https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Deep Temporal Convolutional Networks for F10.7 Radiation Flux Short-Term Forecasting

Luyao Wang[1] Hua Zhang[1,2,3] Xiao xin Zhang[2,3] Guangshuai Peng[1] Zheng Li[1]

[1]Institute of Space Weather, Nanjing University of Information Science & Technology, Nanjing, China.
[2]Key Laboratory of Space Weather, National Center for Space Weather, China Meteorological Administration, Beijing, China.
[3]Innovation Center for Feng Yun Meteorological Satellite (FYSIC), China Meteorological Administration, Beijing, China.

Correspondence to: Hua Zhang（289534957@qq.com）

Abstract. F10.7, the solar radiation flux at a wavelength of 10.7 cm (F10.7), is often used as an important parameter input in various space weather models and is also a key parameter for measuring the strength of solar activity levels. Therefore, it is valuable to study and forecast F10.7. In this paper, the temporal convolutional network (TCN) approach in deep learning is used to predict the daily value of F10.7. The F10.7 series from 1957 to 2019 are used, which the datasets from 1957 to 2008 are used for training and the datasets from 2009 to 2019 are used for testing. The results show that the TCN model of prediction F10.7 with a root mean square error (RMSE) from 5.03 to 5.44sfu and correlation coefficients (R) as high as 0.98 during solar cycle 24. The overall accuracy of the TCN forecasts is better than those of the widely used autoregressive (AR) models and the results of the US Space Weather Prediction Center (SWPC) forecasts especially for 2 and 3 days ahead. In addition ,the TCN model is slightly better than other neural network models like backward propagation network (BP) and long short-term memory network (LSTM) in terms of the solar radiation flux F10.7 forecast. The TCN model predicted F10.7 with a lower root mean square error, a higher correlation coefficient and the better overall model prediction.

1 Introduction

Solar activity has a significant impact on the Earth's climate, electromagnetic fields and communication systems, among other things. F10.7 (2800 MHz, 10.7 cm solar flux) is a good typical parameter for characterizing solar activity levels, and representing the cyclical variability of solar activity (Tapping ,2013). The F10.7 index is an important parameter in atmospheric density calculations for spacecraft orbit forecasting and in ionospheric forecasts affecting communications. For example, F10.7 is used for a control parameter in ionospheric models to calculate the variation of radio signal properties (Ortikov et al.,2003).F10.7 is also widely used for satellite, navigation, communication and terrestrial climate (Huang et al., 2009; Yaya et al., 2017). Therefore, accurate forecasting of F10.7 is not only of great value for the conduct of applicaion, but is also of comparative importance in the scientific study of space weather forecasting. (Swarup et al., 1963; Tapping and DeTracey, 1990; Henney et al., 2012).

Time-series data is data where observations of some process are recorded over the same time interval, and the F10.7 index is a typical type of time-series data. The link between F10.7 at the current moment and F10.7 at the previous moment would be decreasing as the time interval increases, so the core of the F10.7 prediction problem for time series data is to uncover the

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

potential patterns of historical data and predict the future data as far as possible (Lampropoulos et al. , 2016). The forecast models of F10.7 are mainly time series models, and the main research institutions include the SWPC, the US National Oceanic and Atmospheric Administration (NOAA), the National Space Science Center of the Chinese Academy of Sciences and the National Astronomical Observatory of China, etc. There are many researchers who have used different methods to build predictive models for F10.7. Mordvinov et al. (1986) used a multiplicative autoregressive model to forecast the monthly mean of F10.7, but the model had a large error in predicting the monthly mean F10.7.Warren et al. (2017) built optimized independent models for each forecast date, and the results showed that this approach typically predicted better than autoregressive methods. Zhong et al. (2010) utilized the singular spectrum analysis signal processing technique to predict the F10.7 index of solar activity for the next 27 days. The research result indicated that the method performed well in predicting the periodic variations of the F10.7 index. Henney et al. (2012) predicted F10.7 using the global solar magnetic field generated by the energy transport model, with a Pearson correlation coefficient of 0.97 predicted one day in advance. Liu et al. (2018)applied two models by Yeates (Yeates et al., 2007) and Worden (Worden & Harvey, 2000) to predict short-term variability in F10.7. During low levels of solar activity, the predicted values of the model were closer to the observed values.

