



## Article

# Transformer-Based Ionospheric Prediction and Explainability Analysis for Enhanced GNSS Positioning

He-Sheng Wang , Dah-Jing Jwo \*  and Yu-Hsuan Lee

Department of Communications, Navigation, and Control Engineering, National Taiwan Ocean University, Keelung 202, Taiwan; hswang@email.ntou.edu.tw (H.-S.W.); 11167019@mail.ntou.edu.tw (Y.-H.L.)

\* Correspondence: djjwo@mail.ntou.edu.tw

**Abstract:** This study aims to investigate the impact of ionospheric models on Global Navigation Satellite System (GNSS) positioning and proposes an ionospheric prediction method based on a Transformer deep learning model. We construct a Transformer-based deep learning model that utilizes global ionospheric maps as input to achieve spatiotemporal prediction of Total Electron Content (TEC). To gain a deeper understanding of the model's prediction mechanism, we employ integrated gradients for explainability analysis. The results reveal the key ionospheric features that the model focuses on during prediction, providing guidance for further model optimization. This study demonstrates the efficacy of a Transformer-based model in predicting Vertical Total Electron Content (VTEC), achieving comparable accuracy to traditional methods while offering enhanced adaptability to spatial and temporal variations in ionospheric behavior. Furthermore, the application of advanced explainability techniques, particularly the Integrated Decision Gradient (IDG) method, provides unprecedented insights into the model's decision-making process, revealing complex feature interactions and spatial dependencies in VTEC prediction, thus bridging the gap between deep learning capabilities and explainable scientific modeling in geophysical applications. The model achieved positioning accuracies of  $-1.775$  m,  $-2.5720$  m, and  $2.6240$  m in the East, North, and Up directions respectively, with standard deviations of  $0.3399$  m,  $0.2971$  m, and  $1.3876$  m. For VTEC prediction, the model successfully captured the diurnal variations of the Equatorial Ionization Anomaly (EIA), with differences between predicted and CORG VTEC values typically ranging from  $-6$  to  $6$  TECU across the study region. The gradient score analysis revealed that solar activity indicators (F10.7 and sunspot number) showed the strongest correlations ( $0.7$ – $0.8$ ) with VTEC variations, while geomagnetic indices exhibited more localized impacts. The IDG method effectively identified feature importance variations across different spatial locations, demonstrating the model's ability to adapt to regional ionospheric characteristics.



Academic Editor: Angela Aragón-Ángel

Received: 12 December 2024

Revised: 27 December 2024

Accepted: 27 December 2024

Published: 28 December 2024

**Citation:** Wang, H.-S.; Jwo, D.-J.; Lee, Y.-H. Transformer-Based Ionospheric Prediction and Explainability Analysis for Enhanced GNSS Positioning.

*Remote Sens.* **2025**, *17*, 81. <https://doi.org/10.3390/rs17010081>

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** GNSS; transformer; ionospheric effect; explainability; integrated gradient; integrated decision gradient

## 1. Introduction

### 1.1. Introduction to GNSS and XAI

Global Navigation Satellite Systems (GNSSs) have become an integral part of modern life, providing precise positioning and navigation capabilities for a wide range of military, civilian, and commercial applications. Since the launch of the first Global Positioning System (GPS) satellite in 1978 and the completion of the GPS network in 1993, other countries and regions have developed their own systems, including Russia's GLONASS,

the European Union's Galileo, and China's BeiDou. These systems have enabled numerous applications, from vehicle navigation and mobile phone positioning to autonomous vehicles [1].

However, GNSS accuracy is significantly affected by ionospheric delay, which remains a critical challenge in satellite positioning [2,3]. Recent studies have further highlighted these challenges, particularly in the context of precise positioning applications. Lyu et al. [4] investigated the uncertainties in interpolating satellite-specific slant ionospheric delays and their impacts on PPP-RTK, while Tang et al. [5] examined how equatorial plasma bubbles and associated ionospheric gradients affect GNSS PPP-RTK performance. The ionosphere, a layer of the Earth's upper atmosphere, contains free electrons that can alter the speed and direction of radio signals passing through it [6]. This delay varies with factors such as solar activity, geomagnetic activity, and atmospheric disturbances, leading to reduced positioning accuracy.

In recent years, researchers have turned to deep learning methods to predict ionospheric Total Electron Content (TEC) and correct for navigation errors [7,8]. The rise of artificial intelligence, particularly with breakthroughs like OpenAI's GPT models, has further fueled interest in applying these technologies to various domains, including ionospheric prediction [9–14]. Traditional approaches to ionospheric prediction can be broadly categorized into three groups:

- Empirical Models: These include the International Reference Ionosphere (IRI), NeQuick, and various regional models. They rely on historical data and statistical analysis to predict ionospheric behavior.
- Physical Models: These are based on solving continuity, momentum, and energy equations for the ionospheric plasma.
- Data Assimilation Methods: These combine observations with physical or empirical models to provide more accurate predictions. The Global Assimilative Ionospheric Model (GAIM) is a prominent example.

Recent years have seen significant advances in deep learning approaches for ionospheric prediction. These include the following:

- Convolutional Neural Networks (CNNs) for spatial feature extraction from global ionospheric maps;
- Long Short-Term Memory (LSTM) networks for capturing temporal dependencies in TEC variations;
- Hybrid architectures, combining different neural network types for improved prediction accuracy;
- Attention-based models that can capture both local and global ionospheric patterns.

Each approach has its strengths and limitations. Deep learning methods have shown particular promise in capturing complex, non-linear relationships in ionospheric behavior and adapting to varying conditions.

Various artificial intelligence approaches have been applied to ionospheric prediction in recent years. Early efforts focused on traditional neural networks, with Tulunay et al. [7,8] demonstrating the feasibility of neural network techniques for ionospheric forecasting. The field has since evolved to incorporate more sophisticated architectures. Tulası Ram et al. [7] developed an improved two-dimensional artificial neural network model (ANNIM) for ionospheric modeling. Chen et al. [10] applied deep learning techniques to global ionospheric TEC prediction, while Luo et al. [11] introduced the SAM-ConvLSTM model for improved accuracy. Recurrent neural network architectures have also shown promise, with Sun et al. [12] implementing a bidirectional LSTM approach for vertical TEC forecasting. More recent work by Wang et al. [13] has explored automated deep

learning frameworks, combining multiple techniques, including residual networks and attention mechanisms. Ren et al. [14] specifically addressed TEC prediction during storm periods using a mixed CNN-BiLSTM approach. While these studies have demonstrated the potential of AI in ionospheric prediction, they typically lack explainability, making it difficult to understand how these models arrive at their predictions. This gap motivates our current work combining Transformer architecture with explainable AI techniques.

As AI systems become more prevalent in critical applications, there is a growing demand for transparency and trustworthiness in their decision-making processes. This has led to the emergence of eXplainable Artificial Intelligence (XAI), which aims to make AI systems more interpretable and their predictions more understandable to humans. XAI is an emerging field that seeks to make AI models more transparent. It focuses on developing methods to create models that are not only highly accurate but also interpretable to humans. XAI aims to elucidate how AI systems arrive at their decisions, thereby fostering trust among stakeholders and supporting the development of responsible AI [15,16]. By enhancing the comprehensibility of complex AI models, XAI plays a crucial role in bridging the gap between advanced machine learning techniques and human understanding.

This paper presents a novel approach that combines a Transformer-based model for regional ionospheric mapping with XAI techniques. Our work builds upon previous research in ionospheric prediction using neural networks, such as the time-based input features proposed by McKinnell and Poole [17] and the incorporation of geomagnetic indices and solar activity by Sivavaraprasad et al. [18]. In our study, we employ the Transformer deep learning architecture to construct an ionospheric prediction model. To analyze the interpretability of this model, we utilize the Integrated Decision Gradient (IDG) method, an advanced technique in the field of XAI.

The IDG method is an enhancement of the widely used Integrated Gradients (IG) approach, first introduced by Sundararajan et al. in 2017 [19]. IG attributes the prediction of a deep network to its input features by accumulating gradients along the straight-line path from a baseline input to the actual input. While IG has been successful in many applications, it can sometimes produce noisy attributions due to the gradient saturation problem, particularly in deep networks with non-linear activations [20]. To address these limitations, Walker et al. proposed the Integrated Decision Gradient method in 2023 [21]. IDG builds upon the foundational principles of IG but introduces key improvements to enhance the quality and reliability of feature attributions. The primary innovation of IDG lies in its approach to gradient calculation.

Instead of computing gradients with respect to the input features directly, IDG calculates gradients with respect to the decision boundary of the model. This subtle but crucial difference allows IDG to capture more meaningful attributions, especially in cases where the model's decision is primarily influenced by features close to the decision boundary.

In the context of our ionospheric prediction model, IDG allows us to achieve the following:

- Identify the most influential input features for each prediction, providing insights into the model's decision-making process.
- Understand how different ionospheric parameters contribute to the model's predictions across various spatial and temporal scales.
- Detect potential biases or unexpected behaviors in the model, which could lead to improvements in model architecture or training procedures.
- Enhance the interpretability of our Transformer-based model, making it more transparent and trustworthy for stakeholders in the GNSS community.

By employing IDG, we not only improve upon the interpretability offered by traditional methods like IG but also gain a more nuanced understanding of our model's behavior.

This approach aligns with the growing emphasis on responsible AI development in critical applications such as ionospheric modeling for GNSS systems [22,23].

The integration of advanced deep learning techniques like Transformers with state-of-the-art interpretability methods such as IDG represents a significant step forward in the field of ionospheric prediction. It allows us to leverage the power of complex neural networks while maintaining a high degree of model transparency and interpretability, crucial for the continued advancement and acceptance of AI-driven solutions in scientific and engineering domains. By combining advanced deep learning techniques with explainable AI methods, we aim to enhance the accuracy and reliability of GNSS positioning while providing insights into the model's decision-making process. This approach not only contributes to the field of ionospheric prediction but also addresses the broader need for transparent and trustworthy AI systems in critical applications.

### 1.2. Ionospheric Effects on Global Navigation Satellite Systems

The ionosphere, a critical layer of the Earth's atmosphere located between the troposphere and the exosphere, plays a significant role in GNSS positioning accuracy. Composed primarily of ionized gases, including oxygen, hydrogen, and other molecules, the ionosphere's electron density, thickness, and height vary with seasonal changes, diurnal cycles, and solar radiation. This variability introduces challenges in precise GNSS positioning, necessitating accurate monitoring and modeling of ionospheric behavior. The ionosphere is formed when high-energy radiation from the sun, particularly ultraviolet and X-ray, ionizes atmospheric molecules and atoms. This ionization process creates free electrons that affect satellite signals, causing variations in signal propagation speed and path delays [24]. The ionosphere's activity is closely correlated with solar and geomagnetic activities, making these features crucial for predicting the Vertical Total Electron Content (VTEC).

Total Electron Content is a key parameter in describing ionospheric behavior. It represents the total number of electrons present along the path from the satellite to the receiver, typically measured in TEC Units (TECU), where 1 TECU equals  $10^{16}$  electrons/m<sup>2</sup> [25]. TEC can be categorized into Slant TEC (STEC) and Vertical TEC (VTEC), with the relationship between them described by a mapping function. The single-layer ionosphere model simplifies the ionosphere as a thin shell concentrated at a specific height, typically between 350 and 450 km. This model, while neglecting the actual structure and thickness variations of the ionosphere, provides a useful approximation for many applications [2].

The influencing factors can be divided into three groups:

- **Solar Activity:** Solar activity follows an approximate 11-year cycle, characterized by variations in sunspot numbers and solar flux. During periods of high solar activity, phenomena such as solar flares and coronal mass ejections (CMEs) increase, leading to higher levels of ionization in the Earth's upper atmosphere [26]. The F10.7 index, which measures solar radio flux at a wavelength of 10.7 cm, serves as a proxy for solar activity and closely correlates with sunspot numbers.
- **Geomagnetic Activity:** The Earth's magnetic field significantly influences ionospheric plasma behavior. At the geomagnetic equator, the horizontal magnetic field creates an eastward electric current in the ionosphere. Additionally, the interplanetary magnetic field (IMF) interacts with the Earth's magnetosphere, particularly affecting high-latitude regions and contributing to auroral phenomena [27].
- **Anthropogenic Factors:** Human activities, such as nuclear tests and rocket launches, can also impact ionospheric structure. Nuclear tests have been observed to affect D-layer ionization, while rocket launches can create ionospheric holes due to the combustion of rocket fuel in the ionosphere [28,29].

The ionosphere significantly affects GNSS positioning by causing signal delays and bending. This effect, known as ionospheric delay, impacts the accuracy of GNSS signals. To compensate for these effects, researchers have developed various mathematical models to describe ionospheric behavior. Some well-known models include the Klobuchar model, NeQuick model, BeiDou Global Ionospheric delay correction Model (BDGIM), International Reference Ionosphere (IRI) model, and Global Ionosphere Map (GIM). Developed by the U.S. Department of Defense, the Klobuchar model is widely used in GPS systems. It employs eight parameters broadcast by GPS satellites to estimate ionospheric delay. While simple and efficient, its accuracy is limited, making it suitable primarily for single-frequency GPS receivers [2]. The NeQuick model is a global ionospheric model based on radiative transfer theory. It simulates electron density distribution and is commonly used in the Galileo navigation system. NeQuick provides more accurate modeling of electron density at different altitudes, thereby improving compensation for ionospheric delay [30]. The BDGIM [31–33] is designed specifically for the BeiDou Navigation Satellite System (BDS). Like other GNSS ionospheric models, it aims to mitigate ionospheric delay effects on positioning accuracy. BDGIM uses a series of coefficients broadcast in the BDS navigation message to calculate ionospheric delays. The model employs a spherical harmonic function to represent global TEC distribution and has demonstrated good performance in ionospheric delay correction, particularly in the Asia-Pacific region where BDS has stronger coverage. Compared to the Klobuchar model, BDGIM generally achieves better accuracy in ionospheric delay estimation, especially in low and middle latitude regions. The IRI model is an internationally recognized standard ionospheric model based on global measurement data and statistical analysis. Updated annually, it is applicable to a wide range of scenarios, including GPS positioning, communication, and radar systems [34].

