffusion model in the realm of geomagnetic forecasting. This feature is especially valuable in situations where extensive historical data are limited or hard to acquire, such as the infrequent occurrences of geomagnetic storms. Unlike Transformer and LSTM models, which typically need large datasets for effective training and accuracy, our research underscores the impressive po-

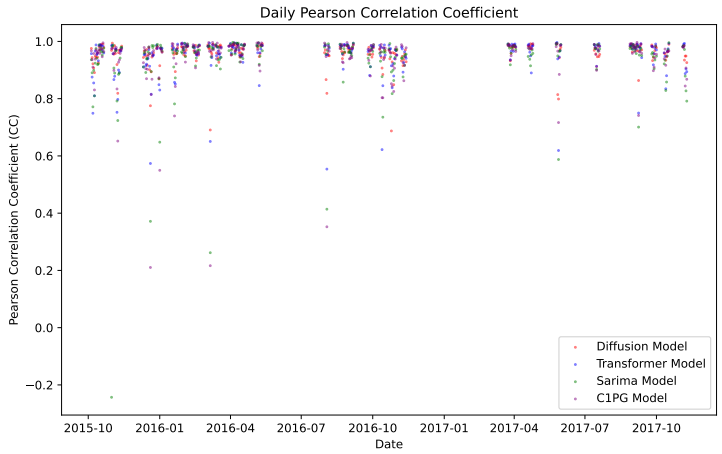


Figure 6. Daily Pearson Correlation Coefficients for strong geomagnetic storm periods, 2015-2017

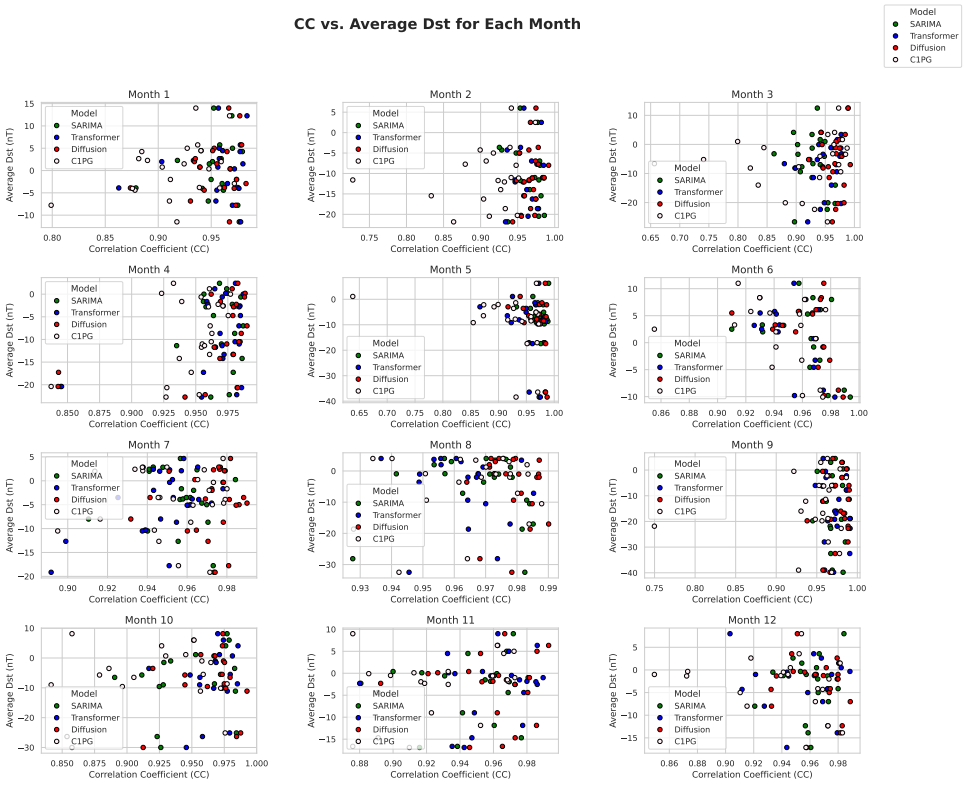


Figure 7. Daily CC values for four predictive models in 2019.

tential of generative modeling techniques, notably the Diffusion model, for forecasting ionospheric disturbances with limited data. Otherwise, SARIMA model is distinguished by its straightforward training process and strong operability, requiring minimal computational resources. Remarkably, it achieves performance comparable to the Transformer and surpasses the C1PG model, demonstrating a 75.66% accuracy in quiet periods, significantly higher than the 51.66% accuracy achieved by C1PG.

Specifically, the use of the Diffusion model has demonstrated how generative probability models can significantly improve forecasting outcomes, even with limited training data. Furthermore, the incorporation of data normalization techniques and the strategic application of the SARIMA model for capturing seasonal patterns and trends have collectively contributed to the robustness and efficacy of the proposed forecasting framework. Our study highlights the impressive potential of generative models, especially the Diffusion model, in accurately forecasting ionospheric disturbances. This advancement represents a significant step in the field of ionospheric prediction with only 1-year training data.