With the rapid development of machine learning and neural networks, the powerful learning capabilities of machine learning are increasingly being intrigued by researchers and used to in the study variaiton of solar activity. The solar activity of daily values F10.7 was predicted using a support vector machine regression method by Wang et al. (2009). Xiao et al. (2017) used back propagation neural network(BP) to predict the solar activity daily mean index F10.7 for short-term forecasting.The results showed that using BP neural networks to predict the solar activity daily index F10.7 was superior to the results of Wang et al(2009). Luo et al. (2020) proposed a multi-step prediction method for the 10.7 cm radio flux. The method is a combination of the Empirical Mode Decomposition (EMD) and Back-Propagation (BP) network to construct an EMD-BP model for predicting F10.7 values. The method significantly reduces the prediction error for high levels of solar activity compared to support vector machine regression (SVR) and backward propagation neural network(BPNN)models. Zhang et al. (2020) proposed a short-term forecast of the solar activity daily mean index F10.7 by a long short-term memory network (LSTM) method. The model of the root mean square error (RMSE) of the forecast was only 6.20-6.35sfu, and the high correlation coefficient (R) was 0.98. Although the above RNN-based architecture and its variants achieved good prediction accuracy of F10.7, the training process of a model often spend a significant amount of time and computational memory, and also frequently encounters issues such as gradient explosion or vanishing gradients during network training. To this end, Bai et al. (2017) proposed a neural network called TCN, in which long input sequences can be processed as a whole in the TCN. The TCN model can read data at a faster rate and therefore has a strong capability of parallel computation. In addition, the back propagation path of TCN is different from the time direction of the sequence, which makes TCN avoid the gradient problem in RNN. In view of the above advantages and for the variability characteristics of F10.7 time-series data, this paper introduces machine learning-based TCN-related theories and techniques into the forecasting of F10.7 and compares the results of TCN prediction with other classical models to verify the effectiveness and feasibility in the short-term forecasting.

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

2 Data and Method

2.1 Data source and Data processing

F10.7 represents the solar ~~radiation~~ flux at a wavelength of 10.7 cm, and the magnitude of this index describes the intensity of solar activity. ~~F10.7 is one of the longest-running indices that records the level of solar activity. And this~~ paper collects daily average data of F10.7 from 1957 to 2019 ~~download~~ from the National Oceanic and Atmospheric Administration website Available ~~from this URL~~ (https://spaceweather.gc.ca/forecast-prevision/solar-solaire/solarflux/sx-5-en.php). ~~The F10.7 series from 1957 to 2019 are used, which the datasets~~ from 1957 to 2008 are used for training ~~and the datasets~~ from 2009 to 2019 are used for testing. Figure 1 shows the processed data. The black line represents the training dataset and the blue line is the testing dataset.