GIMs, unlike the previous three models, are generated using data from GNSS networks. They rely on measurements from a global network of ground-based GNSS receivers. GIMs are produced by analyzing the ionospheric TEC obtained from GNSS signals. These TEC values are then interpolated over a global grid to create a map of the ionosphere's electron content [35]. Organizations like the Center for Orbit Determination in Europe (CODE) and the International GNSS Service (IGS) routinely produce GIMs using their own algorithms and interpolation techniques. GIMs provide near-real-time or post-processed global maps of the ionosphere's TEC. They are crucial for applications that require current ionospheric conditions, such as GNSS positioning and navigation, where ionospheric delays can significantly impact accuracy. Recent studies have evaluated GIM performance under varying conditions. Wielgosz et al. [36] conducted comprehensive validations of GIM accuracy across different solar activity levels, comparing performance during solar minimum (2014) and maximum (2018) periods. Tang et al. [37] specifically investigated GIM-TEC errors during equatorial plasma bubble events in high solar activity conditions, revealing important limitations in GIM accuracy during ionospheric disturbances. GIMs can provide high temporal resolution (e.g., hourly updates), reflecting the ionosphere's current state, including short-term variations and disturbances like geomagnetic storms. They are ideal for real-time applications requiring current ionospheric information, such as precise GNSS positioning, satellite communication adjustments, and space weather monitoring.

Understanding and modeling ionospheric effects are crucial for improving GNSS positioning accuracy. The choice of model depends on the specific application scenario and required precision. The Taiwan region presents particular challenges for VTEC prediction due to its location near the northern crest of the Equatorial Ionization Anomaly (EIA) [38]. The EIA is characterized by two bands of enhanced ionization approximately 15–20 degrees north and south of the magnetic equator, with a trough at the equator [39]. Over Taiwan, the northern EIA crest creates dynamic ionospheric conditions with significant VTEC variations,

even during magnetically quiet periods. This feature typically begins developing during morning hours as solar heating increases, reaches maximum development in the afternoon (around 14:00–16:00 local time), and breaks down during evening hours. The process creates substantial spatial and temporal VTEC gradients that pose significant challenges for prediction. During magnetically disturbed conditions, the EIA becomes even more active and irregular, with enhanced VTEC variations and more unpredictable behavior. These characteristics make our study region particularly suitable for testing ionospheric prediction models, as successful prediction here requires handling both regular daily EIA variations and irregular disturbance effects [40,41].

Our experimental framework consists of several key components. We utilized CORG data spanning from January 2022 to December 2023 for model training and validation, focusing on a regional map around Taiwan (115°E–130°E, 17.5°N–30°N). The study involves three main phases: (1) development and training of the Transformer model using eight input features, including solar activity indicators, geomagnetic indices, and time functions; (2) application of the IDG method to analyze feature importance and model behavior; and (3) validation of the model's performance through VTEC prediction accuracy assessment and GNSS positioning experiments. We evaluate our results using multiple metrics including VTEC prediction errors, positioning accuracy in East–North–Up coordinates, and feature contribution analysis through gradient scores. The positioning experiments were conducted using single-frequency observations, with results compared against dual-frequency solutions and CORG-based corrections to assess practical improvements in GNSS positioning accuracy.

## 2. Transformer Model

The Transformer model, introduced by Vaswani et al. [42] in their 2017 paper “Attention Is All You Need”, revolutionized the field of natural language processing and has since been adapted for various other domains. Unlike previous sequence-to-sequence models that relied on recurrent or convolutional neural networks, the Transformer is based entirely on attention mechanisms, allowing for more parallel processing and better handling of long-range dependencies.

The Transformer architecture comprises an encoder and a decoder, each built from layers that include multi-head self-attention mechanisms and position-wise feed-forward networks.

### A. Encoder

The encoder transforms the input sequence into a continuous representation. It begins with input embeddings, where each element of the input sequence is mapped into a high-dimensional space. Positional encodings are added to these embeddings to retain information about the sequence order, as the model does not inherently capture positional information. The core of the encoder consists of multi-head self-attention layers, allowing the model to focus on different positions within the input sequence to capture various types of dependencies. Following the attention layers are feed-forward networks applied to each position separately, introducing non-linearity and enhancing the model's capacity.

### B. Decoder

The decoder generates the output sequence by attending to both the encoder's output and the previously generated outputs. It includes masked multi-head self-attention layers to prevent the model from attending to future positions, preserving the auto-regressive property necessary for sequence generation. The decoder also incorporates encoder–decoder attention layers, enabling it to focus on relevant parts of the input sequence. The Dst index measures the intensity of the global geomagnetic storm level. The Kp index quantifies

disturbances in the Earth's magnetic field. The Ap index is a daily index derived from the Kp values.

### C. Self-Attention Mechanism

At the heart of the Transformer is the self-attention mechanism, which computes a weighted sum of input representations, where weights are determined by the similarity between elements. For each position in the sequence, the model calculates query, key, and value vectors. The attention scores are computed using scaled dot-product attention:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (1)$$

where  $Q$ ,  $K$ , and  $V$  represent the query, key, and value matrices, respectively,  $d_k$  is the dimension of the key vectors, and softmax is defined as follows:

$$\text{softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

where  $x_i$  represents the  $i$ -th element of the input vector  $x$ , and the sum in the denominator is taken over all elements  $j$ .

### D. Multi-Head Attention

To capture different types of dependencies, the Transformer uses multi-head attention, which allows the model to jointly attend to information from different representation subspaces. Each attention head performs the self-attention mechanism independently, and the outputs are concatenated and linearly transformed:

$$\text{Multihead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (2)$$

where  $\text{Concat}(\text{head}_1, \dots, \text{head}_h)$  denotes the concatenation operation that combines multiple head outputs into a single matrix by joining them along their last dimension. For example, if each  $\text{head}_i$  is a matrix of shape  $(N \times D)$ , then  $\text{Concat}(\text{head}_1, \dots, \text{head}_h)$  produces a matrix of shape  $(N \times hD)$ , where  $h$  is the number of attention heads. Each head is calculated as follows:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V), \quad (3)$$

where  $W^O$  is the weighting matrix for concatenated output, and  $W^Q$ ,  $W^K$ , and  $W^V$  are the weighting matrices for query, key, and value, respectively.

### E. Positional Encoding

Since the Transformer does not inherently account for the order of the sequence, positional encodings are added to the input embeddings to introduce sequence order information. Commonly, sinusoidal functions are used for positional encodings:

$$\text{PE}_{(pos, 2i)} = \sin\left(\frac{pos}{1000^{2i/d_{\text{model}}}}\right), \text{PE}_{(pos, 2i+1)} = \cos\left(\frac{pos}{1000^{2i/d_{\text{model}}}}\right) \quad (4)$$

where  $pos$  is the position and  $i$  is the dimension.

The ionosphere exhibits complex spatial-temporal behaviors influenced by solar activity, geomagnetic conditions, and atmospheric disturbances. Traditional models often fail to capture these complexities due to the following:

- **Spatial-Temporal Complexity:** The ionosphere's behavior varies significantly across different locations and times.

- Long-Range Dependencies: Past ionospheric conditions can influence future states over extended periods and distances.
- Data Representation: GIM data provides TEC values over a global grid, requiring the effective handling of spatial data in conjunction with temporal sequences.

To apply the Transformer model to ionospheric prediction using GIM data, several adaptations are necessary. The GIM provides TEC values on a global grid defined by latitude and longitude. In our study, we focus on a regional map around Taiwan to capture local ionospheric variations pertinent to GNSS applications in this area. Each TEC measurement corresponds to a specific longitude and latitude grid point. The selected features represent the primary physical factors known to influence ionospheric TEC variations:

- Solar Activity Indicators (Sunspot number and F10.7 index): These directly measure solar radiation levels, which are the primary source of ionospheric ionization. The F10.7 index strongly correlates with solar EUV radiation, which drives ionospheric electron production. Using both indicators provides redundancy and robustness in capturing solar activity effects.
- Geomagnetic Indices (Dst and Kp): Dst index captures global-scale magnetic disturbances and storm effects. The Kp index reflects broader geomagnetic activity, which affects electron density distribution. Together, they account for both sudden and gradual geomagnetic influences on the ionosphere.
- Time Functions (DNS, DNC, HRS, HRC): DNS and DNC capture seasonal variations in solar zenith angle and day length. HRS and HRC represent diurnal variations due to the Earth's rotation. These sinusoidal functions effectively model the periodic nature of solar illumination patterns.

While additional parameters like neutral winds or electric fields also affect TEC, their effects are indirectly reflected in the geomagnetic indices, and direct measurements are not typically available in real time for operational GNSS applications. Our experimental results demonstrate that these selected features capture the main drivers of TEC variations sufficiently well for practical prediction purposes.

Here, the time functions DNS and DNC are defined as follows:

$$\text{DNS} = \sin\left(\frac{2\pi \times \text{Day Number}}{365}\right), \text{DNC} = \cos\left(\frac{2\pi \times \text{Day Number}}{365}\right), \quad (5)$$

representing seasonal variations. HRS and HRC are calculated as follows:

$$\text{HRS} = \sin\left(\frac{2\pi \times \text{Hour}}{365}\right), \text{HRC} = \cos\left(\frac{2\pi \times \text{Hour}}{365}\right), \quad (6)$$

representing diurnal variations. These time functions capture the periodic nature of the ionosphere influenced by the Earth's rotation and orbit. The sunspot number is updated every 24 h, reflecting solar activity levels. The F10.7 Index is the solar radio flux at 10.7 cm wavelength, an indicator of solar EUV emissions affecting the ionosphere. The Dst index measures the intensity of the global geomagnetic storm level. The Kp Index quantifies disturbances in the Earth's magnetic field. The Ap index is a daily index derived from the Kp values. By incorporating these features, the model gains a comprehensive view of factors influencing ionospheric conditions.

The Transformer model implementation utilized open-source frameworks, as the focus of our work was not on developing novel neural architectures but rather on understanding and improving model behavior through explainability analysis. The IDG analysis framework, however, was custom-developed for this specific application to ensure proper handling of ionospheric data characteristics. The Transformer architecture has been



adapted for VTEC prediction as follows: For input encoding, each VTEC measurement point is encoded with its corresponding features (solar activity indicators, geomagnetic indices, time functions). Positional encoding represents the spatial relationships between grid points in our Taiwan region (115°E–130°E, 17.5°N–30°N). The Self-Attention Mechanism captures spatial dependencies between different grid points in our region, allowing the model to weigh the importance of neighboring points differently based on ionospheric conditions. This is particularly important for tracking EIA development and movement across our grid. For multi-head attention, different attention heads capture various types of relationships. This is crucial for simultaneously tracking the different ionospheric phenomena affecting our region. For the encoder–decoder structure, the encoder processes input features and their relationships, and the decoder generates VTEC predictions while considering all these relationships. This architecture is particularly suited for VTEC prediction because it can handle multiple input features simultaneously and capture both spatial and temporal dependencies.

### 3. The Need for Explainable AI

Black-box models, such as deep neural networks, have achieved remarkable success across various domains due to their ability to model complex, non-linear relationships. However, their opacity presents significant challenges. Firstly, the internal decision-making processes are often inaccessible or incomprehensible, making it difficult to understand how inputs are transformed into outputs. Moreover, without interpretability, models may inadvertently perpetuate or amplify biases present in the training data, leading to unfair or erroneous outcomes. In critical applications like GNSS, where decisions can have significant consequences, stakeholders may be hesitant to rely on models they do not understand.

Explainable AI seeks to address these limitations by making AI systems more transparent and interpretable. The motivations for XAI include the following:

- **System Verification:** Ensuring that AI models make decisions based on relevant and appropriate factors is essential for validating model behavior and performance.
- **System Improvement:** Understanding the model's reasoning can highlight weaknesses or errors, guiding targeted improvements and refinements.
- **Learning from the System:** Interpretable models can provide novel insights or patterns not immediately apparent to human experts, advancing domain knowledge.
- **Legal and Regulatory Compliance:** Emerging regulations, such as the European Union's General Data Protection Regulation (GDPR), mandate explainability in AI systems to protect user rights and promote ethical practices.

Our Transformer-based model predicts ionospheric TEC using GIM data, focusing on a regional map around Taiwan. The model incorporates several input features as described in the previous section. The model architecture includes an encoder and a decoder with multi-head attention layers, residual connections, layer normalization, and feed-forward networks. Key parameters include eight input features, one output feature, 512 neurons in the embedding layers, a dropout rate of 0.1, and the use of the Adam optimizer with a Mean Squared Error (MSE) loss function. The architecture of the model will be given in the experimental section.

Despite its predictive success (the results are given in the next section), the model's complexity poses challenges for interpretability. First of all, the use of high-dimensional embeddings and multiple layers makes it difficult to trace the influence of individual input features. Next, the model captures complex non-linear relationships, complicating efforts to understand how specific inputs affect outputs. Finally, while attention weights provide some insight, they do not always directly correlate with feature importance or causal influence.