5 Open Research

The global ionosphere maps (GIMs) TEC data sets from global navigation The datasets utilized in this manuscript, specifically the Global Ionosphere Maps (GIMs) Total Electron Content (TEC) data, were sourced from global navigation satellite systems (GNSS). These data are publicly available on the website `http://pub.ionosphere.cn/products/daily/`. Additionally, the Disturbance Storm Time (Dst) index data, crucial for our analysis, can be accessed via `https://omniweb.gsfc.nasa.gov/form/dx1.html`. The C1PG prediction data, referenced in our study and initially proposed by (Schaer & helvétique des sciences naturelles. Commission géodésique, 1999), originates from a global TEC prediction model developed by the Center for Orbit Determination in Europe (CODE). This data can be found at `https://cddis.nasa.gov/archive/gnss/products/ionex/`.

References

Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). Layer normalization. *arXiv preprint arXiv:1607.06450*.

Béniguel, Y. (2002). Global ionospheric propagation model (gim): A propagation model for scintillations of transmitted signals. *Radio Science*, *37*(3), 1–14.

Bjorck, N., Gomes, C. P., Selman, B., & Weinberger, K. Q. (2018). Understanding batch normalization. *Advances in neural information processing systems*, *31*.

Dhariwal, P., & Nichol, A. (2021). Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, *34*, 8780–8794.

Fubin, Z., Chen, Z., Cheng, W., Jiaqi, Z., Yi, L., Guozhen, X., & Zhengyu, Z. (2021). Global ionospheric tec prediction based on deep learning. *Chinese Journal of Radio Science*, *36*(4), 553–561.

Han, Y., Wang, L., Fu, W., Zhou, H., Li, T., & Chen, R. (2021). Machine learning-based short-term gps tec forecasting during high solar activity and magnetic storm periods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 115–126.

Lakhina, G. S., & Tsurutani, B. T. (2016). Geomagnetic storms: historical perspective to modern view. *Geoscience Letters*, *3*, 1–11.

Lastovicka, J., Urbar, J., & Kozubek, M. (2017). Long-term trends in the total electron content. *Geophysical Research Letters*, *44*(16), 8168–8172.

Lean, J., Emmert, J., Picone, J., & Meier, R. (2011). Global and regional trends in ionospheric total electron content. *Journal of Geophysical Research: Space Physics*, *116*(A2).

Li, Y., Lu, X., Wang, Y., & Dou, D. (2022). Generative time series forecasting with

diffusion, denoise, and disentanglement. *Advances in Neural Information Processing Systems*, *35*, 23009–23022.

Li, Z., Wang, N., Wang, L., Liu, A., Yuan, H., & Zhang, K. (2019). Regional ionospheric tec modeling based on a two-layer spherical harmonic approximation for real-time single-frequency ppp. *Journal of Geodesy*, *93*, 1659–1671.

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A*, *379*(2194), 20200209.

Luo, P., Wang, X., Shao, W., & Peng, Z. (2018). Towards understanding regularization in batch normalization. *arXiv preprint arXiv:1809.00846*.

Nath, S., Chetia, B., & Kalita, S. (2023). Ionospheric tec prediction using hybrid method based on ensemble empirical mode decomposition (eemd) and long short-term memory (lstm) deep learning model over india. *Advances in Space Research*, *71*(5), 2307–2317.

Nobre, F. F., Monteiro, A. B. S., Telles, P. R., & Williamson, G. D. (2001). Dynamic linear model and sarima: a comparison of their forecasting performance in epidemiology. *Statistics in medicine*, *20*(20), 3051–3069.

Pérez-Cruz, F. (2008). Kullback-leibler divergence estimation of continuous distributions. In *2008 ieee international symposium on information theory* (pp. 1666–1670).

Perreault, P., & Akasofu, S. (1978). A study of geomagnetic storms. *Geophysical Journal International*, *54*(3), 547–573.

Ratnam, D. V., Vishnu, T. R., & Harsha, P. B. S. (2018). Ionospheric gradients estimation and analysis of s-band navigation signals for navic system. *IEEE Access*, *6*, 66954–66962.

Reinisch, B. W., & Galkin, I. A. (2011). Global ionospheric radio observatory (giro). *Earth, planets and space*, *63*, 377–381.

Schaer, S., & helvétique des sciences naturelles. Commission géodésique, S. (1999). *Mapping and predicting the earth's ionosphere using the global positioning system* (Vol. 59). Institut für Geodäsie und Photogrammetrie, Eidg. Technische Hochschule ....

Şentürk, E. (2020). Ionospheric tec prediction performance of arima and lstm methods in different space weather conditions. *Intercontinental Geoinformation Days*, *1*, 32–35.

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*, 105524.

Sola, J., & Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Transactions on nuclear science*, *44*(3), 1464–1468.

Valipour, M. (2015). Long-term runoff study using sarima and arima models in the united states. *Meteorological Applications*, *22*(3), 592–598.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wang, C., Xin, S., Liu, X., Shi, C., & Fan, L. (2018). Prediction of global ionospheric vtec maps using an adaptive autoregressive model. *Earth, Planets and Space*, *70*, 1–14.

