AGU Space Weather

Supporting Information for

Neural Network Models for Ionospheric Electron Density Prediction: A Neural Architecture Search Study

Yang Pan[1], Mingwu Jin[1], Shunrong Zhang[2], Simon Wing[3], and Yue Deng[1]

[1] University of Texas at Arlington, Arlington, Texas, USA.

[2] Haystack Observatory, Massachusetts Institute of Technology, Westford, Massachusetts, USA

[3] Applied Physics Laboratory, The Johns Hopkins University, Laurel, MD, USA

Contents of this file

Introduction

Figures and tables are related to the electron densities in curve and color maps, statistical errors.

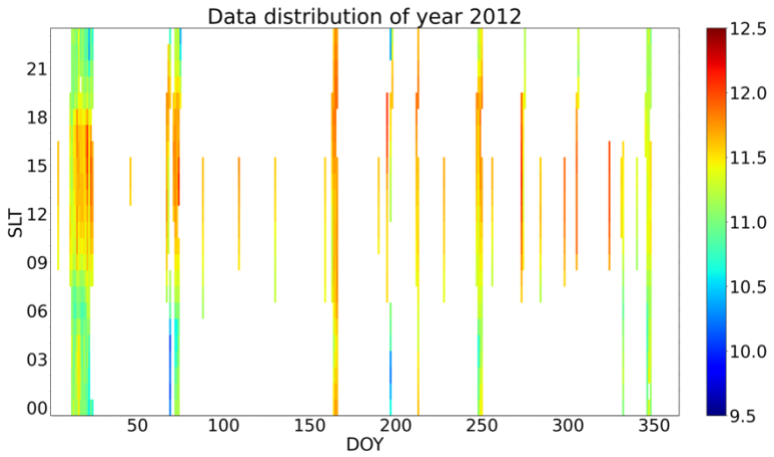


Figure S1. The ISR records of Ne in the logarithmic scale around 350 km altitude in 2012. Horizontal axis: day of year (DOY); vertical axis: solar local time (SLT); the intensity represents logarithmic electron density ($log_{10}Ne$), while the blank space represents missing records. Most of the region is in blank, indicating the irregularity of ISR's operation.

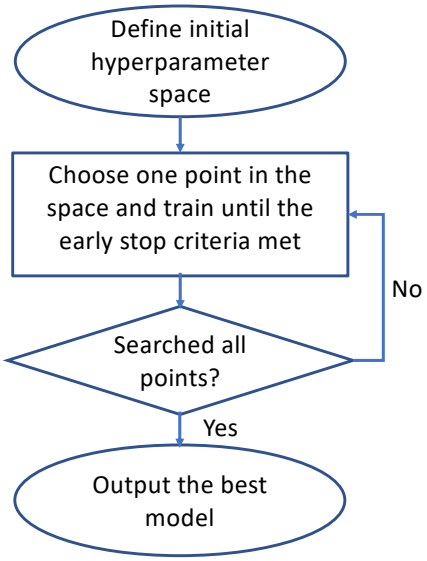


Figure S2. Flow chart of Neural Architecture Search (NAS).

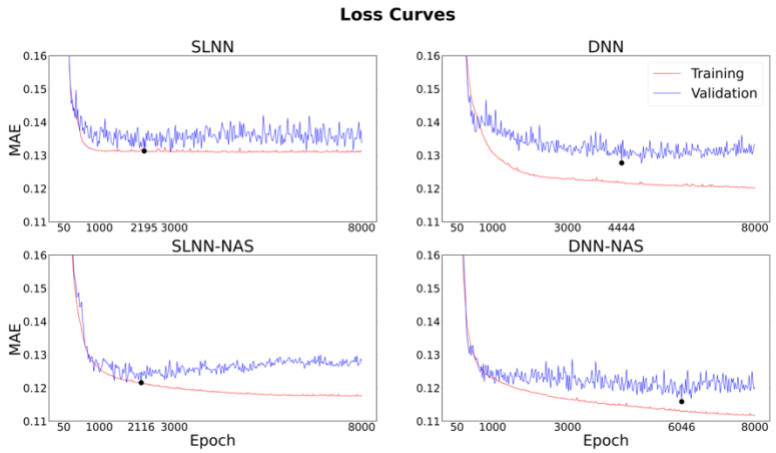


Figure S3. The training (red) and validation (blue) loss curves of four NN models (the optimal number of epochs marked as the black dot). The two DNN models take more epochs to evolve the optimal results due to more complexity than SLNNs, while the NAS guided models lead to better model generality (lower possible validation loss).