To address these challenges, we turn to gradient-based explanation methods, which have shown promise in interpreting deep learning models. These methods analyze the gradients of the output with respect to the input features, identifying how changes in inputs affect the output.

Integrated Gradients, introduced by Sundararajan et al. [19], compute the average gradients along the path from a baseline input to the actual input. The integrated gradient along the  $i$ th dimension for an input  $x$  and baseline  $x'$  (usually a zero vector) is defined as follows:

$$\text{IG}_i(x) = (x_i - x'_i) \times \int_{\alpha=0}^1 \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} d\alpha, \quad (7)$$

where  $x_i$  denotes the  $i$ th dimension,  $F$  is the model's output function (the model to be explained),  $\alpha$  is the interpolation constant, and  $i$  indexes the input features. In practice, the integral is approximated using a summation:

$$\text{IG}_i^{\text{approx}}(x) = (x_i - x'_i) \times \sum_{k=1}^m \frac{\partial F(x' + \frac{k}{m}(x - x'))}{\partial x_i} \times \frac{1}{m} \quad (8)$$

While IG provides valuable insights, it can suffer from issues such as gradient saturation and sensitivity to the choice of baseline, potentially leading to noisy or misleading attributions. To overcome these limitations, we employ the Integrated Decision Gradient (IDG) method, which enhances IG by focusing on gradients near the decision boundary. IDG calculates gradients with respect to the model's decision function, capturing more meaningful attributions, especially in deep models with non-linear activations.

IDG introduces an additional weighting factor based on the gradient with respect to the interpolation constant:

$$\text{ImportanceFactor}(\alpha) = \frac{\partial F(x' + \alpha(x - x'))}{\partial \alpha} \quad (9)$$

$$\text{IDG}_i(x) = \int_{\alpha=0}^1 \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} \frac{\partial F(x' + \alpha(x - x'))}{\partial \alpha} d\alpha, \quad (10)$$

The approximation for IDG is as follows:

$$\text{IDG}_i^{\text{approx}}(x) = (x_i - x'_i) \times \sum_{k=1}^m \frac{\partial F(x' + \frac{k}{m}(x - x'))}{\partial x_i} \times \frac{\partial F(x' + \frac{k}{m}(x - x'))}{\partial \alpha} \times \frac{1}{m} \quad (11)$$

By integrating gradient-based explanation techniques like IDG into our Transformer model, we can identify which input features contribute most to the model's predictions, aiding in understanding and validating the model's behavior. It is also possible to uncover potential biases or unintended influences in the model, facilitating corrective measures. The method may gain insights into the model's weaknesses, guiding targeted refinements and enhancements.

IDG is generally considered a model-agnostic method. It is built upon the principles of gradient-based attribution methods, which are typically applicable to a wide range of differentiable models. This includes most neural network architectures. IDG is an extension of the IG method, which is itself model-agnostic. IG can be applied to any differentiable model, and IDG inherits this property. Like other gradient-based methods, IDG primarily relies on analyzing the relationship between inputs and outputs of the model, rather than requiring access to specific internal structures or parameters. However, it is important to note that IDG, like other gradient-based methods, requires that the model be differentiable. This means it is not applicable to non-differentiable models like decision

trees or random forests. While theoretically applicable to many models, the computational efficiency of IDG might vary depending on the model architecture and complexity. In practice, some adaptations might be necessary when applying IDG to different types of models or data, especially when moving beyond the image processing domain, where it has shown clear benefits.

In our VTEC prediction application, IDG reveals how the model processes different input features to make its predictions. The method helps us understand why the model predicts certain VTEC values by showing the relative importance of solar activity indicators, geomagnetic indices, and time functions at each grid point. For example, at the northern grid points (25°N), the interpolation weights show different patterns compared to southern points (22.5°N), suggesting the model adapts its prediction strategy based on geographic location. The gradient scores reveal how the model responds to changes in solar and geomagnetic conditions, while cumulative gradients demonstrate the long-term impact of each feature. These insights not only validate that our model learns physically meaningful relationships but also identify which features drive predictions under different conditions, providing guidance for potential model improvements.

#### 4. Experimental Results

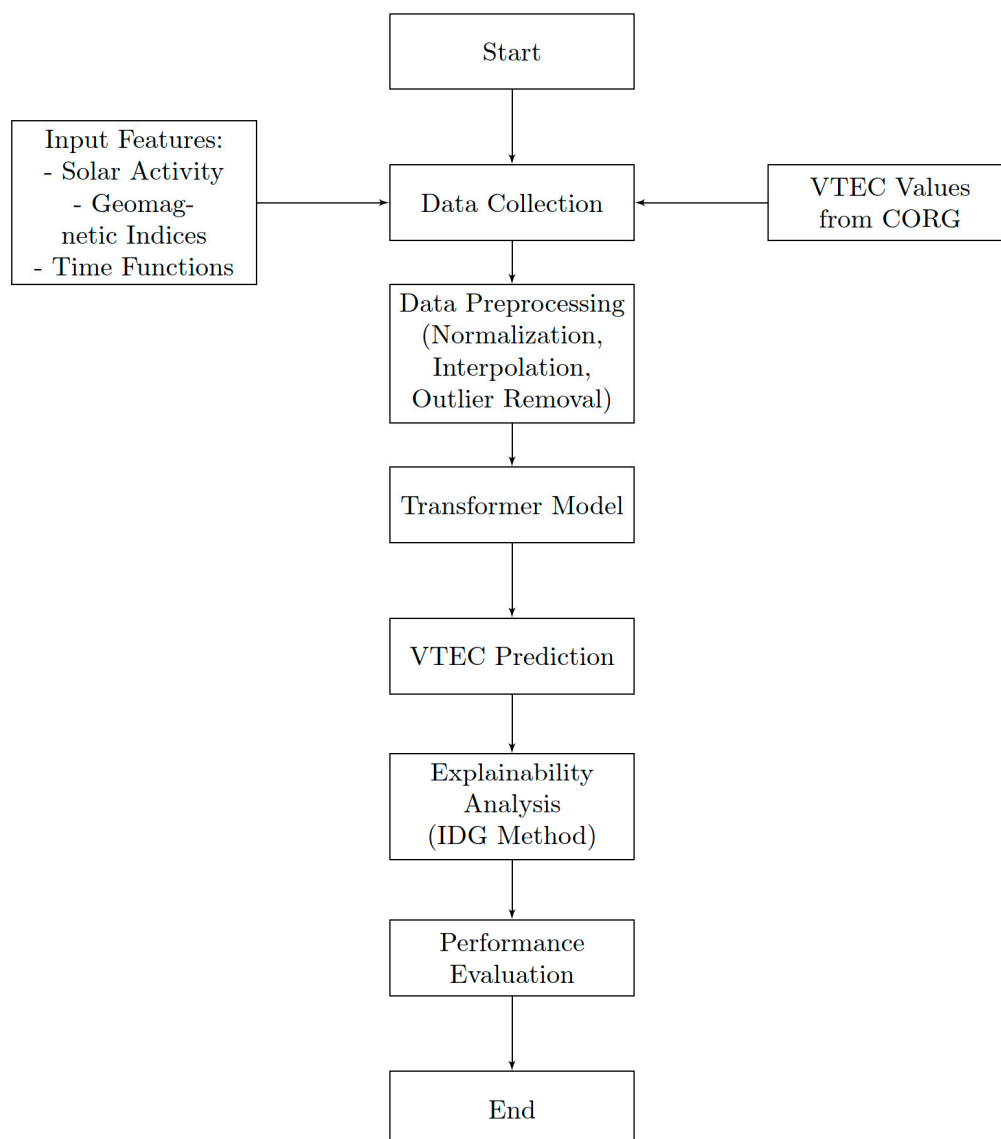
The flow of the experiment is shown in Figure 1, while the flow of the transformer model is shown in Figure 2. In order to analyze the explainability of deep learning models in the context of ionospheric predictions, we first employed a Transformer model to predict VTEC. This approach serves a dual purpose: firstly, to demonstrate the efficacy of the Transformer architecture in capturing the complex spatio-temporal dynamics of the ionosphere, and secondly, to provide a robust foundation for subsequent explainability analyses. By first validating the accuracy of our VTEC predictions, we ensure that the model we seek to interpret is indeed capturing meaningful patterns in the data. This step is crucial, as interpreting a poorly performing model would yield little insight into the underlying physical processes. Once the model's predictive capabilities are established, we can then confidently apply interpretability techniques to unravel the decision-making processes within the model. This approach allows us to not only improve VTEC predictions but also to gain deeper insights into the factors influencing ionospheric behavior, potentially advancing our understanding of space weather phenomena.

Input features of the model are as described in Section 3. The study utilized the rapid ionosphere map product (CODE Orbit Rapid GIM, abbreviated as CORG) retrieved from the CODE website. The prediction area covered longitudes from 115°E to 130°E and latitudes from 17.5°N to 30°N, encompassing 24 grid points (See Figure 3).

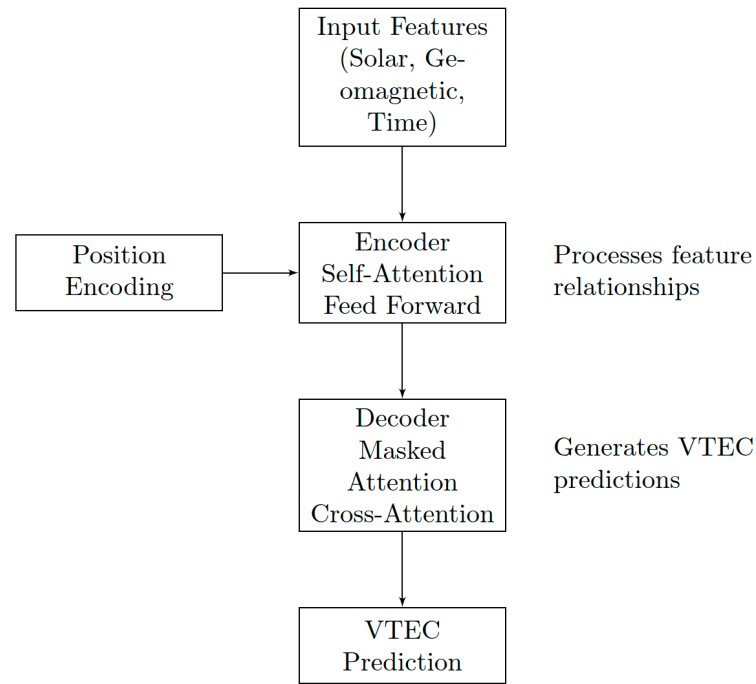
The model was trained using data from the following sources and periods. Some key parameters are summarized as follows:

- (a) Training Data:
  - Time Period: January 2022 to December 2023
  - Temporal Resolution: 24 data points per day (hourly samples)
  - Spatial Coverage: Regional map around Taiwan (115°E–130°E, 17.5°N–30°N)
  - Number of Grid Points: 24 points total (4 × 6 grid)
- (b) Input Features (eight features total):
  - Solar Activity Indicators
  - Sunspot Number
  - F10.7 index
  - Geomagnetic Indices
  - Dst index

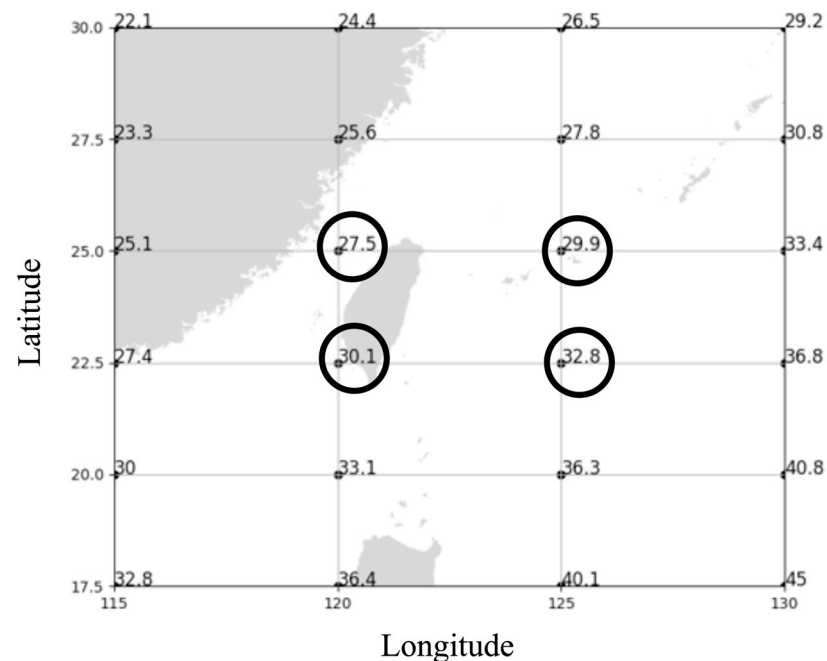
- Kp index
- (c) Time Functions
- DNS (Day Number Sine)
  - DNC (Day Number Cosine)
  - HRS (Hour Sine)
  - HRC (Hour Cosine)
- (d) Target Variable: VTEC values from CODE's rapid ionosphere map product (CORG)
- (e) Training Parameters:
- Training/Validation Split: 80%/20%
  - Batch Size: 64
  - Learning Rate: 0.001
  - Number of Epochs: 20
  - Optimizer: Adam
  - Loss Function: Mean Squared Error (MSE)



**Figure 1.** Flowchart of the proposed VTEC prediction algorithm using the Transformer model with explainability analysis.



**Figure 2.** Transformer model architecture for VTEC prediction, showing the flow from input features through encoder–decoder structure to final VTEC predictions. The model incorporates positional encoding and multiple attention mechanisms to capture spatial-temporal relationships in ionospheric behavior.