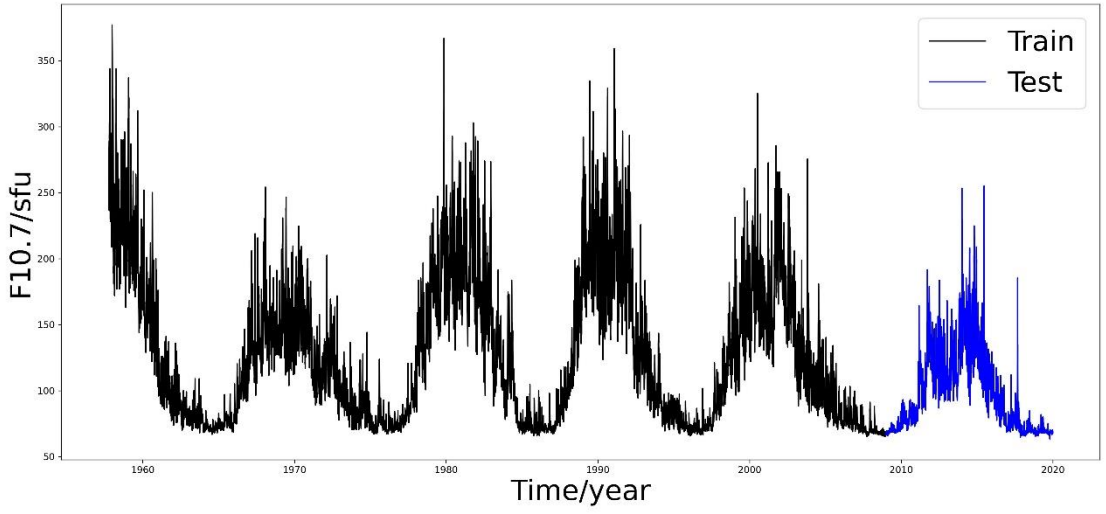


Figure 1: The Daily average value of F10.7 index from 1957 to 2019

2.2 Introduction to the experimental environment

The parameters related to the hardware and software environment for this experiment are shown in Table 1.

Table 1. Table of experimental environment parameters

Category	Configuration
Hardware Environment	CPU：Inter(R) Core(TM) i5-6200 GPU：NVIDIA GeForce 940MX

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Software Environment	Development software: Jupyter Notebook、Matlab Compiler environment: Python 3.7、Matlab Data processing frameworks: Pandas, Numpy Mapping frameworks: Matplotlib Machine learning frameworks: Tensorflow, Sklearn

2.3 Method

TCN was proposed by Bai et al. (2018). Since its introduction, TCN has caused a huge response, and some scholars believe it will replace RNN as the king of the temporal prediction field. TCN combines both RNN and CNN architectures and is a convolutional neural network variant designed to handle time series modelling problems. TCN is well adapted to the temporal nature of the data by using both causal and extended convolutional structures to extract feature information. The convolutions in TCN are causal, meaning there is no information leakage from future time steps. This distinguishes TCN from other recurrent neural networks such as LSTM, GRU, which require gate mechanisms. As a result, TCN achieves higher accuracy and longer memory without the need for gate mechanisms. Because of its a long sequence can be treated as a whole in TCN, and TCN does not have the advantages of gradient disappearance and gradient explosion problems. Here, TCN is introduced to model the prediction of F10.7.

For the prediction of a univariate time series, the TCN model takes lagged observations of the time series as inputs and predicts future F10.7 sequence values as outputs. The each set of input patterns consists of moving a fixed length window in the time series. The principle of forecasting is represented in Fig.2.

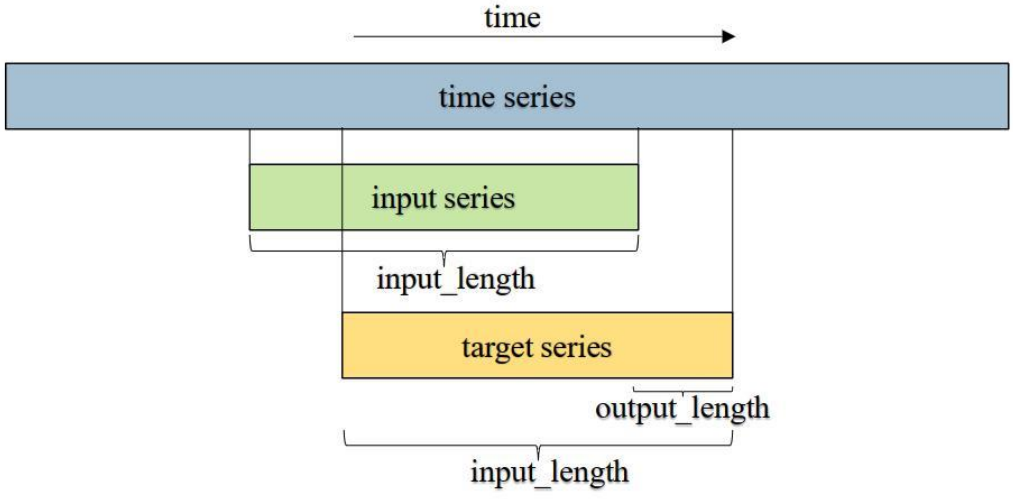


Figure 2: Diagram of F10.7 sequence data prediction

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Supposed the input of F10.7 is $x = (x_0, x_1, \ldots, x_T)$, the desired output sequence is $y = (y_0, y_1, \ldots, y_T)$, where the two sequences x, y satisfy the causal relationship. That is, the input $x_0, x_1, \ldots, x_{t-1}$ observed at the previous moment be used to predict the output y_t at moment t. The modeling objective of the TCN network is to generate any hidden function mapping, which means that the prediction of the F10.7 sequence can be represented as:

$$\hat{y}_1, \ldots, \hat{y}_{T+1} = f(x_0, x_1, \ldots, x_i, \ldots, x_T) \quad (1)$$

where x_i and $\hat{y}_i$ are the observed and predicted values of F10.7 at time i, respectively, and f is the mapping of the function trained by the TCN network.