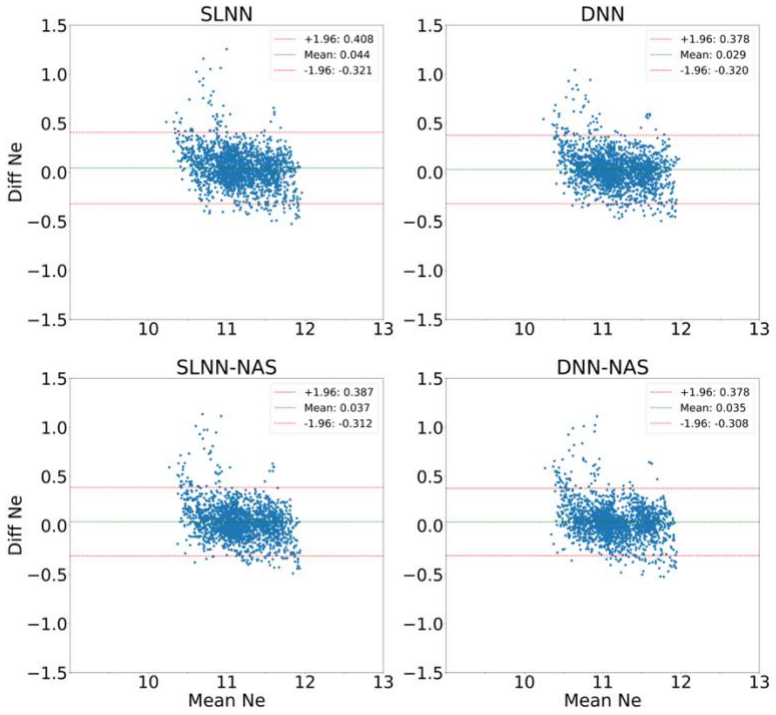
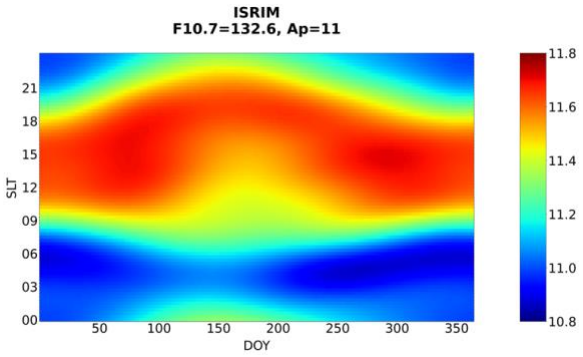
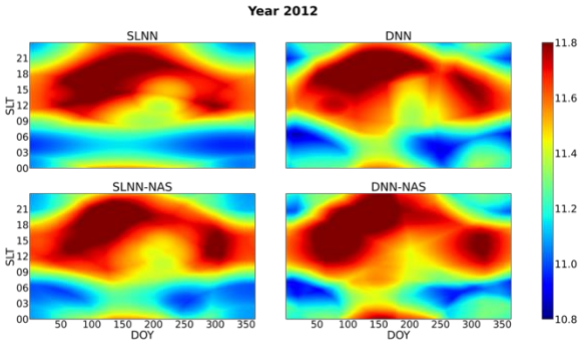


Figure S4. BA-plots of the four optimal models (SLNN, DNN, SLNN-NAS, and DNN-NAS), in which the calculations are based on the test set. DNN tends to have the lowest averaged difference (green line in the upper right subplot) and the DNN-NAS owns the narrowest limits of agreements (distance between two red lines in the lower right subplot). The Y-axis is the Ne difference between the model prediction and the observation. The X-axis is the average of the model prediction and the observation.

(a) ISRIM climatological pattern of medium solar activity.



(b) semi-annual patterns of climatological study.



(c) semi-annual patterns based on external geophysical indices.