Xia, G., Liu, M., Zhang, F., & Zhou, C. (2022). Caitst: Conv-attentional image time sequence transformer for ionospheric tec maps forecast. *Remote Sensing*, *14*(17), 4223.

Xia, G., Zhang, F., Wang, C., & Zhou, C. (2022). Ed-convlstm: A novel global ionospheric total electron content medium-term forecast model. *Space Weather*, *20*(8), e2021SW002959.

Xiong, R., Yang, Y., He, D., Zheng, K., Zheng, S., Xing, C., ... Liu, T. (2020). On layer normalization in the transformer architecture. In *International conference on machine learning* (pp. 10524–10533).

Yuan, Y., Xia, G., Zhang, X., & Zhou, C. (2023). Synthesis-style auto-correlation-based transformer: A learner on ionospheric tec series forecasting. *Space Weather*, *21*(10), e2023SW003472.

figure1.

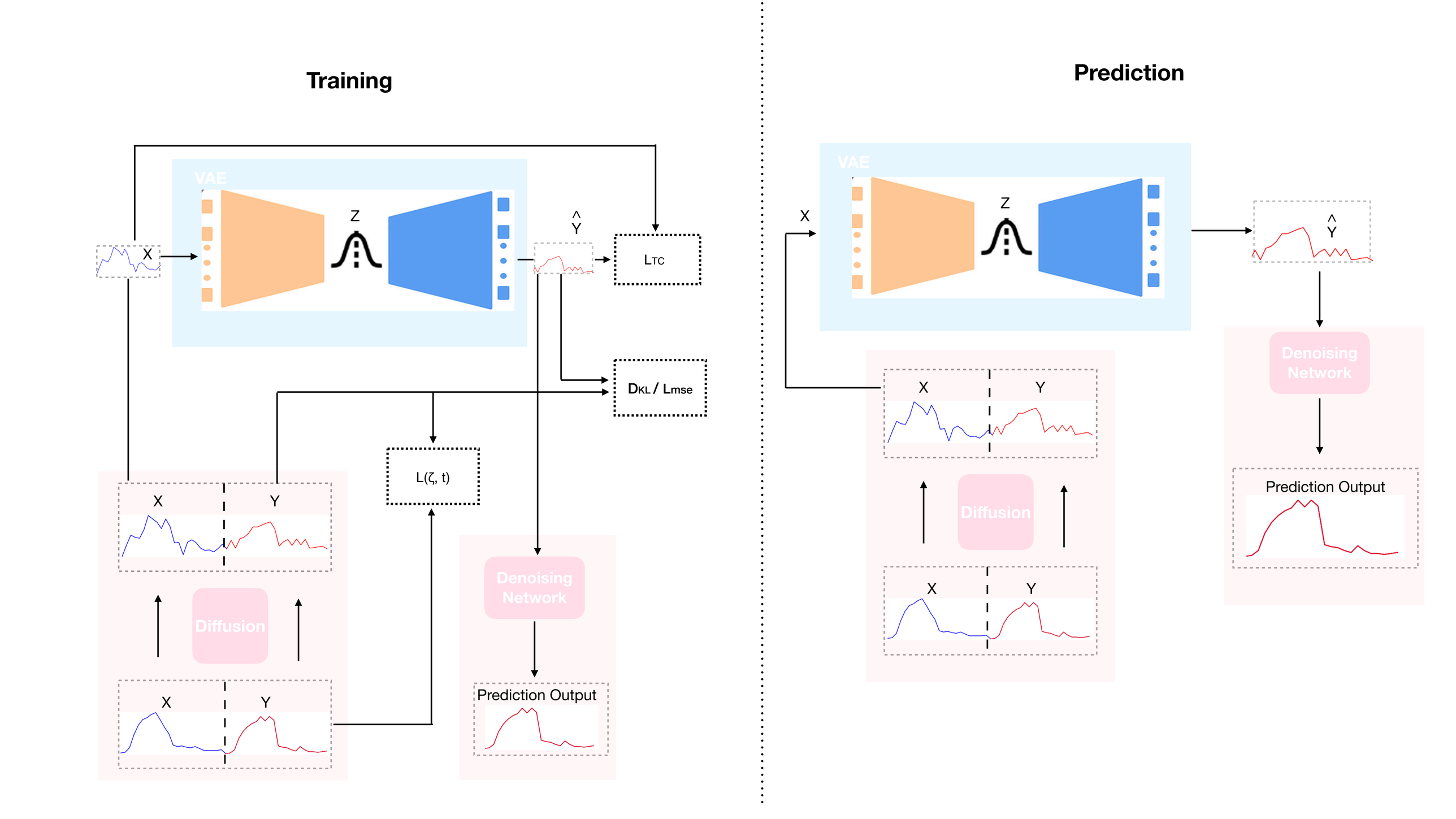
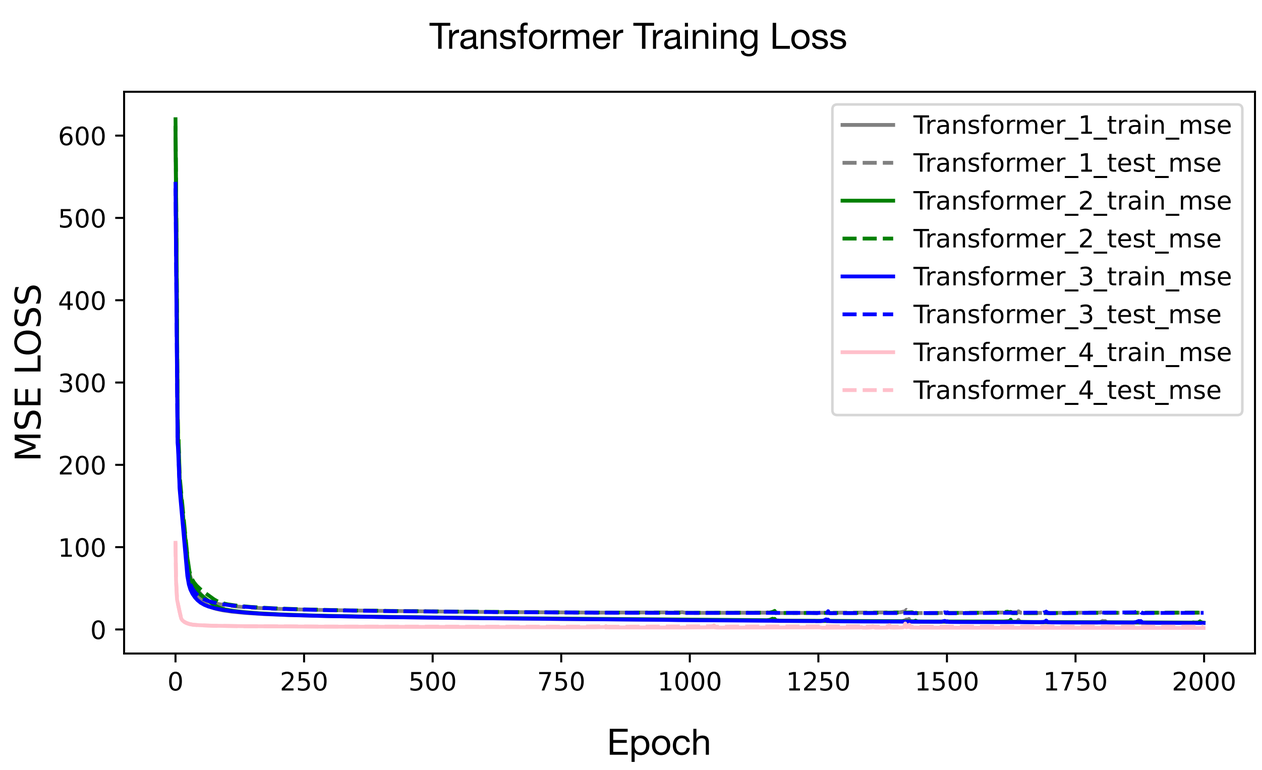
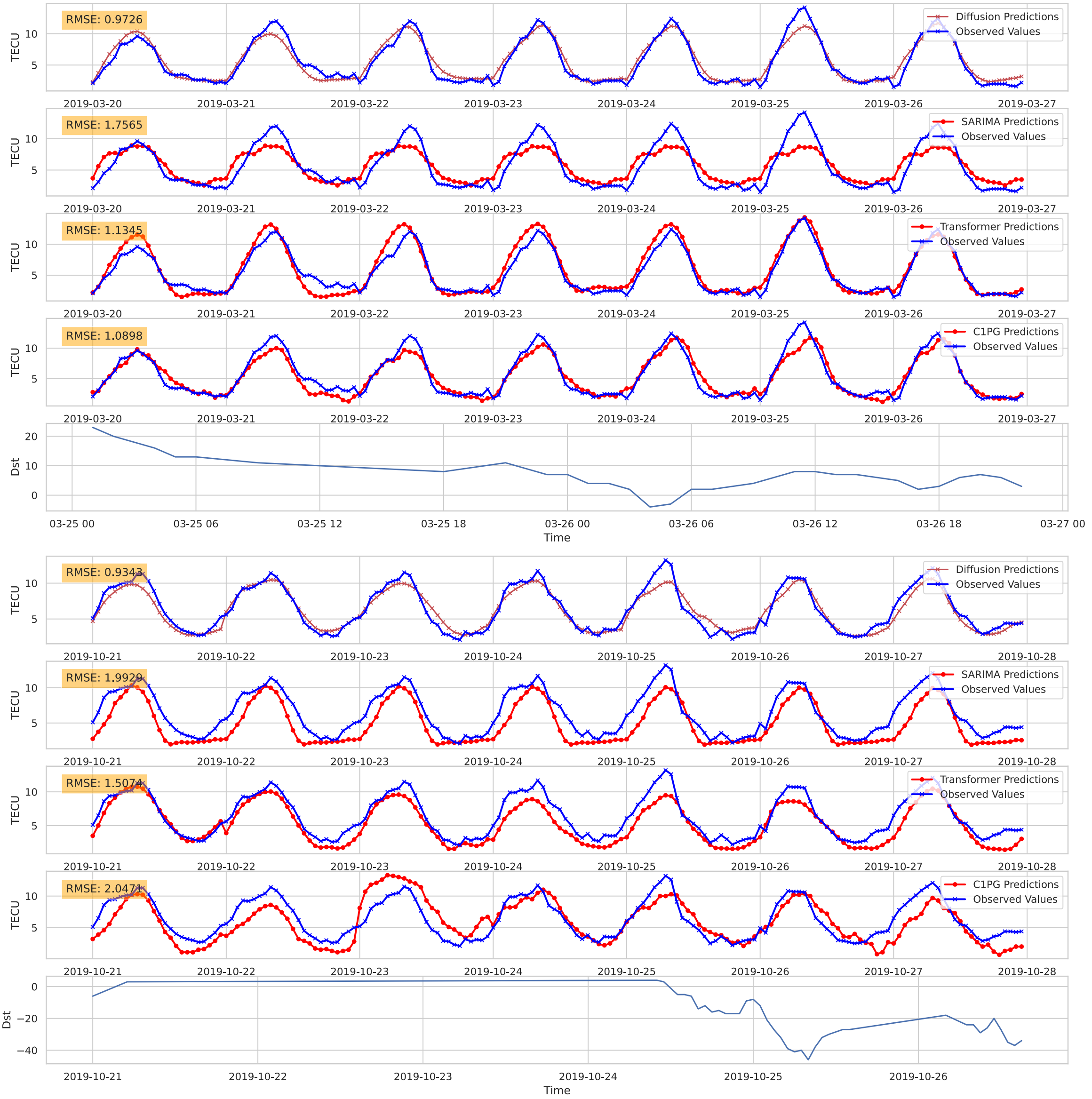

