Using Transformers to Improve Ionospheric Modeling and Enable a Wider Input Space

Liam Smith (lsmith377@gatech.edu), Morris Cohen
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

CV Link

UCAR SW Workshop 2025
Poster 30

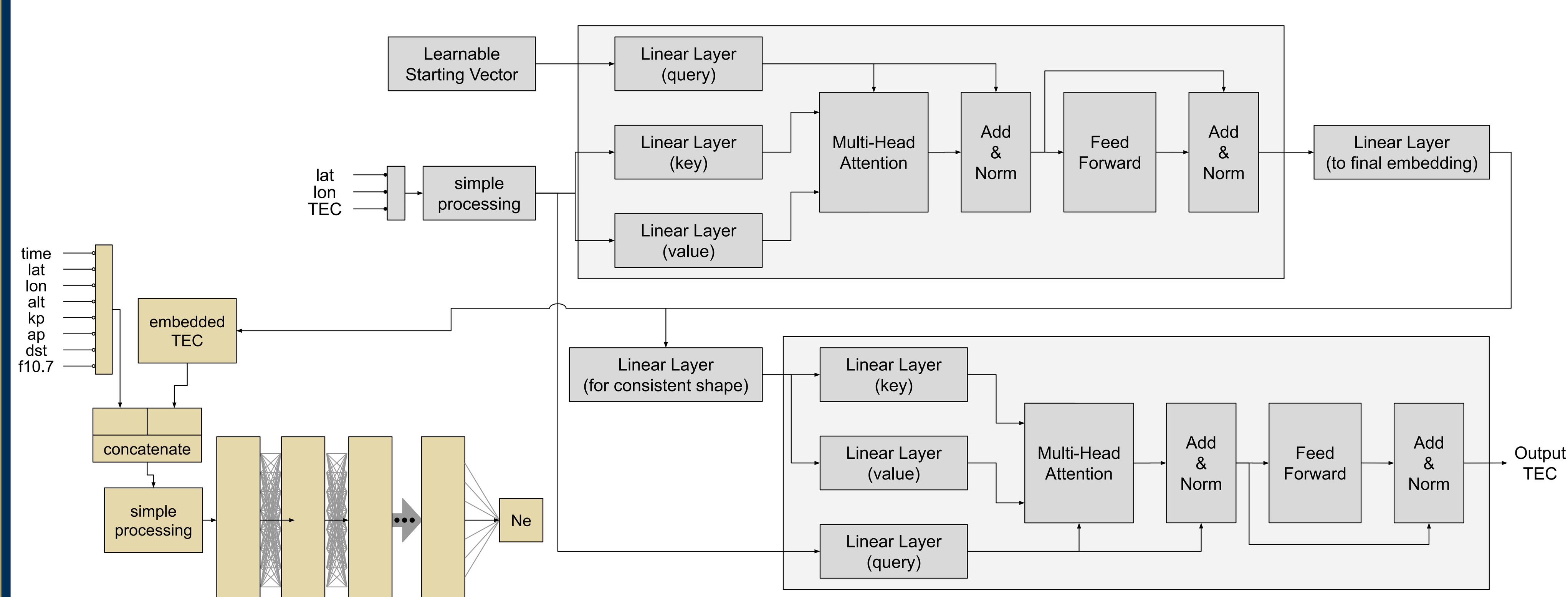
Motivation

Modeling ionospheric electron density is critical for improving satellite communication and understanding space weather. The integral of electron density, Total Electron Content (TEC), is heavily correlated and much more readily available. However, TEC is oddly shaped, which poses difficulties for traditional ML methods because of irregular shape and length. To address this, we focus on two tasks:

- Develop a technique to embed irregularly shaped data into fixed-length vectors using cross-attention.
- Apply the embedding approach to TEC and use the embeddings to improve ionospheric electron density predictions.