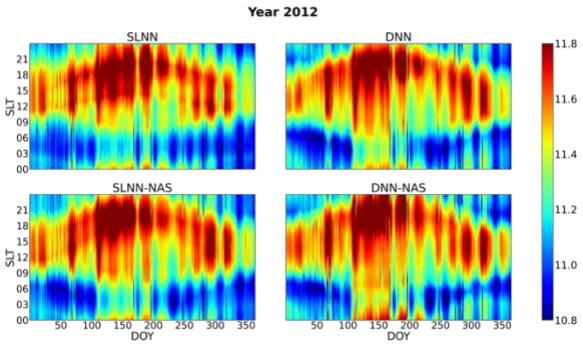
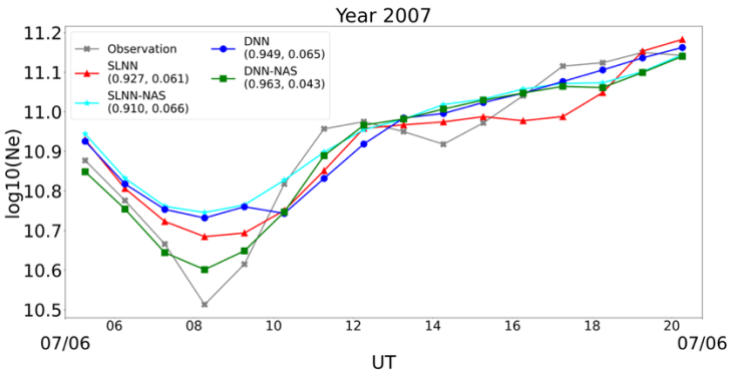
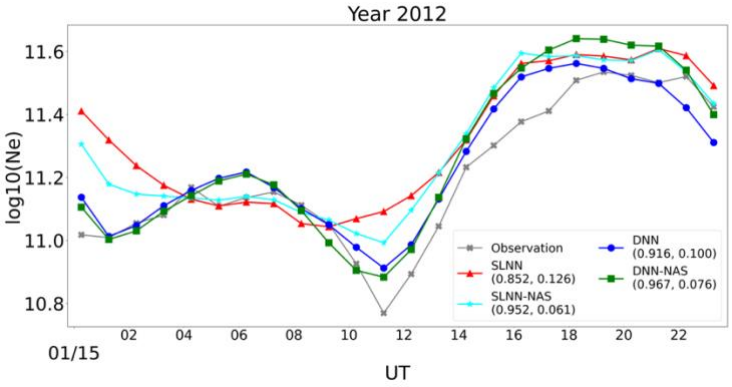


Figure S5. Annual electron density patterns of year 2012 from different sources: (a) ISR empirical model (ISRIM), (b) four model predictions based on the fixed F10.7 and Ap3, (c) four model predictions based on the real-time F10.7 and Ap3. Based on the nature of neural network models, the input can be arbitrary values. We set the evenly distributed temporal information to get the time related drivers (year, DOY, and SLT), while comparison between (a) and (b) serves as the comparison on the climatological study, while (c) demonstrates a more realistic case of Ne annual pattern with real-time F10.7 and Ap3 inputs.

(a) 2007-07-06



(b) 2012-01-15



(c) 2012-08-01

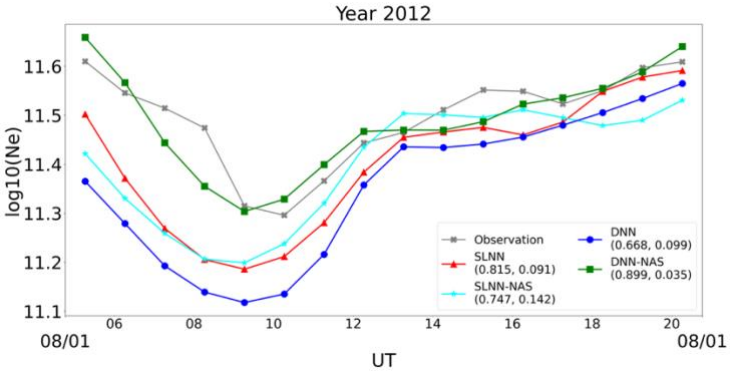


Figure S6. Daily Ne pattern prediction on three different days: (a) 2007-07-06, (b) 2012-01-15, and (c) 2012-08-01. Gray cross: the ISR observation; red triangle: SLNN; cyan star: SLNN-NAS; blue circle: DNN; green square: DNN-NAS. The two parameters (Pearson correlation coefficients and MAE) help evaluate how well model outputs predict the observed diurnal Ne pattern. Generally, all model outputs follow the observed diurnal pattern well, while DNN-NAS predicts the best.