**Figure 3.** Example of 24 grid points covering the area around Taiwan. The values next to the grid points show VTEC values retrieved from the CODE website at 2022 11/25 00:00. The four grids enclosed with small circles will be used to compute the correlations later on.

The model was evaluated on test data from January 2024, which was completely separate from the training/validation period. The model architecture includes an encoder and a decoder with multi-head attention layers, residual connections, layer normalization, and feed-forward networks. Key parameters include eight input features (four GIM parameters and four time functions). For GIM parameters, the  $A_p$  index is not used as an input feature,

since it is a daily average of the geomagnetic activity, providing a linear measure of the overall level of geomagnetic disturbances. It is derived from the Kp index by converting each three-hour Kp value into a linear scale. In this case, the Kp index should have enough information for daily geomagnetic activity), one output feature (VTEC), 512 neurons in the embedding layers, a dropout rate of 0.1, and the use of the Adam optimizer with a Mean Squared Error (MSE) loss function. Model parameters are summarized in Table 1.

**Table 1.** Key parameters of the Transformer model.

Transformer Parameters and Information			
Input Feature	9	Batch Size	64
Output Feature	1	Activation	ReLU
Optimizer	Adam	Loss Function	MSE
Epoch	20	Time Step	24
Neuron Number	512	Dropout	0.1

The choice of 512 neurons in the embedding layers was determined through empirical testing and consideration of several factors:

- **Model Capacity:** A total of 512 dimensions provide sufficient representational capacity to capture the complex relationships in our input features while avoiding excessive parameters that could lead to overfitting. We experimented with different dimensions (128, 256, 512, and 1024) and found that 512 offered the best balance between model performance and computational efficiency.
- **Common Practice:** This dimension is a standard choice in Transformer architectures, as established in the original ‘Attention is All You Need’ paper by Vaswani et al. [31]. It provides adequate capacity for most sequence modeling tasks while maintaining reasonable computational requirements.
- **Memory Constraints:** While larger dimensions (e.g., 1024) might theoretically capture more subtle patterns, they significantly increase memory usage and computation time without providing substantial improvements in our prediction accuracy.

The model predicted VTEC values for the 24 grid points in the region, with bilinear interpolation used to fill in the remaining areas. We present the results for 25 November 2022, at different times to demonstrate the model’s performance. The formula for bilinear interpolation is as follows (See Figure 4):

$$\overline{VTEC}(\lambda, \psi, t) = (1 - p)(1 - q)\overline{VTEC}(\lambda_1, \psi_1, t) + (1 - p)q\overline{VTEC}(\lambda_1, \psi_2, t) + p(1 - q)\overline{VTEC}(\lambda_2, \psi_1, t) + pq\overline{VTEC}(\lambda_2, \psi_2, t) \quad (12)$$

where  $t$  is the time of observation;  $\lambda$  and  $\psi$  are longitude and latitude, respectively; and  $p$  and  $q$  are defined as follows:

$$p = \frac{\lambda - \lambda_i}{\lambda_j - \lambda_i}, q = \frac{\psi - \psi_i}{\psi_j - \psi_i}$$

Figure 5a shows the interpolated regional ionospheric map using the data taken from the CORG file. Figure 5b shows the model predictions. The predicted VTEC values closely match the CORG file data, with slight variations in some areas. The interpolated maps show similar patterns of VTEC distribution across the region. The result reflects the ionospheric effect at midnight local time, which usually has a lower VTEC value in the Taiwan region. The 24 h difference averages between the CORG and predicted VTEC values (CORG–Predicted) in TECU at 24 grid points are shown in Table 2.

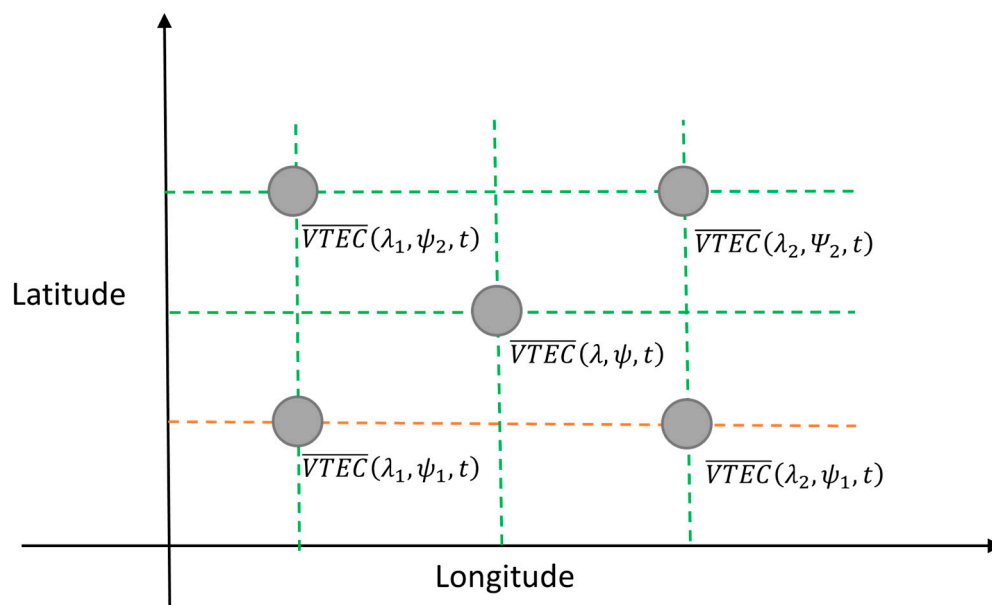


Figure 4. Bilinear interpolation using four grid points for interpolation of VTECs inside the rectangle.

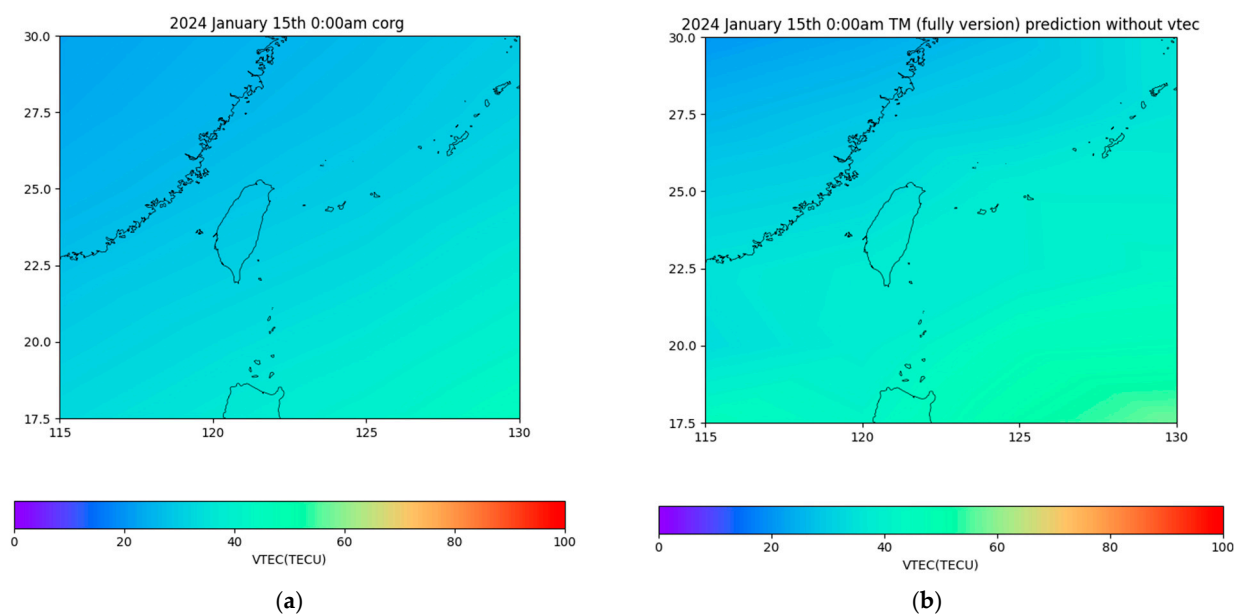


Figure 5. Local ionospheric map with lower VTEC value. (a) CORG data. (b) Prediction results.

Table 2. Differences between CORG and predicted VTEC values (CORG–Predicted) in TECU at 24 grid points.

Lat/Long	115° E	120° E	125° E	130° E
30.0° N	6.04	2.33	3.70	0.12
27.5° N	−1.37	2.35	1.96	−1.39
25.0° N	−2.07	−5.01	−3.00	−1.09
22.5° N	−6.19	−1.52	−1.28	−0.27
20.0° N	−2.74	−3.48	−1.95	−2.42
17.5° N	−6.45	0.61	−4.99	−1.45

To demonstrate our model's ability to capture the dynamic evolution of ionospheric conditions, particularly the development of the EIA over Taiwan, we present a series of comparisons between CORG data and model predictions throughout key periods on 15 January 2024. Figure 6 shows the progression from the early morning through the afternoon hours. Throughout this sequence, our model successfully reproduces both the magnitude and spatial distribution of VTEC variations, demonstrating its capability to track the complex temporal evolution of ionospheric conditions, particularly during the crucial period of EIA development and intensification. This sequence provides strong evidence that our model can effectively capture not just static conditions but also the dynamic processes driving ionospheric variability over our region of interest.

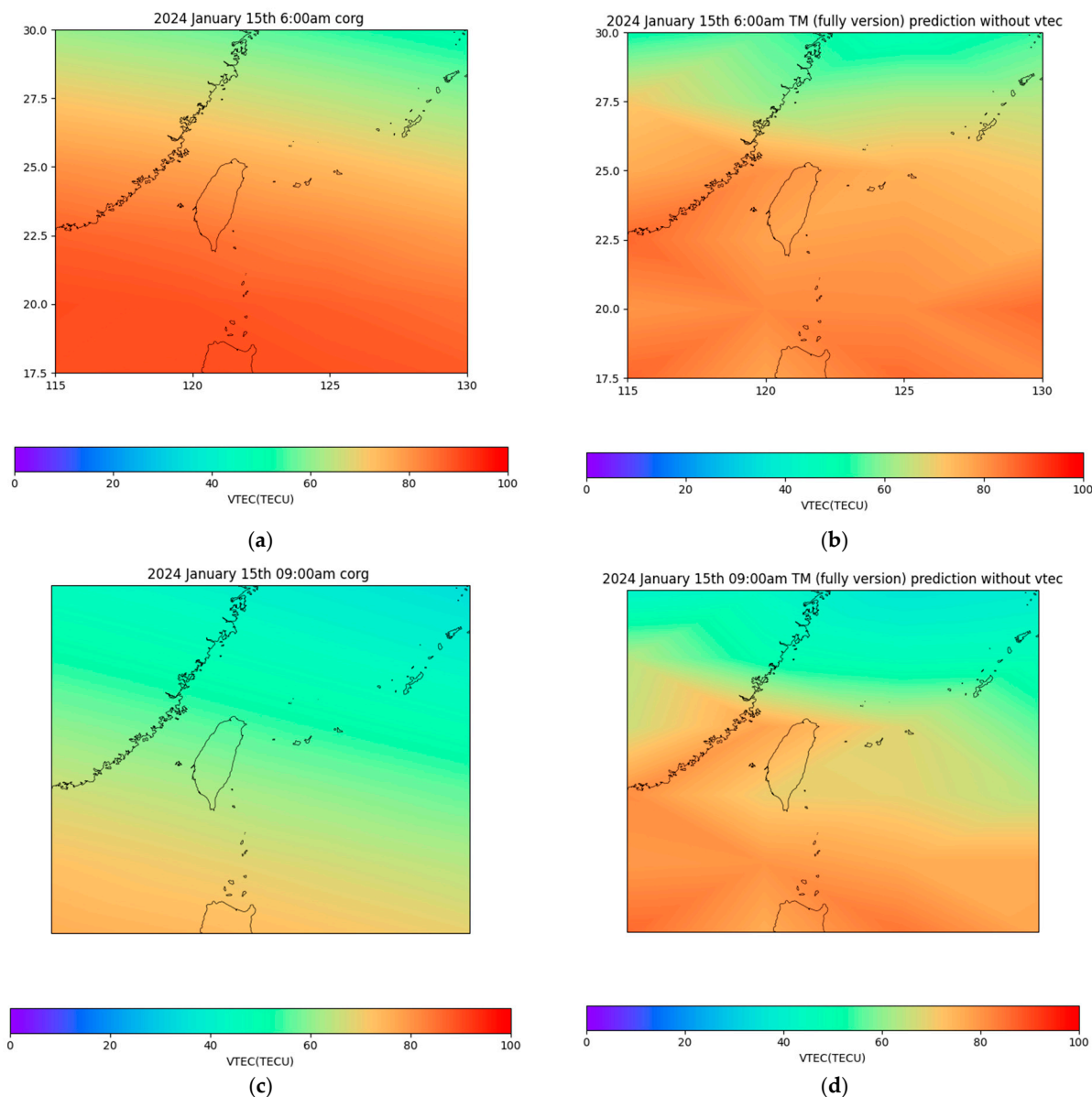
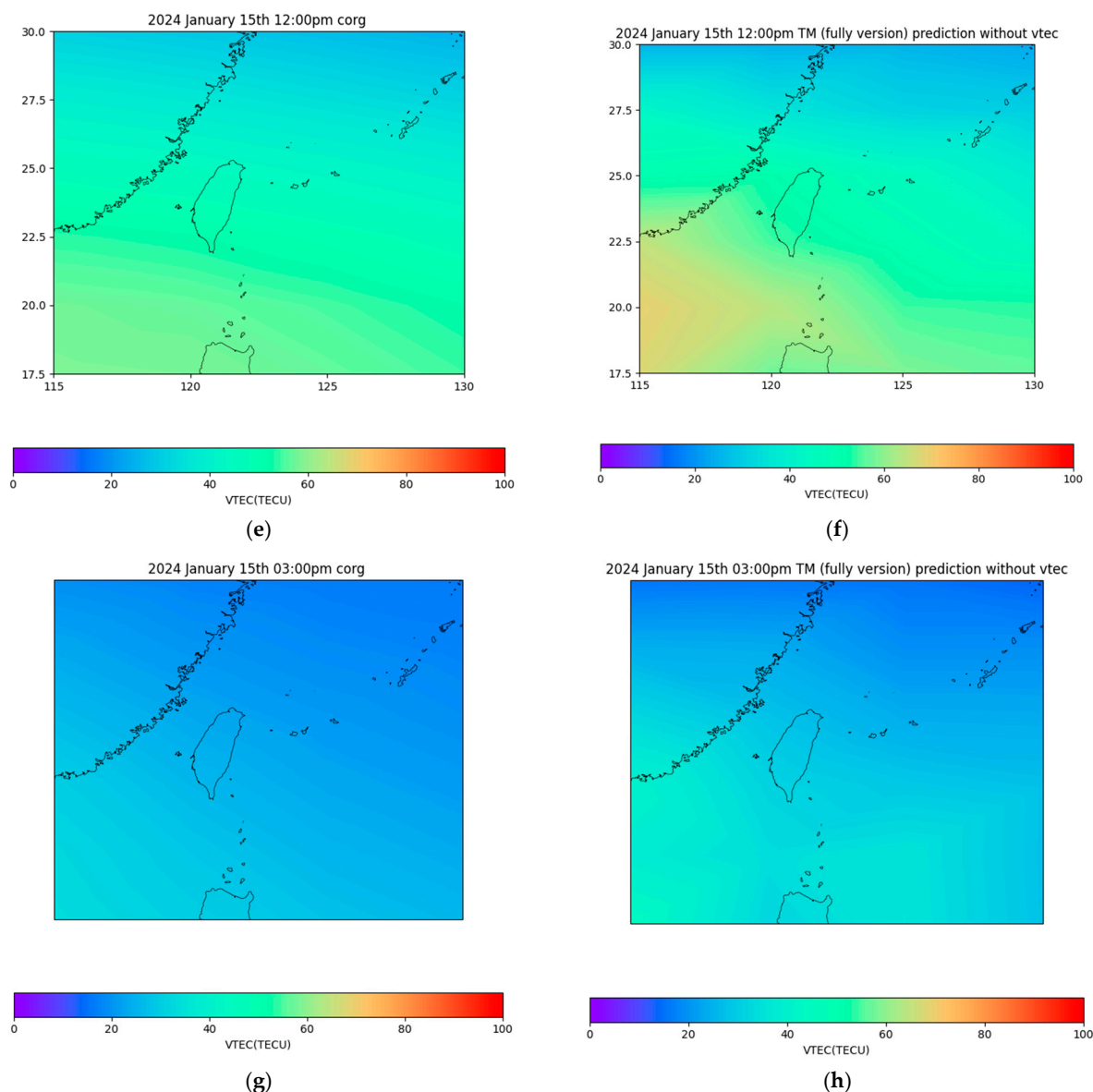