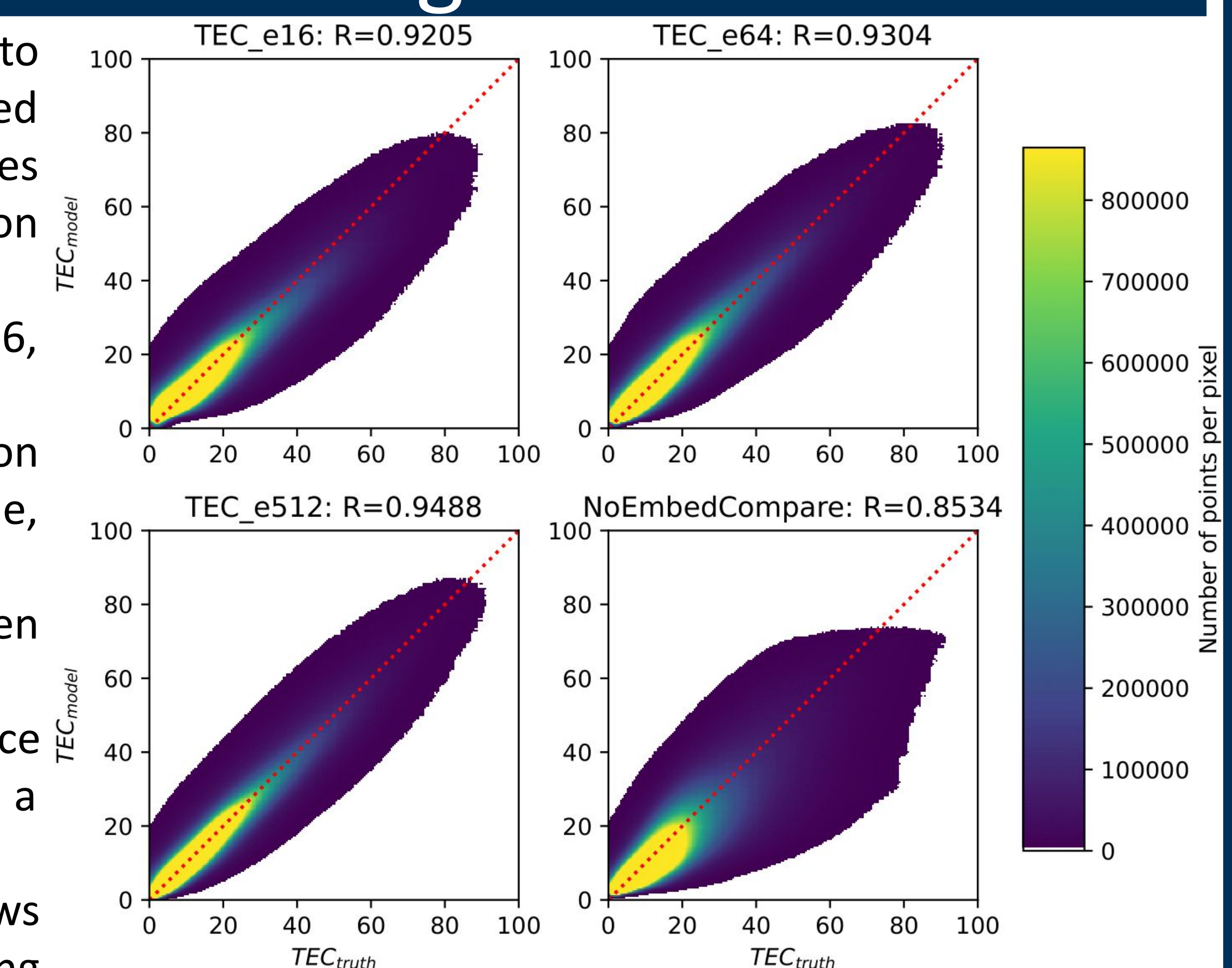
Methodology

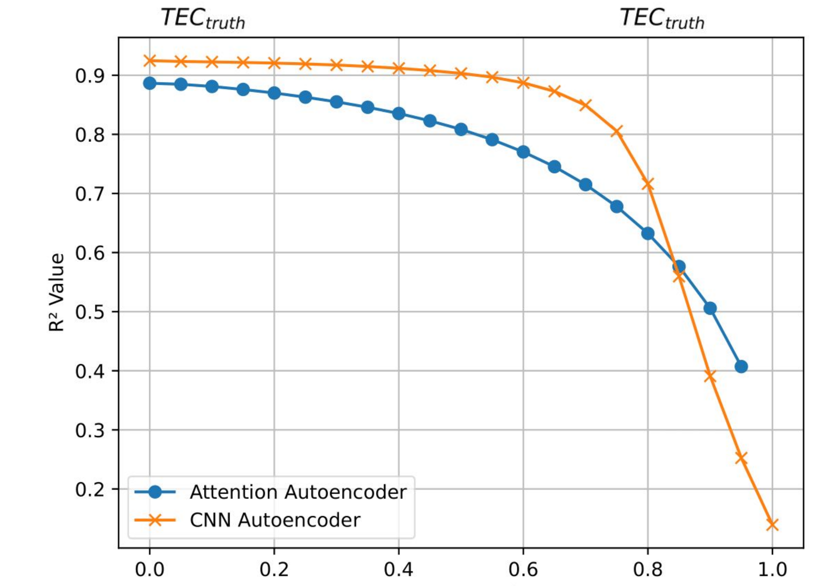
- Embedding TEC data allows us to convert irregularly shaped sequences into fixed-length vectors for machine learning models, which we train in an autoencoding approach using transformer-based models.
  - An encoder creates fixed-length embeddings, which can be used in other models.
  - A decoder reconstructs TEC from embeddings, which allows for training of the encoder.
- We model electron density with a feed-forward neural network.
  - Standard Inputs include location, time, and various indices (kp, ap, dst, and f10.7).
  - We augment the standard inputs with the embedded TEC.
- Data: TEC (Madrigal), e^- density (CHAMP, GRACE, COSMIC-1/2), hmf2/nmf2 (GIRO ionosondes)
  - Splits: training (pre 2020), validation (2021), testing (2022–2023).
  - Preprocessing: TEC normalized; electron density and ionosonde data filtered and scaled.
  - Inputs: Geomagnetic indices (kp, ap, dst, f10.7) from NASA OMNIWeb