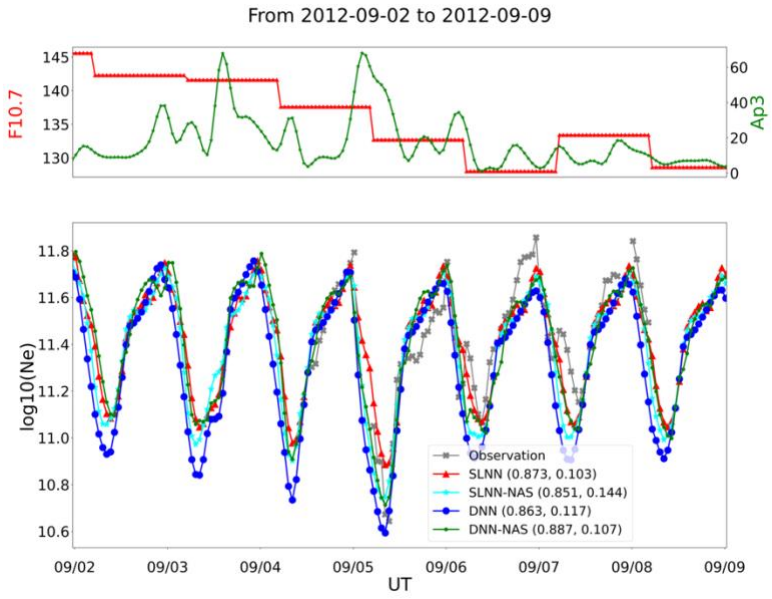


Figure S7. Ne patterns during 2012-09-02 to 2012-09-09. The two geophysical drivers are drawn in the upper panel. Four model outputs are of different markers followed with CCs and MAEs (based on observational values) in parentheses. Clearly, we see the Ap3 serves as the major driver effect to the model outputs as the predictions dip down when Ap3 reaches its peak at early time of September 5th.

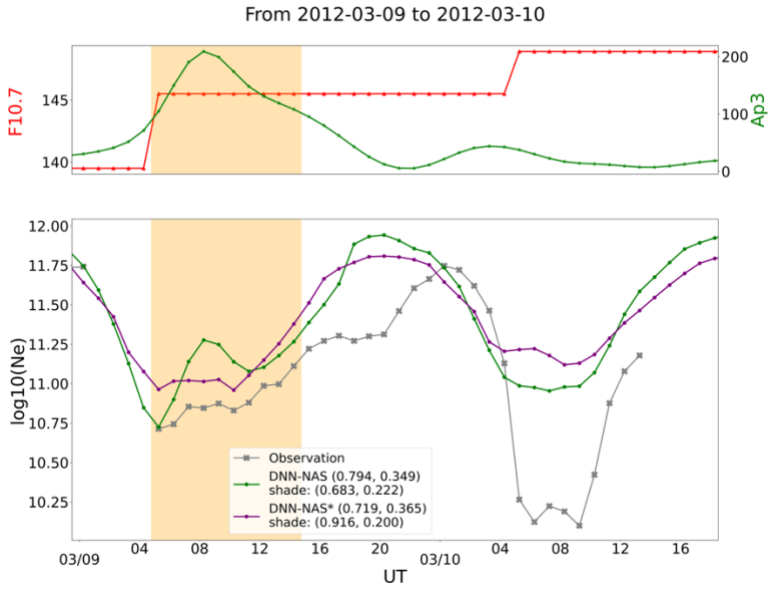


Figure S8. DNN-NAS trained with Ap3≤80 and DNN-NAS* trained without the restriction on Ap3., the DNN-NAS models trained with and without filter on Ap3 have the prediction results in green and purple color. The CC and MAE calculated on the observational data are in the parentheses (the whole curve after the model name and the shade region after "shade").