Figure 6. Cont.

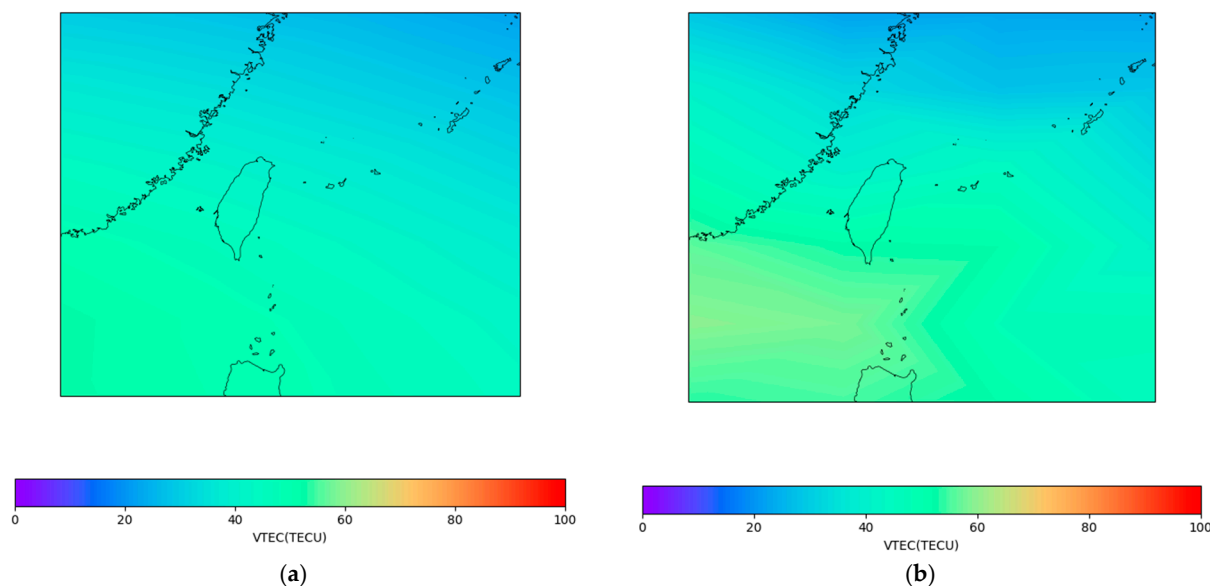




**Figure 6.** Diurnal variation of VTEC over the Taiwan region on 15 January 2024, showing the development and peak of the EIA. Left column: CORG VTEC maps. Right column: Transformer model predictions. From top to bottom: (a,b) 06:00 a.m. local time, before EIA development; (c,d) 09:00 a.m. local time, during EIA buildup; (e,f) 12:00 p.m. local time, approaching peak EIA intensity; (g,h) 15:00 local time, EIA transition. Color scale indicates VTEC values in TECU. The sequence demonstrates the model’s capability to capture the temporal evolution of the EIA, from minimal VTEC values in early morning to the formation of the characteristic enhanced VTEC band during peak hours.

To evaluate the practical impact of the VTEC predictions, positioning experiments were conducted using data from 21 February 2023, between 04:20 and 04:30 UTC. VTEC maps for this time period are shown in Figure 7. The study compared three methods for ionospheric delay correction:

- I. Dual-frequency receiver solution;
- II. CORG file VTEC values;
- III. Transformer model predictions.

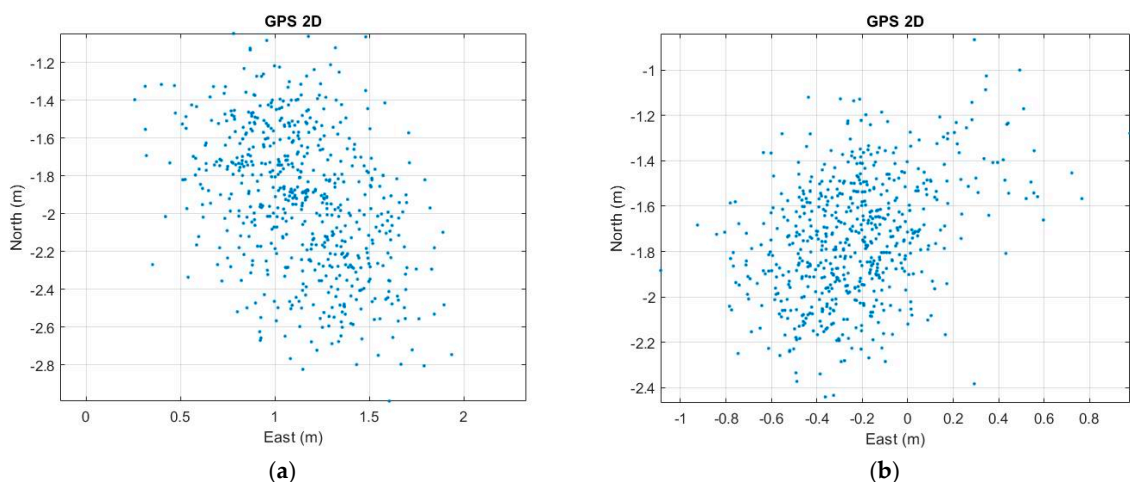


**Figure 7.** VTEC maps over Taiwan region during local midday (12:20–12:30 LT, 04:20–04:30 UT) on 21 February 2023, showing the northern EIA crest. (a) CORG VTEC map showing actual values. (b) Predicted VTEC map from our Transformer model. The band of enhanced VTEC values visible in both maps demonstrates the model’s ability to capture the EIA’s northern crest during its peak development period. Color scale indicates VTEC values in TECU.

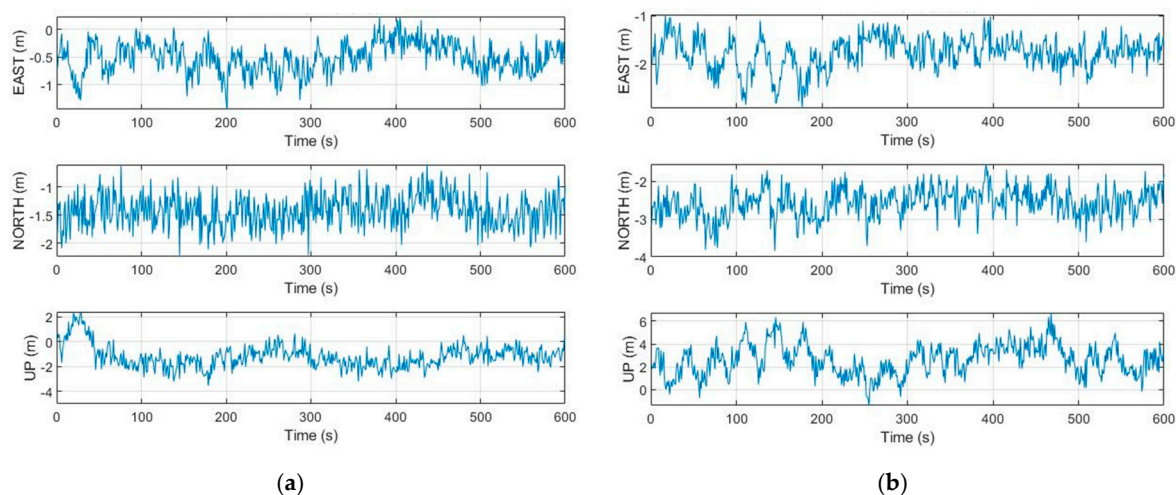
In all cases, we only used standard single-frequency code-phase measurements for position calculation. For the dual-frequency case, the L1–L2 combination is utilized only to eliminate the ionospheric effect. For the positioning experiments, we employed the Single Point Positioning (SPP) model using single-frequency code observations. We used broadcast ephemeris data to calculate satellite positions and clock corrections. We chose SPP with broadcast ephemeris to evaluate our model’s performance in a common user scenario, as this represents the most widely used positioning approach in mass-market GNSS receivers. While precise products could potentially provide better positioning accuracy, our focus was on comparing the relative performance of different ionospheric correction methods under typical user conditions. We emphasize that, to evaluate the practical impact of the VTEC predictions, our primary focus is on comparing the positioning results obtained using VTEC values from the CORG file and those predicted by our Transformer model. The dual-frequency receiver solution serves as a baseline for comparison but is not the main focus of our analysis. The positioning experiments were conducted under the following conditions:

- Test Site Details
  - Location: NTOU GNSS base station (25°09′00″N, 121°46′48″E)
  - Antenna Type: Trimble Zephyr Geodetic 2
  - Receiver Type: Trimble NetR9
  - Height: 32.5 m above sea level
- Data Collection Period
  - Date: 21 February 2023
  - Time: 04:20–04:30 UTC (12:20–12:30 Local Time)
  - Total Duration: 10 min
  - Satellite Configuration
- Number of visible satellites: 8–10 GPS satellites
  - Elevation cutoff angle: 10 degrees
  - PDOP range: 2.1–2.8

The positioning results shown in Figures 8 and 9 and the statistics presented in Tables 2 and 3 were computed using all available epochs during this period. The experiment used a single station to ensure consistent conditions for comparing different ionospheric correction methods. Figure 8 shows the 2D scatter plot of the GPS positioning. Figure 8b uses the Transformer model's predicted VTEC values. For comparison, Figure 8a presents the corresponding results using CORG file VTEC values. Figure 9 shows East–North–Up (ENU) error plots using the Transformer model's predicted VTEC values and the corresponding results using CORG file VTEC values. Tables 3 and 4 summarize the Root Mean Square (RMS) errors and standard deviations (STD) of the positioning errors for all methods, including the dual-frequency solution as a baseline.



**Figure 8.** The 2D scatter plot for GPS positioning. (a) Positioning results with the ionospheric effect calculated using CORG VTEC value. (b) Positioning results with the ionospheric effect calculated using the Transformer model's predicted VTEC value.



**Figure 9.** (a) ENU error plots for positioning using CORG file VTEC values. The graphs show the time series of positioning errors in the East, North, and Up directions. (b) ENU error plots for positioning using the Transformer model's predicted VTEC values.