Training
VAE
X
Z
^
Y
LTC
DKL / Lmse
L(ζ, t)
X
Y
Diffusion
X
Y
Denoising
Network
Prediction Output
Prediction
VAE
X
Z
^
Y
X
Y
Diffusion
X
Y
Denoising
Network
Prediction Output

figure2.


Transformer Training Loss
MSE LOSS
Epoch
0
100
200
300
400
500
600
0
250
500
750
1000
1250
1500
1750
2000
Transformer_1_train_mse
Transformer_1_test_mse
Transformer_2_train_mse
Transformer_2_test_mse
Transformer_3_train_mse
Transformer_3_test_mse
Transformer_4_train_mse
Transformer_4_test_mse

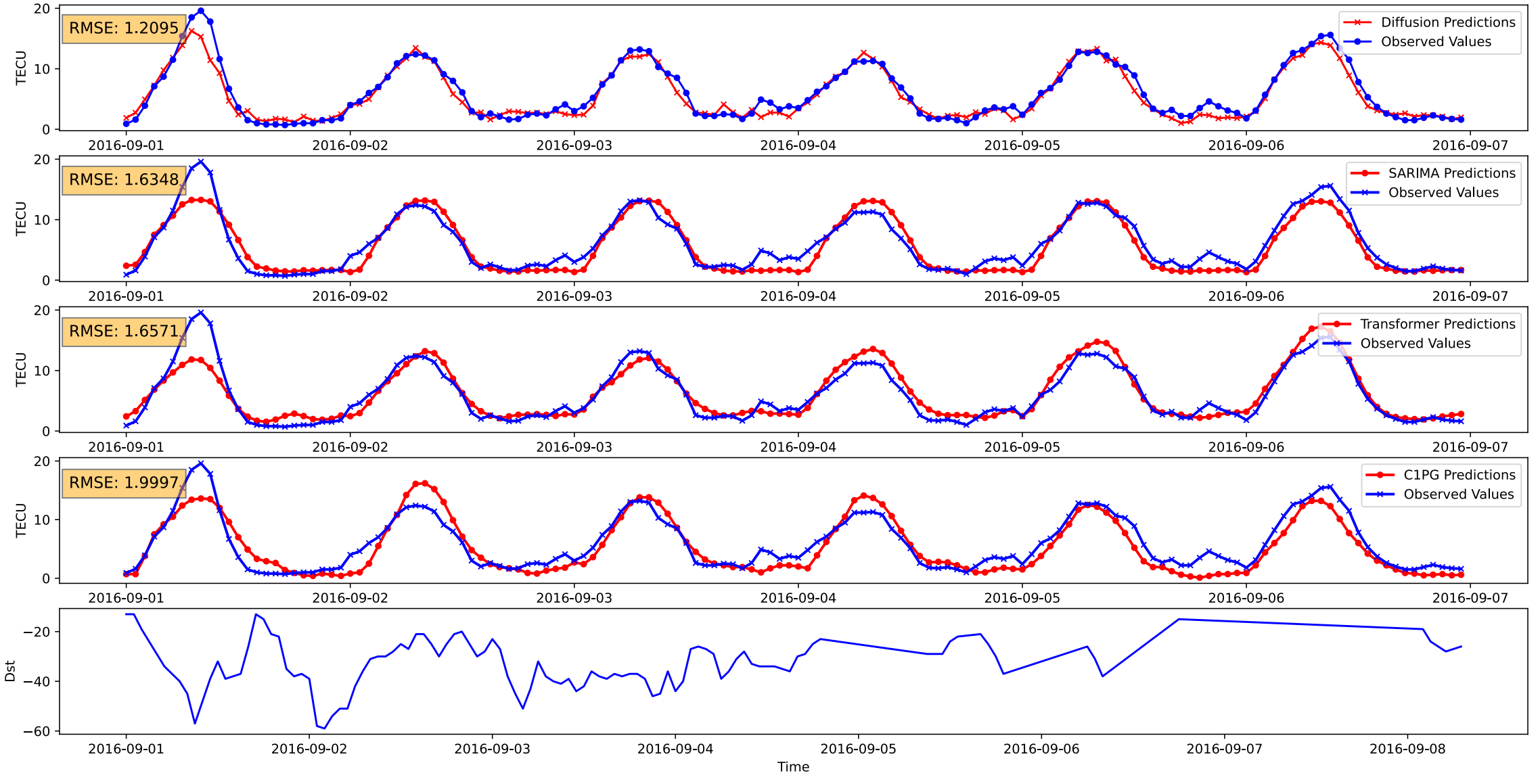

Diffusion Training Loss
Epoch
0
1
2
3
4
5
0
20
40
60
80
Diffusion_1_train_mse
Diffusion_2_train_mse
Diffusion_3_train_mse
Diffusion_4_train_mse
Diffusion_1_vali_mse
Diffusion_2_vali_mse
Diffusion_3 vali_mse
Diffusion_4_vali_mse

figure3.


RMSE: 0.9726
Diffusion Predictions
Observed Values
RMSE: 1.7565
SARIMA Predictions
Observed Values
RMSE: 1.1345
Transformer Predictions
Observed Values
RMSE: 1.0898
C1PG Predictions
Observed Values
TECU
Dst
Time
2019-03-20
2019-03-21
2019-03-22
2019-03-23
2019-03-24
2019-03-25
2019-03-26
2019-03-27
03-25 00
03-25 06
03-25 12
03-25 18
03-26 00
03-26 06
03-26 12
03-26 18
03-27 00

RMSE: 0.9343
Diffusion Predictions
Observed Values
RMSE: 1.9929
SARIMA Predictions
Observed Values
RMSE: 1.5074
Transformer Predictions
Observed Values
RMSE: 2.0471
C1PG Predictions
Observed Values
TECU
Dst
Time
2019-10-21
2019-10-22
2019-10-23
2019-10-24
2019-10-25
2019-10-26
2019-10-27
2019-10-28

figure4.


RMSE: 1.2095
Diffusion Predictions
Observed Values
RMSE: 1.6348
SARIMA Predictions
Observed Values
RMSE: 1.6571
Transformer Predictions
Observed Values
RMSE: 1.9997
C1PG Predictions
Observed Values
TECU
Dst
Time
2016-09-01
2016-09-02
2016-09-03
2016-09-04
2016-09-05
2016-09-06
2016-09-07
2016-09-08