TCN is one of the algorithms developed on the basis of convolutional neural network (CNN). That uses a one-dimensional convolutional network, consisting of an inflated causal convolution and a residual module.

One-dimensional convolution operates on time series and extracts various features, but as the length of the time series grows, a regular convolutional network requires more convolutional layers to receive longer sequences. Extended convolution, on the other hand, improves on convolution by allowing interval sampling of the input for convolution with a number of layers L and a convolution kernel of size k with an acceptance domain of:

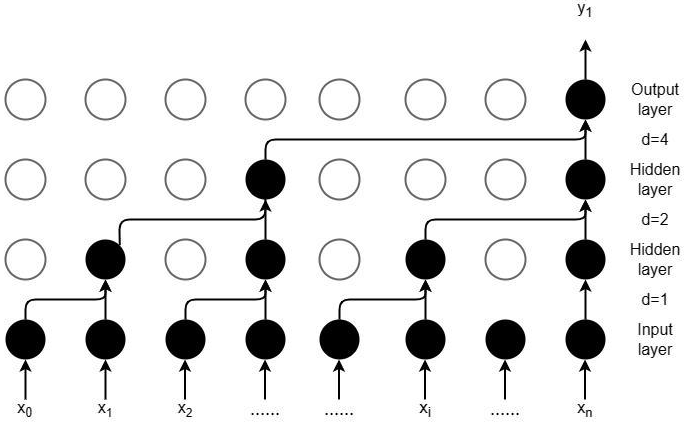
$$r = 2^{(L-1)}k \quad (2)$$

The causal extended convolution operation F for element s in a time series is defined as:

$$F(s) = (\vec{x} * f)(s) = \sum_{i=0}^{k-1} f(i) \cdot \vec{x}_{s-d \cdot i} \quad (3)$$

where: $x = (x_0, x_1, \ldots, x_T)$ is the input vector, d is the expansion factor, $*$ is the causal expansion convolution operator, f is the convolution kernel vector, k is the convolution kernel size, and $s - d \cdot i$ indicates the past direction of the input

The expanded causal convolution structure shown in Fig.3 can be adjusted by varying the number of layers, perceptual field size, convolution kernel size, and expansion coefficient. This helps to address the challenge in CNNs where the length of temporal modeling is limited by the size of the convolution kernel. Compared to traditional neural networks like LSTM and BP, TCN overcomes issues such as gradient vanishing and exploding. At the same time, TCN possesses advantages such as lower memory consumption, stable gradient, improved parallelism, and flexible perceptual field.



https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Figure 3: Expanded causal convolution

The structure of the residual module in the TCN is shown in Fig.4. The residual links allow the network to pass information across layers, thus avoiding information loss due to too many layers. Residual convolution is introduced for layer hopping and 1×1 convolution is performed to ensure that the input and output remain consistent.