**Table 3.** Positioning results—RMS (m).

Method	E	N	U
Dual-Frequency	0.4406	−1.1318	0.1653
CORG	−0.4008	−1.4018	−1.5746
Transformer	−1.775	−2.5720	2.6240

**Table 4.** Positioning results—STD (m).

Method	E	N	U
Dual-Frequency	0.2703	0.2659	0.5482
CORG	0.2442	0.3060	0.8494
Transformer	0.3399	0.2971	1.3876

The positioning errors show a bias rather than fluctuating around zero. After careful analysis, we believe this systematic error can be attributed to several factors:

- Use of Broadcast Ephemeris: The broadcast ephemeris contains residual orbit and clock errors that can introduce systematic biases in the position solution. While these errors are typically at the meter level, they do not average to zero over short time spans.
- Code Multipath Effects: In our experimental setup, code measurements are susceptible to multipath effects. Without employing multipath mitigation techniques, these effects can manifest as systematic errors in the position solution.
- Hardware Biases: Code measurements contain various hardware-related biases (both receiver and satellite) that were not explicitly accounted for in our processing strategy.

The experimental results show that while the Transformer model’s predictions for VTEC lead to positioning accuracy at the meter level, they do not yet achieve the same level of accuracy as the CORG data. The positioning errors using our model are approximately 1–2 m larger than those obtained using CORG data. This difference suggests that while our approach demonstrates the feasibility of using Transformer-based models for VTEC prediction, there is still room for improvement. Potential enhancements could include expanding the training dataset, incorporating additional input features, fine-tuning the model architecture, and improving the handling of extreme ionospheric conditions.

The current results should be viewed as a proof of concept, demonstrating that deep learning models can provide reasonable ionospheric corrections for GNSS positioning, while acknowledging the need for further refinement to match the performance of established methods. The Transformer model’s ability to learn from historical data implies that its predictive capabilities could be extended to longer time intervals. This scalability is a significant advantage over some traditional prediction methods. If provided with a more extensive training dataset covering a longer historical period, the Transformer model could potentially extend its prediction horizon. This could lead to more accurate long-term VTEC forecasts, which would be invaluable for various GNSS applications and space weather prediction.

Next, we would like to investigate the explainability of the proposed model. Prior to conducting the main experiment, we performed a correlation analysis on the input features using Spearman’s rank correlation coefficient. This analysis serves as a reference point for interpreting the explainability results of our model. Spearman’s rank correlation coefficient is a measure of the strength and direction of association between two variables. It is defined as the Pearson correlation coefficient between the rank variables. The coefficient ranges from  $-1$  to  $+1$ , where  $+1$  indicates a perfect positive correlation,  $-1$  indicates a perfect

negative correlation, and 0 indicates no correlation. The formula for Spearman's rank correlation coefficient (SRCC) is as follows:

$$\text{SRCC} = \text{Pearson}(S(X), S(Y)) = \frac{\text{cov}(S(X), S(Y))}{\sigma(S(X))\sigma(S(Y))}$$

where  $S(X)$  and  $S(Y)$  are the ranked versions of the original variables  $X$  and  $Y$ ,  $\text{cov}$  is the covariance, and  $\sigma$  is the standard deviation.

The data used for training span from 2022 to 2023, while the prediction data are from January 2024, comprising 24 data points per day. In this study, we focus exclusively on four grid points near Taiwan. These points are arranged from left to right and top to bottom, as follows (See Figure 1):

Grid Point 10: 120°E, 25°N

Grid Point 11: 125°E, 25°N

Grid Point 14: 120°E, 22.5°N

Grid Point 15: 125°E, 22.5°N

Figure 10a–d illustrates the rank correlation coefficients among these four grid points. In all cases, the last row indicates the correlation coefficients with VTEC of the present grid. These figures serve as references and comparisons for interpreting the experimental results. After reviewing the Spearman's rank correlation coefficients for the four grid points near Taiwan, it can be seen that the correlation patterns are largely consistent across all four grid points, suggesting that the relationships between variables are relatively stable in the region around Taiwan. Some specific correlations are summarized as follows:

(a) VTEC correlations:

- Strong positive correlation with F10.7 and Sunspot Number (around 0.7 to 0.8), confirming the known relationship between solar activity and ionospheric electron content.
- Moderate positive correlation with HRS (sine of hour, around 0.4 to 0.5), indicating a clear diurnal pattern in VTEC values.
- Weak to moderate negative correlation with HRC (cosine of hour, around  $-0.2$  to  $-0.3$ ), further supporting the diurnal variation.
- Very weak correlations with DNS and DNC (day number sine and cosine), suggesting that seasonal variations might be less prominent than daily variations in this dataset.

(b) Solar activity indicator:

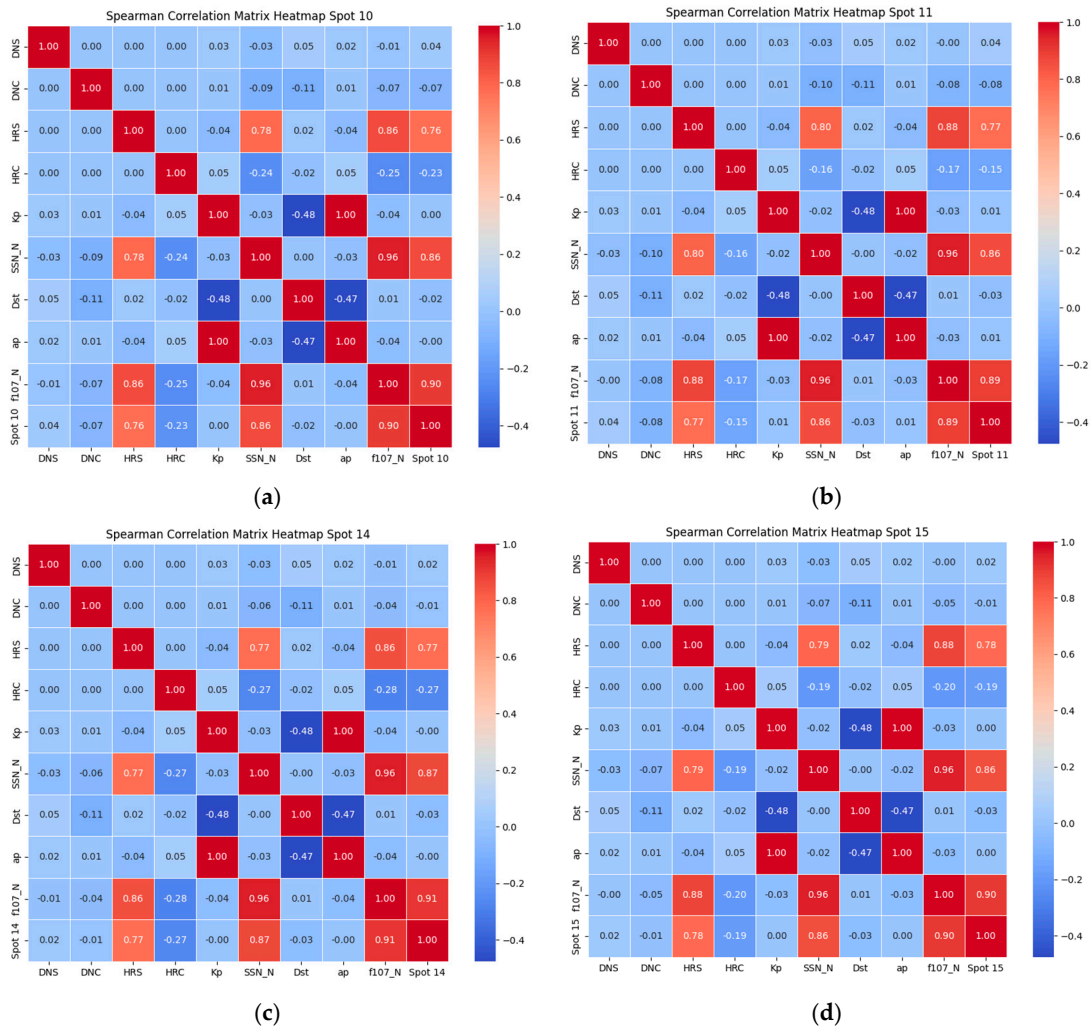
- Very strong correlation between F10.7 and Sunspot Number ( $>0.9$ ), as expected, since both are indicators of solar activity.
- These solar activity indicators show similar correlation patterns with other variables, reinforcing their interchangeability in many applications.

(c) Geomagnetic indices:

- $A_p$  and  $K_p$  indices show very strong correlation ( $>0.9$ ), which is expected, as they are closely related measures of geomagnetic activity.
- Dst index shows moderate negative correlations with  $A_p$  and  $K_p$  (around  $-0.5$  to  $-0.6$ ), aligning with the understanding that strong geomagnetic storms (indicated by negative Dst) often correspond to higher  $A_p$  and  $K_p$  values.
- Interestingly, the geomagnetic indices ( $A_p$ ,  $K_p$ , Dst) show very weak correlations with VTEC in this dataset, which might be unexpected given the known influence of geomagnetic activity on the ionosphere.

(d) Time functions:

- HRS and HRC show the expected orthogonal relationship (correlation close to 0), validating their use as independent components of the diurnal cycle.
- Similarly, DNS and DNC show near-zero correlation, correctly representing independent components of the annual cycle.



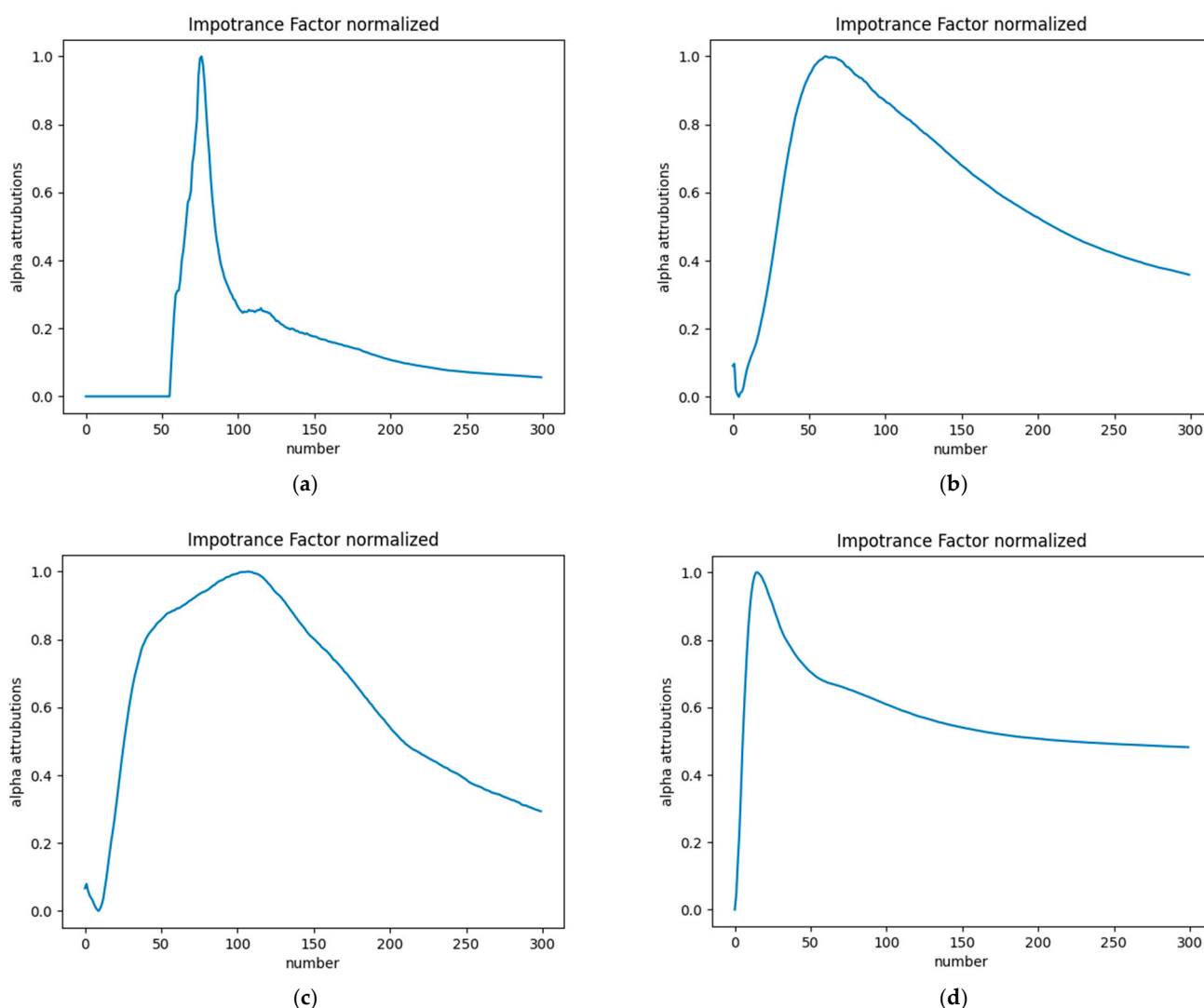
**Figure 10.** Spearman’s rank correlation coefficient matrix for input features and VTEC at four grid points. The matrix displays the strength and direction of correlations between VTEC, solar activity indicators, geomagnetic indices, and time functions. Color intensity represents the magnitude of correlation, with red indicating positive correlations and blue indicating negative correlations. (a) Grid point 10. (b) Grid point 11. (c) Grid point 14. (d) Grid point 15.

These observations provide a valuable baseline for interpreting the Transformer model’s behavior and feature importance. Any significant deviations from these correlation patterns in the model’s learned representations would be particularly interesting to investigate further.