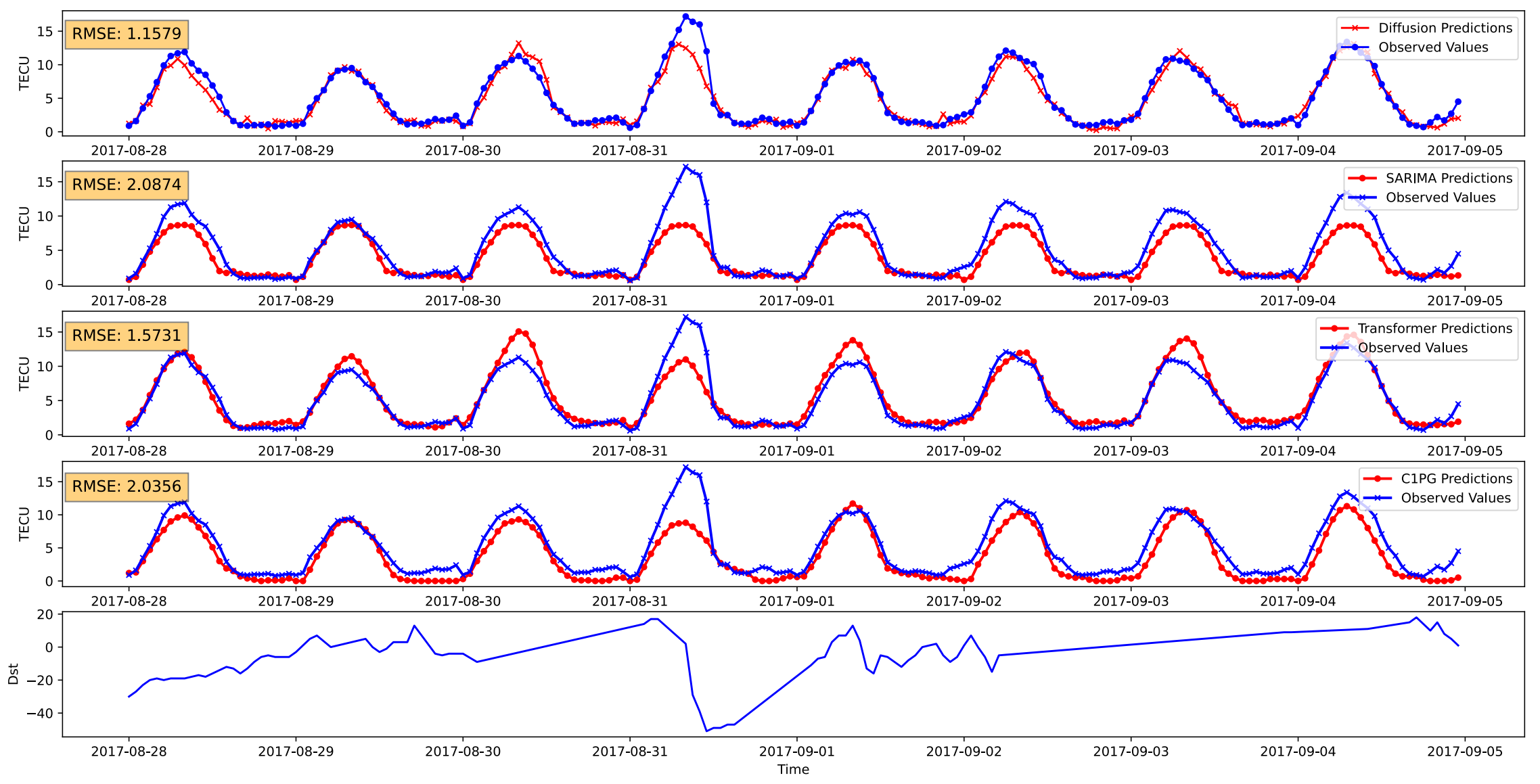

RMSE: 1.1579
Diffusion Predictions
Observed Values
RMSE: 2.0874
SARIMA Predictions
Observed Values
RMSE: 1.5731
Transformer Predictions
Observed Values
RMSE: 2.0356
C1PG Predictions
Observed Values
TECU
Dst
Time
2017-08-28
2017-08-29
2017-08-30
2017-08-31
2017-09-01
2017-09-02
2017-09-03
2017-09-04
2017-09-05

figure5.