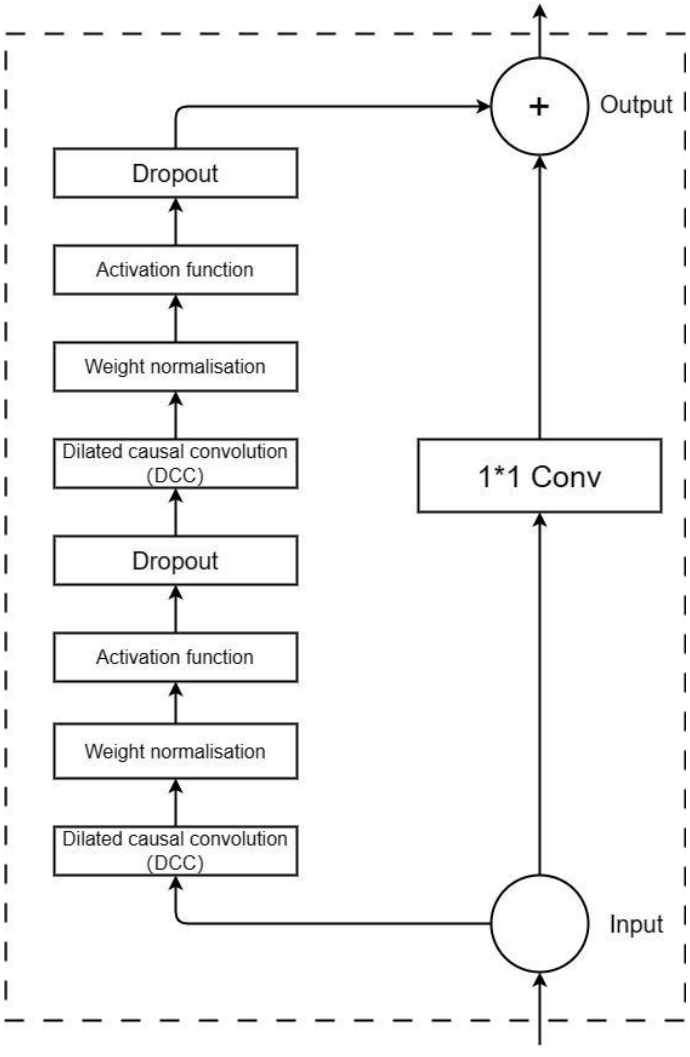


Figure 4: Structure of the TCN residual module

2.4 Selection of training parameters

A key component of the machine learning model training process is called the loss function, which gives direction to the optimization of the model by measuring the difference between the model output $\hat{y}$ and the observation y. The smaller the loss function, and the better the robustness of the model. The L1 norm loss function is extensively utilized in deep learning tasks.

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

It possesses a notable advantage of being insensitive to outliers and exceptional values, consequently avoiding the gradient explosion issue. Moreover, The loss function provides a more robust solution by offering stability. Therefore, the L1 loss function is chosen to construct the loss function for the predicted and observed values of the F10.7 sequence. The function is defined as:

$$L(\hat{y}, y) = \sum_{i=0}^{n} |\hat{y}_i - y_i| \tag{4}$$

where $\hat{y}_i$ is the predicted value of F10.7 at moment i, and y_i is the observed value of F10.7 at moment i.

In order to build the TCN model that is not merely a linear regression model, it is essential to introduce non-linearity by adding a Relu activation function at the top of the convolutional layers. To counteract the problem of gradient explosion, weights are normalized at each convolutional layer. To prevent overfitting, each convolutional layer is followed by a dropout for regularization. After several training sessions, the optimal parameters for model training are shown in Table 2:

Table 2. Training parameters of the TCN model

Parameter	Value	Parameter explaination
batch_size	None	Batch size
time_steps	20	Step length
epochs	30	Number of training session
input_dim	1	Dimension
input_shape	20	Input shape size
tcn_layer.receptive_field	/	The perceptual wildness of the convolutional layer
Dense(1)	/	Fully connected layer
optimizer	adam	Optimizer
loss	L1	Loss function
activation=	relu	Activation function
filters	64	Number of channels for the input and output of the convolution kernel
kernel_size	3	Convolution kernel size
stacks	1	Determining the depth of the network
dilations	{1,2,4,8,16,32}	Expansion coefficient
padding	causal	Fill factor

2.5. Forecast evaluation criteria

In order to quantify the forecast performance of the model, three evaluation measures, namely the mean absolute error (MAE), the root mean square error (RMSE), and the correlation coefficient (R) have been chosen, which are three common