We are now ready to employ the IDG method. This approach calculates an additional gradient of the interpolation constant with respect to the sample to obtain a weight for each interpolated sample. By multiplying the interpolation contribution with this weight, we effectively reduce the impact of noise on the explanations. The IDG method integrates the entire gradient path and additionally integrates each interpolation sample step to minimize noise, resulting in cleaner explanation results. In our experiment, we used 300 interpolation samples. The interpolation samples are defined using the integral gradient to determine

their contribution. The IDG method uses important factors to eliminate the saturation effect that usually occurs in the integrated gradient approach.

Figure 11 presents the results of the IDG method as applied to the VTEC prediction model for four specific grid points. These figures illustrate the standardized interpolation sample weights, providing insights into the model's decision-making process and the relative importance of different interpolation steps in VTEC prediction. The differences in weight distributions across grid points indicate that the model adapts its decision-making process based on geographical location.



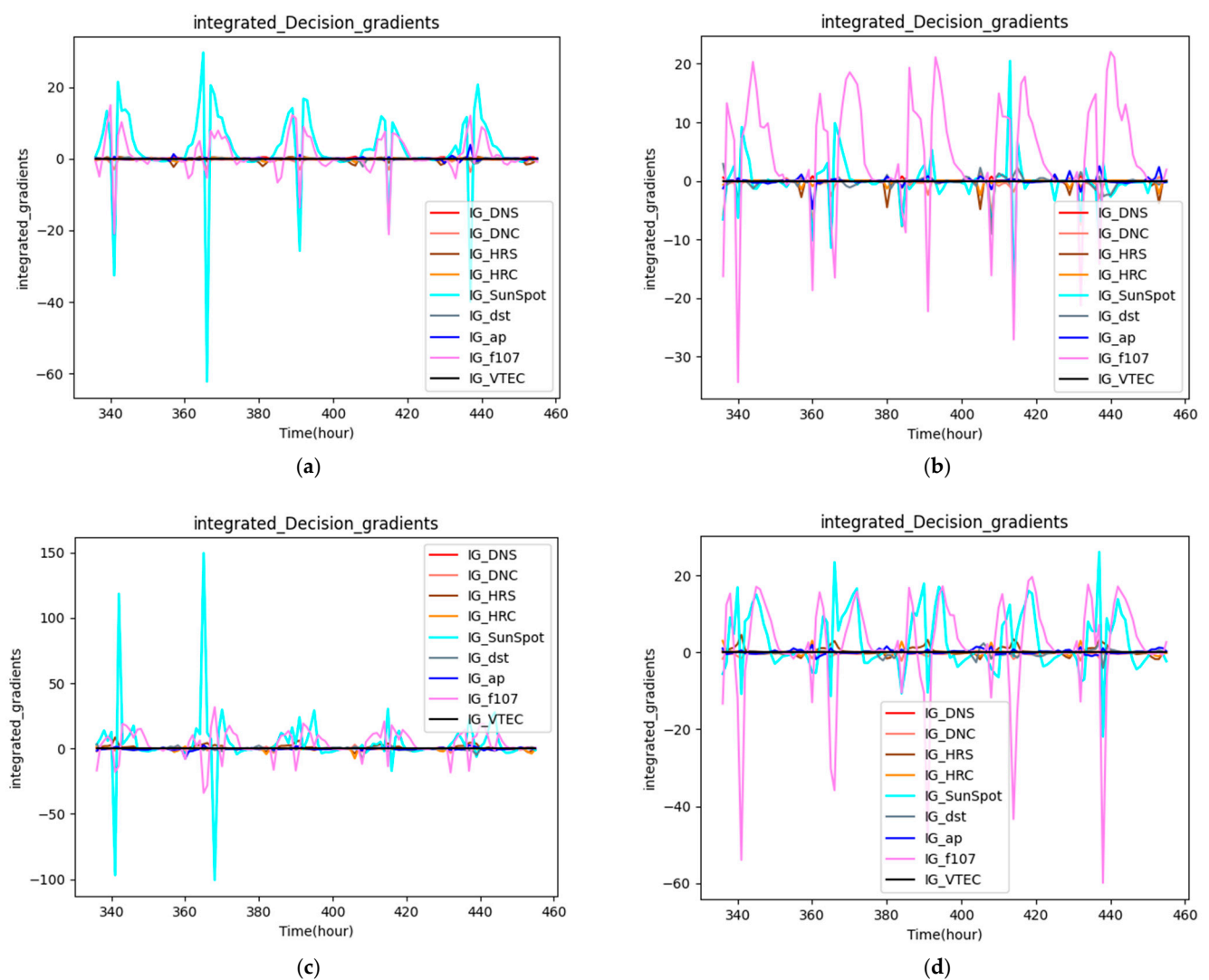
**Figure 11.** Standardized interpolation sample weights derived from IDG method for four grid points in the VTEC prediction model. (a) Grid point 10 (120°E, 25°N); (b) Grid point 11 (125°E, 25°N); (c) Grid point 14 (120°E, 22.5°N); (d) Grid point 15 (125°E, 22.5°N). The x-axis represents 300 interpolation steps from the baseline (step 0) to the actual input (step 300). The y-axis shows the standardized weight (0–1), indicating the relative importance of each interpolation step in the model's decision-making process. These plots illustrate the spatial variability in feature importance and the model's adaptive behavior across different geographical locations. Note the varying patterns between northern (a,b) and southern (c,d) grid points, suggesting latitude-dependent prediction strategies in the VTEC model.

All four figures show very low weights for the initial interpolation steps, suggesting that baseline values have minimal impact on the model's predictions. The generally

increasing trends in all figures imply that the model considers cumulative changes in features rather than relying on specific threshold values.

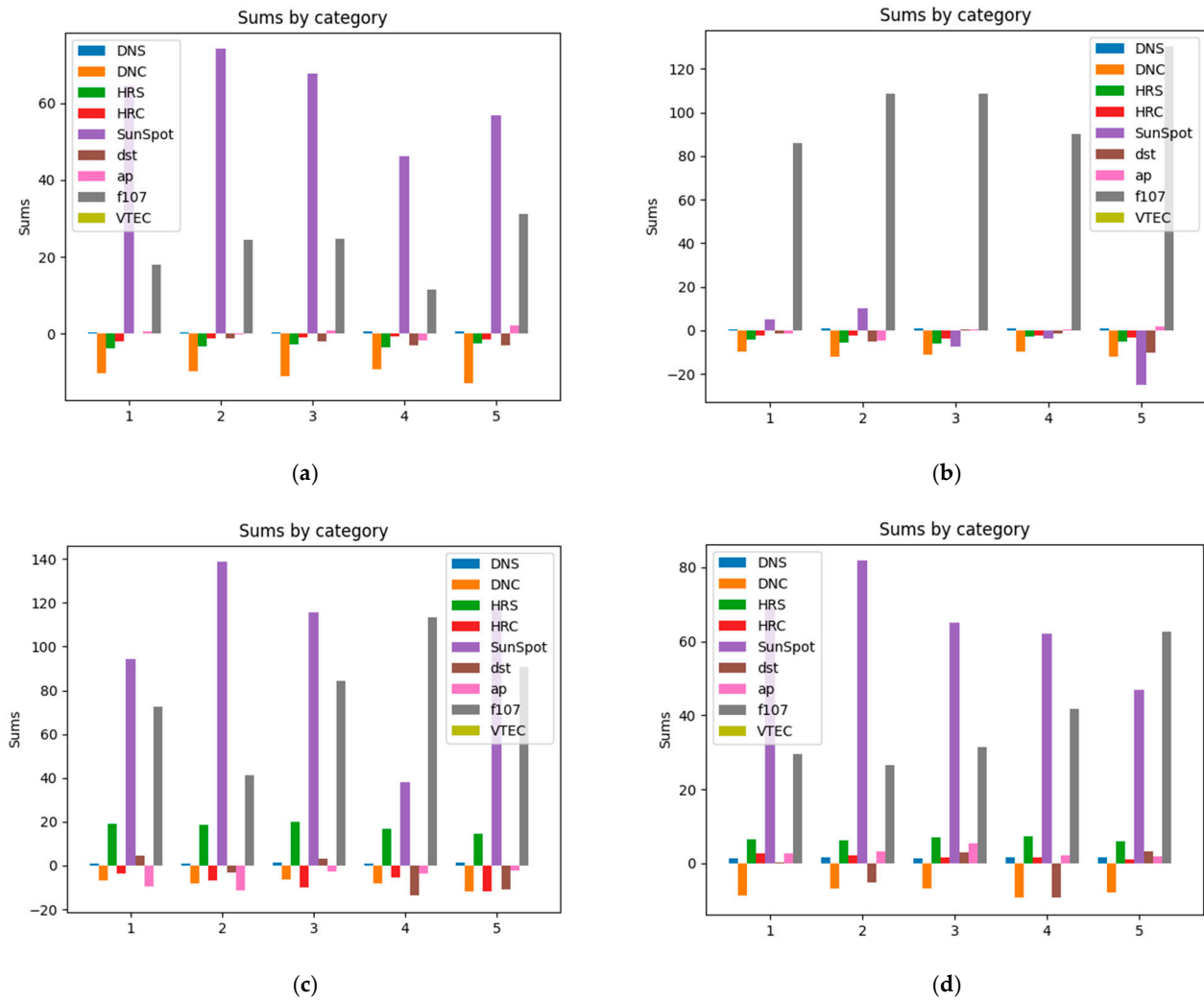
The southern grid points (14 and 15) show more linear weight increases, while the northern points (10 and 11) exhibit more pronounced peaks, suggesting latitude-dependent prediction strategies. The varied patterns across grid points highlight the complexity of the model and its ability to capture location-specific ionospheric behaviors. These figures provide valuable insights into the inner workings of the VTEC prediction model, demonstrating its adaptive nature and the complex spatial dependencies in ionospheric modeling.

The next experimental results show the gradient scores (Figure 12) and cumulative gradient scores (Figure 13) for each feature at four grid points over a five-day period. These figures provide insights into the importance and impact of different features on the VTEC prediction model. Each figure contains multiple lines, representing different features: Sunspot Number, F10.7, Dst, Ap, DNS, DNC, HRS, and HRC.



**Figure 12.** Gradient scores for VTEC prediction features across four grid points. (a) Grid point 10 (120°E, 25°N); (b) Grid point 11 (125°E, 25°N); (c) Grid point 14 (120°E, 22.5°N); (d) Grid point 15 (125°E, 22.5°N). Each line represents a different feature: Sunspot Number, F10.7, Dst, Ap, DNS, DNC, HRS, and HRC. The y-axis shows the gradient score, indicating the instantaneous impact of each feature on the VTEC prediction. Positive values suggest an increase in VTEC, while negative values indicate a decrease. Note the varying patterns and magnitudes across different grid points, revealing the spatial dependency of feature importance in the model.





**Figure 13.** Cumulative gradient scores for VTEC prediction features across four grid points. (a) Grid point 10 ( $120^{\circ}\text{E}$ ,  $25^{\circ}\text{N}$ ); (b) Grid point 11 ( $125^{\circ}\text{E}$ ,  $25^{\circ}\text{N}$ ); (c) Grid point 14 ( $120^{\circ}\text{E}$ ,  $22.5^{\circ}\text{N}$ ); (d) Grid point 15 ( $125^{\circ}\text{E}$ ,  $22.5^{\circ}\text{N}$ ). Each line represents the cumulative effect of a different feature: Sunspot Number, F10.7, Dst, Ap, DNS, DNC, HRS, and HRC. The y-axis shows the cumulative gradient score, illustrating the overall impact of each feature on the VTEC prediction over time. The diverging lines highlight the relative importance and long-term influence of different features. Observe the dominance of solar activity indicators (Sunspot Number and F10.7) and the consistent negative contribution of DNC across all grid points.

In these figures, solar activity indicators show similar patterns across all grid points. They often have opposite contributions (one positive, one negative) in the gradient scores. In corresponding cumulative plots, they tend to have the largest magnitudes, indicating their significant impact on VTEC predictions.

Geomagnetic indices generally show smaller contributions compared to solar activity indicators. The Dst index often displays negative contributions, while the Ap index contributions vary. Their impact becomes more noticeable in the cumulative plots, especially during periods of geomagnetic disturbance.

Time functions DNS and DNC show relatively small but consistent contributions. HRS and HRC display more pronounced oscillations in gradient scores. DNC consistently shows a notable negative contribution across all grid points. Clear daily cycles are visible in the gradient scores, particularly for HRS and HRC. The cumulative plots show how these daily variations accumulate over time.

While overall patterns are similar, there are noticeable differences between grid points. Grid points 14 and 15 (southern points) often show more pronounced variations in gradient scores compared to grid points 10 and 11 (northern points).

## 5. Discussion

This study has explored the application of a Transformer-based model for VTEC prediction, coupled with advanced explainability techniques to understand the model's decision-making process. The research has yielded plausible insights into both the effectiveness of deep learning for ionospheric modeling and the interpretability of complex AI models in geophysical applications. The Transformer model demonstrated impressive capabilities in VTEC prediction, showing comparable performance to traditional methods. The model showed the ability to capture spatial variations in VTEC, adapting its predictions across different grid points. This adaptability is crucial for accurate regional ionospheric mapping. While already competitive, the results suggest that with more extensive training data and further refinement, the Transformer model could potentially outperform traditional methods, especially in capturing complex, non-linear ionospheric behaviors.

The IDG analysis offered deep insights into the model's internal decision-making process. The analysis revealed how the importance of different features varied across the interpolation steps, providing a nuanced view of feature contributions. Differences in IDG patterns across grid points demonstrated the model's ability to adapt its decision-making process based on geographical location. The IDG method effectively reduced noise in the explanations, providing clearer insights into feature contributions compared to traditional gradient-based methods.