TEC Embedding

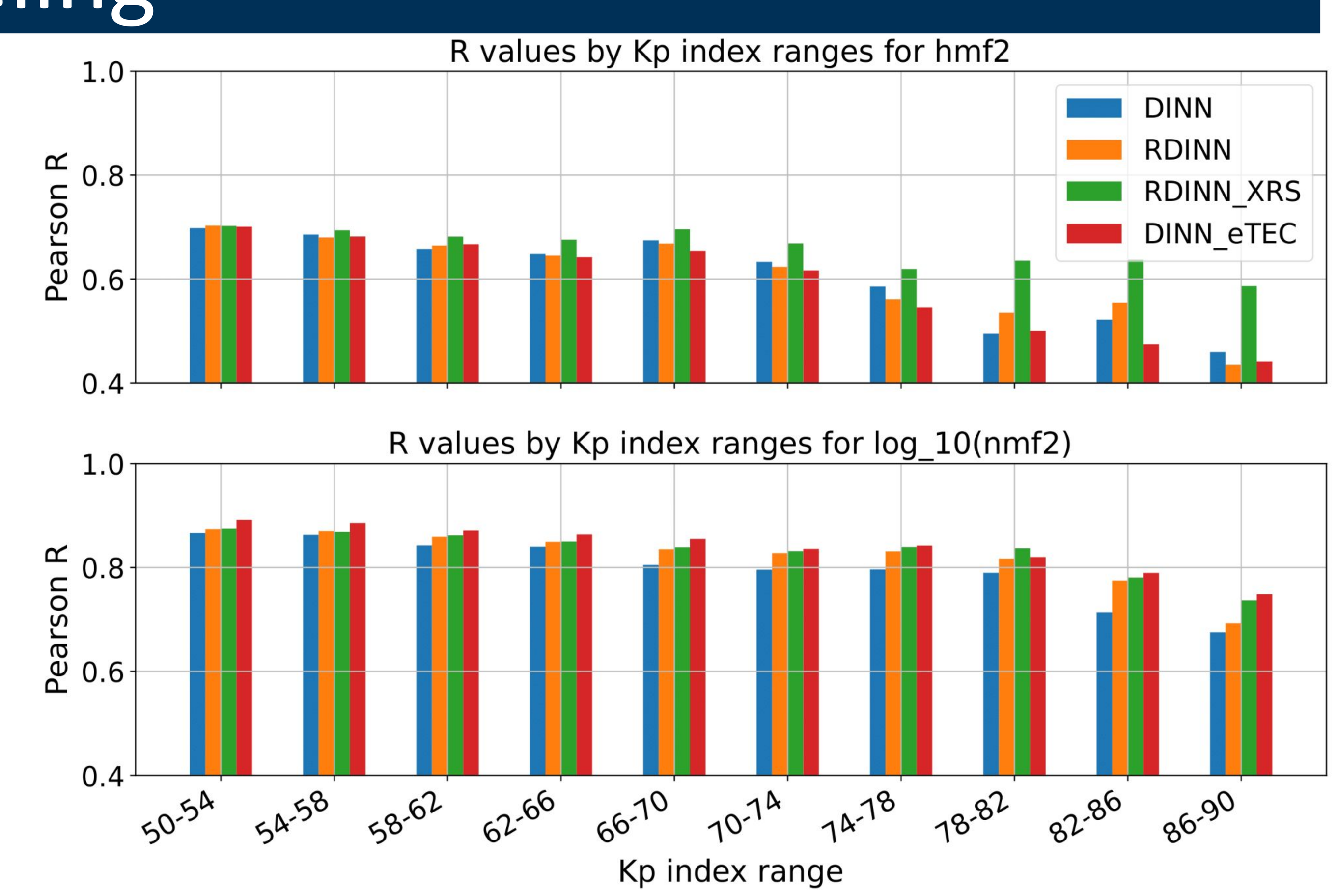
- We compare embedding models to demonstrate that data can be embedded effectively and larger embedding spaces capture more information (Pearson correlation coefficients shown to the right).
  - We have various sizes of embeddings: 16, 64, and 512 (noted as TEC_e#).
  - We also have a non-embedding comparison that directly maps input parameters (time, location, and geomagnetic indices)
- Embedding size acts as a trade-off between accuracy and computational resources.
  - Even using a small embedding space drastically improves reconstruction over a climatological approach.
  - Embedding and using only TEC data allows for better modeling than a model using geomagnetic indices instead.
- Compared to traditional CNN-based approaches, we notice worse performance at low sparsity, but at high sparsity the attention-based approach performs better.
- Correlation of the TEC_e16 vectors with commonly used electron density model inputs is low, indicating the model is capturing additional non-high-level information.