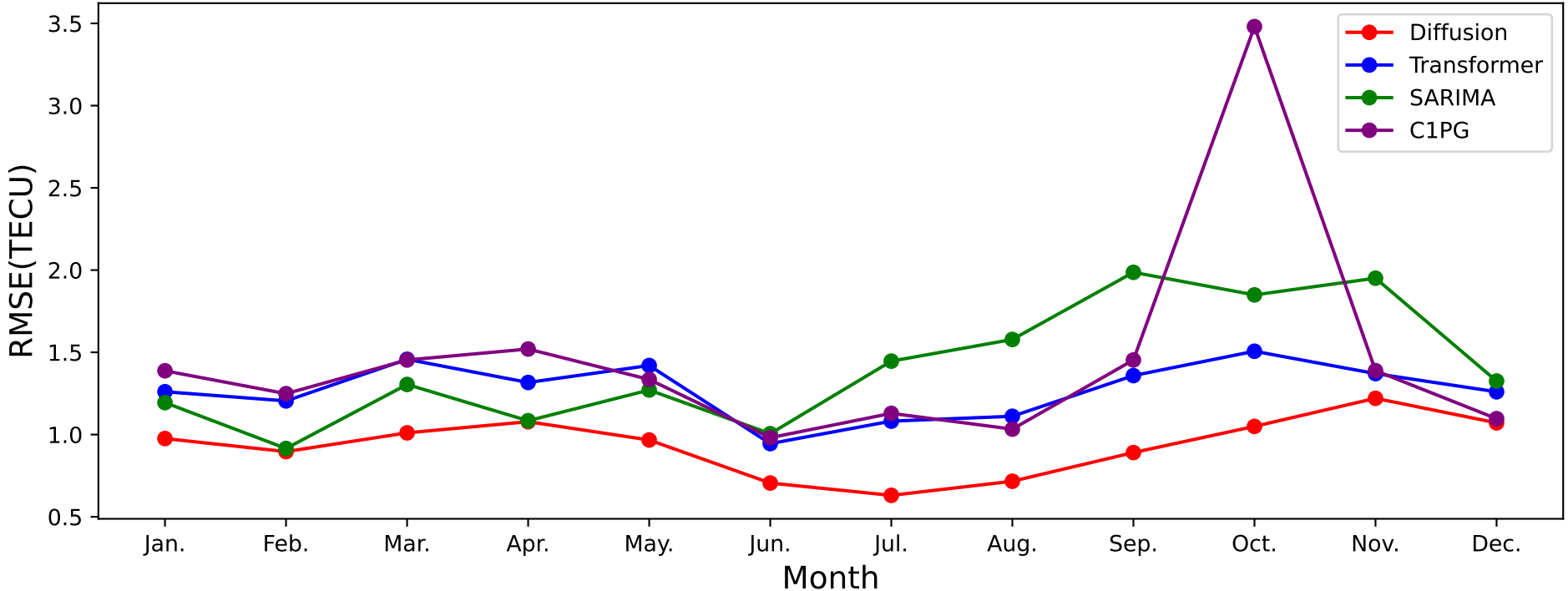

Diffusion
Transformer
SARIMA
C1PG
RMSE(TECU)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
Jan.
Feb.
Mar.
Apr.
May.
Jun.
Jul.
Aug.
Sep.
Oct.
Nov.
Dec.
Month

figure6.

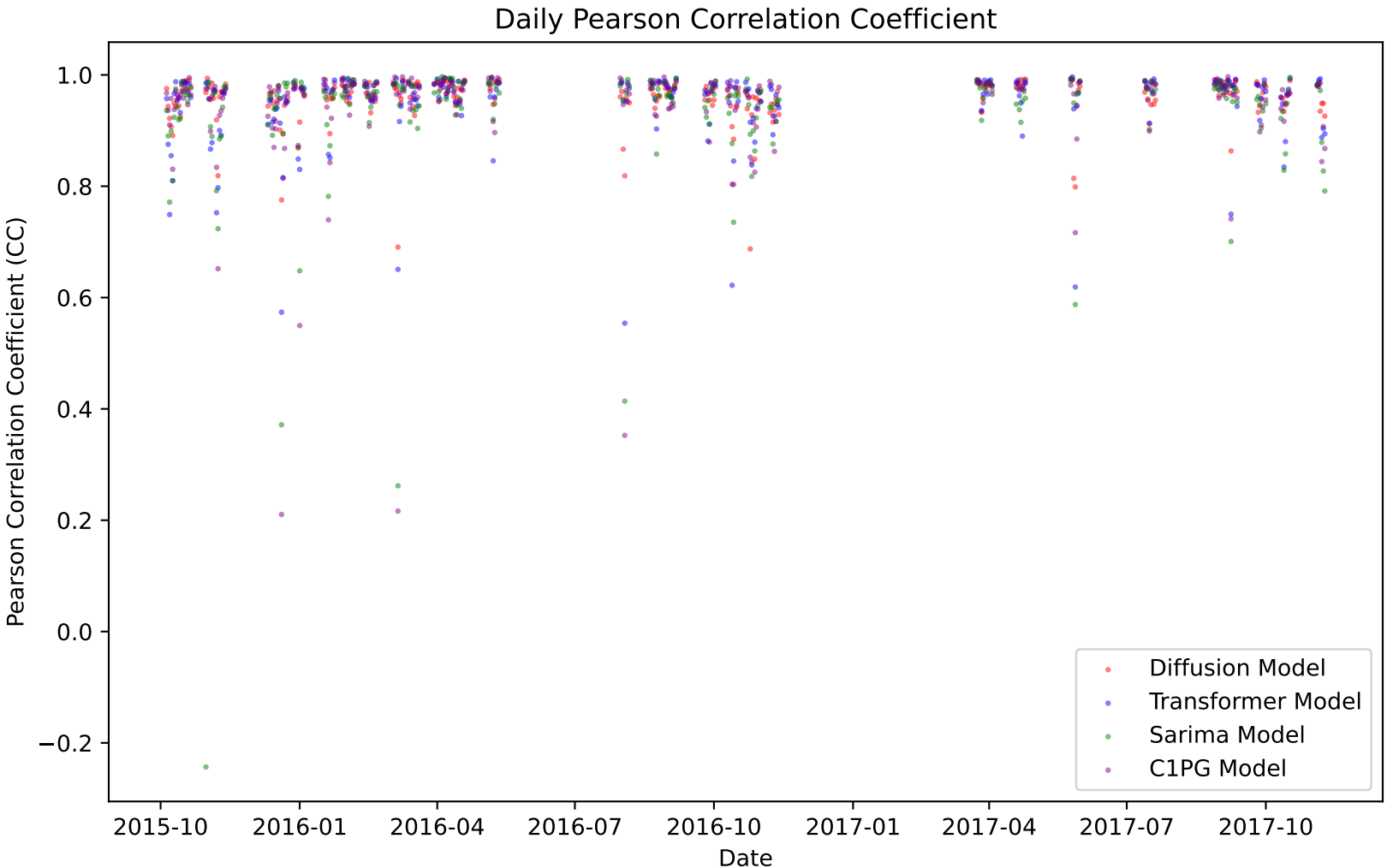

Daily Pearson Correlation Coefficient
Pearson Correlation Coefficient (CC)
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
2015-10
2016-01
2016-04
2016-07
2016-10
2017-01
2017-04
2017-07
2017-10
Date
Diffusion Model
Transformer Model
Sarima Model
C1PG Model

figure7.

CC vs. Average Dst for Each Month