The gradient score analysis further elucidated the model's behavior and feature utilization. Consistent with correlation analysis, solar activity indicators emerged as the most influential features, but the analysis also revealed the significant contributions of other features like DNC, which were not apparent from simple correlation studies. The analysis captured both the instantaneous (gradient scores) and cumulative (cumulative gradient scores) impacts of features, providing a comprehensive view of their influence over time. Variations in gradient scores across grid points further confirmed the model's ability to adapt to different geographical locations, potentially capturing localized ionospheric phenomena. The often opposing contributions of related features (e.g., Sunspot Number and F10.7) highlighted the model's ability to capture subtle, complex interactions between features.

An interesting observation from the gradient score analysis is the relatively small variations and contributions shown by the DNS and DNC feature lines (Figure 9). While it might be tempting to consider removing these seasonal indicators from the model, several factors support their retention. First, these features capture fundamental physical processes—specifically, seasonal variations in solar zenith angle and day length that are known to influence ionospheric behavior. Their smaller contributions in our study might be particularly related to our focus on the region around Taiwan, where seasonal effects are less prominent than diurnal variations. Second, while their direct contributions appear small, these features provide stable baseline information that helps the model maintain awareness of the longer-term seasonal context. Finally, the gradient scores only show direct contributions and may not fully capture how these features interact with other inputs to improve overall prediction accuracy. A comprehensive ablation study would be necessary to definitively determine the impact of these seasonal indicators on model performance.

## 6. Conclusions

In this paper, we demonstrate the potential of deep learning, specifically Transformer models, in advancing the field of ionospheric modeling and VTEC prediction. The application of explainability techniques (IDG and gradient score analysis) to a complex geophysical model may open new research, interesting for interpreting AI models in earth and space sciences. The combination of deep learning and explainable AI techniques in this geophysical application shows the potential for similar approaches in other scientific domains dealing with complex, multi-variate time series data. Future work should include comparative analysis with other deep learning architectures such as LSTM, CNN-LSTM hybrid, and BiLSTM models. Such comparisons would provide quantitative insights into the relative strengths and limitations of different approaches for VTEC prediction. Performance metrics like RMSE, MAE, and MAPE could be used to evaluate prediction accuracy, while also considering computational efficiency and model complexity. This would help establish more comprehensive guidelines for choosing appropriate deep learning architectures for ionospheric prediction tasks.

Additionally, future work might include comprehensive ablation studies to quantify the impact of XAI on model performance. Such studies would compare different model configurations: a base Transformer model without XAI guidance, a model with basic gradient-based analysis, and our full model with IDG-based XAI. This would help isolate the effects of different XAI components and their contributions to prediction accuracy. The ablation analysis could examine how XAI influences feature selection, model refinement, and the handling of edge cases. Additionally, it could investigate whether insights from XAI can guide architectural improvements and hyperparameter optimization. Such studies would contribute to our understanding of how explainability techniques can be leveraged to enhance deep learning models for ionospheric prediction.

**Author Contributions:** Conceptualization, H.-S.W. and D.-J.J.; methodology, H.-S.W. and Y.-H.L.; software, Y.-H.L. and H.-S.W.; validation, H.-S.W. and D.-J.J.; writing—original draft preparation, H.-S.W.; writing—review and editing, H.-S.W. and D.-J.J.; supervision, H.-S.W. and D.-J.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors gratefully acknowledge the support of the National Science and Technology Council, Taiwan, under grant number NSTC 113-2221-E-019-059.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

## References

1. Kaplan, E.; Hegarty, C. *Understanding GPS/GNSS: Principles and Applications*, 3rd ed; Artech House: Boston, MA, USA, 2017.
2. Klobuchar, J.A. Ionospheric time-delay algorithm for single-frequency GPS users. *IEEE Trans. Aerosp. Electron. Syst.* **1987**, *3*, 325–331. [[CrossRef](#)]
3. Dubey, S.; Wahi, R.; Gwal, A.K. Ionospheric Effects on GPS Positioning. *Adv. Space Res.* **2006**, *38*, 2478–2484. [[CrossRef](#)]
4. Lyu, S.; Xiang, Y.; Soja, B.; Wang, N.; Yu, W.; Truong, T.-K. Uncertainties of interpolating satellite-specific slant ionospheric delays and impacts on PPP-RTK. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *60*, 490–505. [[CrossRef](#)]
5. Tang, L.; Zhang, F.; Li, P.; Deng, Y.; Chen, W. Effects of equatorial plasma bubble-induced ionospheric gradients on GNSS PPP-RTK. *GPS Solut.* **2024**, *28*, 124. [[CrossRef](#)]
6. Hagfors, T. 2.1. The Ionosphere. In *Methods in Experimental Physics*; Meeks, M.L., Ed.; Astrophysics; Academic Press: Cambridge, MA, USA, 1976; Volume 12, pp. 119–135.
7. Tulunay, Y.; Tulunay, E.; Senalp, E.T. The neural network technique—1: A general exposition. *Adv. Space Res.* **2004**, *33*, 983–987. [[CrossRef](#)]
8. Tulunay, Y.; Tulunay, E.; Senalp, E.T. The neural network technique—2: An ionospheric example illustrating its application. *Adv. Space Res.* **2004**, *33*, 988–992. [[CrossRef](#)]

9. Tulasi Ram, S.; Sai Gowtam, V.; Mitra, A.; Reinisch, B. The Improved Two-Dimensional Artificial Neural Network-Based Ionospheric Model (ANNIM). *J. Geophys. Res. Space Phys.* **2018**, *123*, 5807–5820. [[CrossRef](#)]
10. Chen, Z.; Liao, W.; Li, H.; Wang, J.; Deng, X.; Hong, S. Prediction of Global Ionospheric TEC Based on Deep Learning. *Space Weather* **2022**, *20*, e2021SW002854. [[CrossRef](#)]
11. Luo, H.; Gong, Y.; Chen, S.; Yu, C.; Yang, G.; Yu, F.; Hu, Z.; Tian, X. Prediction of Global Ionospheric Total Electron Content (TEC) Based on SAM-ConvLSTM Model. *Space Weather* **2023**, *21*, e2023SW003707. [[CrossRef](#)]
12. Sun, W.; Xu, L.; Huang, X.; Zhang, W.; Yuan, T.; Yan, Y. Bidirectional LSTM for ionospheric vertical Total Electron Content (TEC) forecasting. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10–13 December 2017; pp. 1–4.
13. Wang, H.; Liu, H.; Yuan, J.; Le, H.; Shan, W.; Li, L. MAOOA-Residual-Attention-BiConvLSTM: An Automated Deep Learning Framework for Global TEC Map Prediction. *Space Weather* **2024**, *22*, e2024SW003954. [[CrossRef](#)]
14. Ren, X.; Zhao, B.; Ren, Z.; Xiong, B. Ionospheric TEC Prediction in China during Storm Periods Based on Deep Learning: Mixed CNN-BiLSTM Method. *Remote Sens.* **2024**, *16*, 3160. [[CrossRef](#)]
15. Ali, S.; Abuhmed, T.; El-Sappagh, S.; Muhammad, K.; Alonso-Moral, J.M.; Confalonieri, R.; Guidotti, R.; Del Ser, J.; Díaz-Rodríguez, N.; Herrera, F. Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Inf. Fusion* **2023**, *99*, 101805. [[CrossRef](#)]
16. Markus, A.F.; Kors, J.A.; Rijnbeek, P.R. The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. *J. Biomed. Inform.* **2021**, *113*, 103655. [[CrossRef](#)] [[PubMed](#)]
17. McKinnell, L.A.; Poole, A.W.V. Predicting the ionospheric *F* layer using neural networks. *J. Geophys. Res.* **2004**, *109*, A08308. [[CrossRef](#)]
18. Sivavaraprasad, G.; Deepika, V.S.; SreenivasaRao, D.; Ravi Kumar, M.; Sridhar, M. Performance evaluation of neural network TEC forecasting models over equatorial low-latitude Indian GNSS station. *Geod. Geodyn.* **2020**, *11*, 192–201. [[CrossRef](#)]
19. Sundararajan, M.; Taly, A.; Yan, Q. Axiomatic Attribution for Deep Networks. In Proceedings of the International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017.
20. Smilkov, D.; Thorat, N.; Kim, B.; Viégas, F.; Wattenberg, M. Smoothgrad: Removing noise by adding noise. *arXiv* **2024**, arXiv:1706.03825.
21. Walker, C.; Jha, S.K.; Chen, K.; Ewetz, R. Integrated Decision Gradients: Compute Your Attributions Where the Model Makes Its Decision. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 20–27 February 2024.
22. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [[CrossRef](#)]
23. Samek, W.; Montavon, G.; Lapuschkin, S.; Anders, C.J.; Müller, K.R. Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications. *Proc. IEEE* **2021**, *109*, 247–278. [[CrossRef](#)]
24. Kelley, M.C. *The Earth's Ionosphere: Plasma Physics and Electrodynamics*, 2nd ed.; Academic Press: San Diego, CA, USA, 2009.
25. Mannucci, A.J.; Wilson, B.D.; Yuan, D.N.; Ho, C.H.; Lindqwister, U.J.; Runge, T.F. A global mapping technique for GPS-derived ionospheric total electron content measurements. *Radio Sci.* **1998**, *33*, 565–582. [[CrossRef](#)]
26. Tsurutani, B.T.; Gonzalez, W.D.; Tang, F.; Akasofu, S.I.; Smith, E.J. Origin of interplanetary southward magnetic fields responsible for major magnetic storms near solar maximum (1978–1979). *J. Geophys. Res.* **1988**, *93*, 8519–8531. [[CrossRef](#)]
27. Basu, S.; Basu, S.; Groves, K.M.; Yeh, H.-C.; Su, S.-Y.; Rich, F.J.; Sultan, P.J.; Keskinen, M.J. Response of the equatorial ionosphere in the South Atlantic Region to the Great Magnetic Storm of July 15, 2000. *Geophys. Res. Lett.* **2001**, *28*, 3577–3580. [[CrossRef](#)]
28. Willard, H.R.; Kenney, J.K. Ionospheric effects of high-altitude nuclear tests. *J. Geophys. Res.* **1963**, *68*, 2053–2056. [[CrossRef](#)]
29. Mendillo, M.; Hawkins, G.S.; Klobuchar, J.A. A sudden vanishing of the ionospheric *F* region due to the launch of Skylab. *J. Geophys. Res.* **1975**, *80*, 2217–2228. [[CrossRef](#)]
30. Nava, B.; Coisson, P.; Radicella, S.M. A new version of the NeQuick ionosphere electron density model. *J. Atmos. Sol. Terr. Phys.* **2005**, *70*, 1856–1862. [[CrossRef](#)]
31. Yuan, Y.; Wang, N.; Li, Z.; Huo, X. The BeiDou global broadcast ionospheric delay correction model (BDGIM) and its preliminary performance evaluation. *Navigation* **2019**, *66*, 55–69. [[CrossRef](#)]
32. Li, Z.; Yuan, Y.; Wang, N.; Hernandez-Pajares, M.; Huo, X. SHPTS: Towards a new method for generating precise global ionospheric TEC map based on spherical harmonic and generalized trigonometric series functions. *J. Geod.* **2015**, *89*, 331–345. [[CrossRef](#)]
33. Chen, J.; Zhang, Y.; Wang, J.; Yang, S.; Dong, D.; Zhou, J.; Wang, M.; Wang, J. A simplified and unified model of multi-GNSS precise point positioning. *Adv. Space Res.* **2015**, *55*, 125–134. [[CrossRef](#)]
34. Bilitza, D. IRI the International Standard for the Ionosphere. *Adv. Radio Sci.* **2018**, *16*, 1–11. [[CrossRef](#)]

35. Hernández-Pajares, M.; Juan, J.M.; Sanz, J. GPS ionospheric tomography: A review of real time implementation and challenges. *GPS Solut.* **2009**, *13*, 103–112.
36. Wielgosz, P.; Milanowska, B.; Krypiak-Gregorczyk, A.; Jarmołowski, W. Validation of GNSS-derived global ionosphere maps for different solar activity levels: Case studies for years 2014 and 2018. *GPS Solut.* **2021**, *25*, 103. [[CrossRef](#)]
37. Tang, L.; Zhang, F.; Chen, W. The Error of Global Ionospheric Map-TEC During Equatorial Plasma Bubble Event in the High Solar Activity Year. *Space Weather* **2024**, *22*, e2023SW003714. [[CrossRef](#)]
38. Huang, Y.N.; Cheng, K. Solar cycle variations of the equatorial ionospheric anomaly in total electron content in the Asian region. *J. Geophys. Res. Space Phys.* **1996**, *101*, 24513–24520. [[CrossRef](#)]
39. Anderson, D.N. A theoretical study of the ionospheric F region equatorial anomaly—I. Theory. *Planet. Space Sci.* **1973**, *21*, 409–419. [[CrossRef](#)]
40. Lin, C.H.; Liu, J.Y.; Tsai, H.F.; Cheng, C.Z. Variations in the equatorial ionization anomaly peaks in the Western Pacific region during the geomagnetic storms of April 6 and July 15, 2000. *Earth Planets Space* **2007**, *59*, 401–405. [[CrossRef](#)]
41. Mo, X.; Zhang, D. A comparative study of the northern and southern equatorial ionization anomaly crests in the East-Asian sector during 2006–2015. *Adv. Space Res.* **2021**, *68*, 1461–1472. [[CrossRef](#)]
42. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, CA, USA, 4–9 December 2017.

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