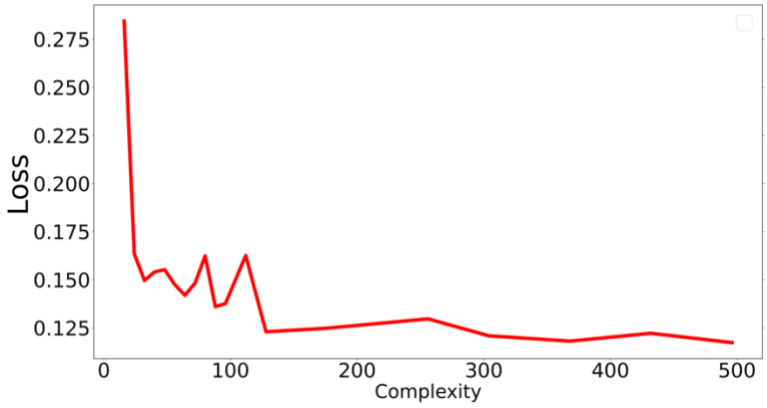
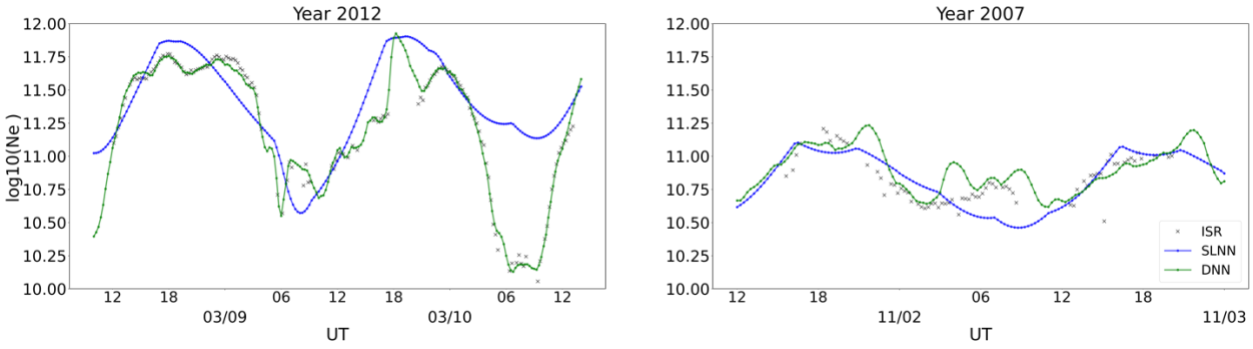


Figure S9. Prediction performance changes along with the model complexity. The complexity is defined as the total number of trainable weights of the NN model. The mean absolute error of the validation set serves as the loss function, where the less loss indicates the better performance.



(a) (b)

Figure S10. Overfitting of DNN (architecture: [512, 512, 512, 512, 32], green) (a) fitting and (b) prediction. SLNN (18 hidden neuron, blue) is served as a benchmark. DNN can fit the ISR data more closely than SLNN as shown in (a). However, DNN leads to an unrealistic wavy pattern for prediction as shown in (b).

Hyperparameter	Range
Number of layers	SLNN: [1] DNN: [2, 3, 4]
Neuron number	[16, 18, 20, …, 64]
Learning rate	SLNN: 9e-04, 8e-04, …, 1e-04
	DNN: 5e-04, 4e-04, …, 5e-05

Table S1. Hyperparameter space of AutoKeras. The candidates in each hyperparameter poll are the optimal results of multiple trials. For instance, the single layered architecture prefers a larger learning rate than the deep neural architecture.

Parameter	Values	
Years	Training	2003 to 2018 except the val&test sets
	Validation	[2010, 2015]
	Test	[2007, 2012]
F10.7	≤ 300 sfu	
Ap3	≤ 80	
Altitude	~350 km	
$\boldsymbol{Ne}$	$[\log_{10}(5 \times 10^{9}), \log_{10}(3 \times 10^{12})]$ el/m^3	

Table S2. Data setting and the conditions to clean ISR data. The ISR data has the greatest number of observations near height of 350km, which indicates the data availability is of our major consideration. The filters on two F10.7 and Ap3 would rule out high intensity geophysical events.

	SLNN	DNN	SLNN-NAS	DNN-NAS
# of layers and neurons	[18]	[24, 22, 20]	[52]	[60, 32]
Learning rate	5e-04	9e-05	1.6e-04	7.7e-05
# of epochs	2195	4444	2116	6046

Table S3. The hyperparameters for four NN models, which are the optimal results of each category in architecture, learning rate, and validation loss dip epoch.

	SLNN	*DNN*	*SLNN-NAS*	*DNN-NAS*
MAE	0.1399	0.1312	0.1307	0.1250
RMSE	0.1908	0.1805	0.1821	0.1784
RE (%)	1.2667	1.1872	1.1844	1.1327

Table S4. Prediction errors for four models in mean absolute error (MAE), root mean square error (RMSE), and relative error (RE) percentage.

		SLNN	**DNN**	**SLNN-NAS**	**DNN-NAS**
CC	Rank 1	25	16	26	61 (48%)
	Rank 2	26	35	41	26
	Rank 3	32	48	31	17
	Rank 4	45	29	30	24
MAE	Rank 1	17	30	27	54 (42%)
	Rank 2	34	32	33	29
	Rank 3	29	32	44	23
	Rank 4	48	34	24	22

Table S5. The number of ranks for daily pattern prediction. Among the 128 days in the test set, the Pearson correlation coefficients (CCs) and mean absolute errors (MAEs) are calculated and sorted from best (highest CC or lowest MAE). The DNN-NAS shows the greatest number of rank 1 cases.