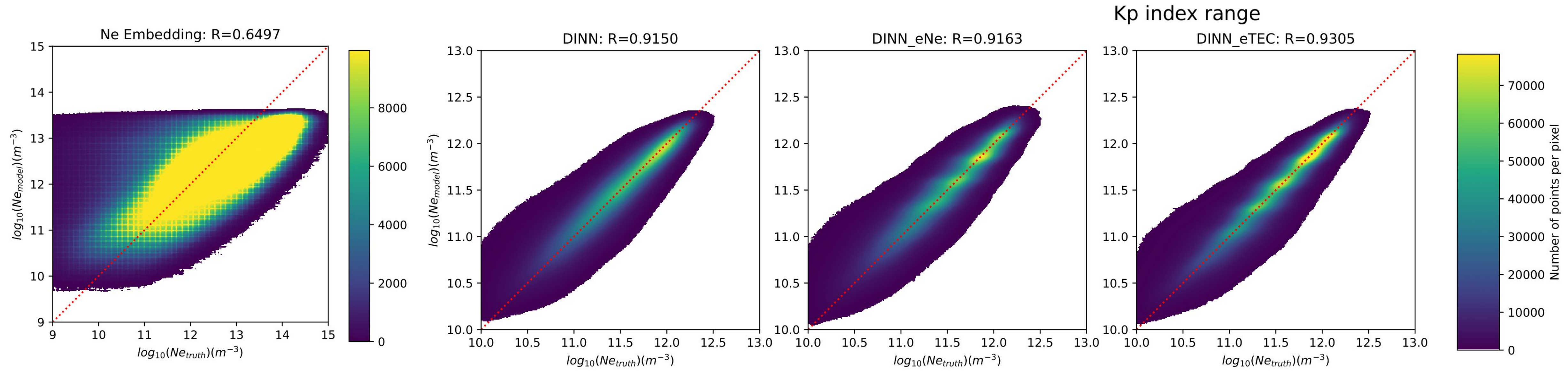




Electron Density Modeling

- We previously modeled electron density without TEC with Deeper Ionospheric Neural Network (DINN).
  - Models use location, time, and geomagnetic indices (kp, ap, dst, and f10.7) as inputs.
  - These models produce reasonable results but lack more descriptive ionospheric observations as inputs.
- We use embedded TEC information to supplement typical model inputs (DINN_eTEC).
  - We use a small embedding space of size 16 due to computational time, but plan to expand this.
  - Embeddings provide richer input data without significantly increasing model complexity.
- DINN_eTEC outperforms DINN in the general case.
- For storm-time predictions (high kp values), we have more mixed results.
  - DINN_eTEC shows better accuracy for peak electron density (nmf2), outperforming even models with historical and X-ray flux data (RDINN and RDINN_XRS, architecture not shown), likely due to a strong correlation between TEC and nmf2.
  - Peak altitude (hmf2) predictions are not as promising, likely due to the lack of vertical profile information in TEC that may be overfit.
- We perform a similar approach done with TEC data with historical electron density data.
  - We use the last complete hour of electron density data after embedding as an additional input.
  - The performance increases here are much less noticeable than with embedded TEC information.
  - Embedding electron density also ends up being worse than embedding TEC, likely because of the extremely high sparsity.





Conclusions and Future Work

- Embeddings enable improved modeling of electron density, with richer input features for ionospheric electron density models.
  - Using embedded TEC provides clear performance improvements over baseline models, particularly for nmf2.
- This work has an upcoming paper; we are currently working on:
  - Combining TEC and electron density histories together in a joint model that forecasts a few hours ahead TEC, electron density, and various geomagnetic indices.
- We also aim to:
  - Explore the possibility of using this approach to complete TEC maps (such as below, which uses the 512 sized embedding, but has not yet been validated for accuracy).