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

model evaluation metrics to evaluate the forecast performance, referring to the evaluation methods available in the business sector such as the SWPC website.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | f_i - F_i | \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - F_i)^2} \quad (6)$$

$$R = \frac{\sum_{i=1}^{N}(f_i - \bar{f})(F_i - \bar{F})}{\sqrt{\sum_{i=1}^{N}(f_i - \bar{f})^2}\sqrt{\sum_{i=1}^{N}(F_i - \bar{F})^2}} \quad (7)$$

where MAE denotes mean absolute error, RMSE denotes root mean square error, R denotes linear correlation coefficient, N denotes number of samples, f_i denotes forecast and F_i denotes observation, $\bar{f}$ is the mean of f_i, and $\bar{F}$ is the average of F_i. Each indicator evaluates the model in a different perspective. Among them, MAE represents the average absolute error between predicted values and actual values. RMSE represents the root mean square error between predicted values and actual values. R represents the degree of trend fitting between predicted values and actual values. Therefore, the smaller the MAE and RMSE and the larger the R, the better the model prediction.

3. Results and Discussions

The TCN model is used to predict the values of F10.7 for 1-3 days ahead. Table 3 presents the statistical parameters between the predicted values and observed values of the TCN model for different years during 24 solar cycle. The table represents the performance of the TCN model in different years. In Table 3, it can be seen that the TCN model predicts F10.7 with a root mean square error (RMSE) ranging from 1 to 9 sfu for 1-day ahead, and an average absolute error (MAE) ranging from 0 to 6 sfu. The highest correlation coefficient reaches up to 0.98. For 2 and 3 days ahead, the RMSE ranges from 1 to 9 sfu, the MAE ranges from 0 to 6 sfu, and the highest correlation coefficient remains at 0.98. Irrespective of the lead time, be it one, two, or three days, the TCN model demonstrates consistent performance with relatively small ranges of root mean square error and mean absolute error, accompanied by a consistently high correlation coefficient. The results demonstrate the stability of the TCN model. However, the magnitude of prediction errors for 1-3 days ahead forecasts varies across different year. For example, the RMSE for a 1-day ahead forecast is 1.02 sfu in 2009, while its value is 8.80 sfu in 2014. Zhang et al. (2020) showed that the variation in error follows the same trend as the sunspot number, meaning that the magnitude of error is related to the year of high and low solar activity.

Table 3. The prediction errors (MAE, RMSE) and R of the TCN model for the F10.7 data during 2009–2019

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Year	1-Day ahead			2-Days ahead			3-Days ahead		
	MAE (sfu)	RMSE (sfu)	R	MAE (sfu)	RMSE (sfu)	R	MAE (sfu)	RMSE (sfu)	R
2009	0.71	1.02	0.9302	1.07	1.30	0.9313	0.73	1.03	0.9295
2010	1.55	2.15	0.9154	1.62	2.19	0.9139	1.58	2.19	0.9119
2011	3.45	5.22	0.9776	3.45	5.02	0.9785	3.41	5.14	0.9774
2012	4.47	6.59	0.9374	4.38	6.39	0.9403	4.41	6.61	0.9359
2013	3.63	4.88	0.9683	3.63	4.84	0.9690	3.63	4.96	0.9682
2014	5.77	8.80	0.9458	5.68	8.57	0.9486	5.91	8.86	0.9463
2015	4.31	8.37	0.9630	4.49	8.59	0.9037	4.45	8.80	0.8981
2016	2.17	3.03	0.9656	2.25	3.01	0.9664	2.17	2.95	0.9659
2017	2.02	5.32	0.8778	2.04	4.63	0.9067	1.92	4.48	0.9116
2018	0.88	1.15	0.9317	1.22	1.47	0.9332	0.87	1.15	0.9328
2019	0.79	1.16	0.9059	1.17	1.49	0.9092	0.81	1.19	0.9022
Total	2.77	5.44	0.9837	2.82	5.03	0.9861	2.72	5.12	0.9855

The high solar activity years of 2013- 2014, and the low solar activity year of 2018 are chosen for comparison in solar cycle 24. Figure 5 shows the predicted effects for solar activity high years in the Panel (a)-(b) and solar activity low year in the panel(c) in solar cycle 24. Observed values are represented by the black line, and predicted values are represented by the blue dots. As can be seen from Fig. 5, it shows that the TCN model effectively predicts the trend of F10.7 and exhibits good agreement in terms of magnitude between the actual and predicted values for the majority of the time. Especially during the peak of F10.7, the TCN model's predictions align well with the actual values, and it performs exceptionally well during periods of high solar activity.

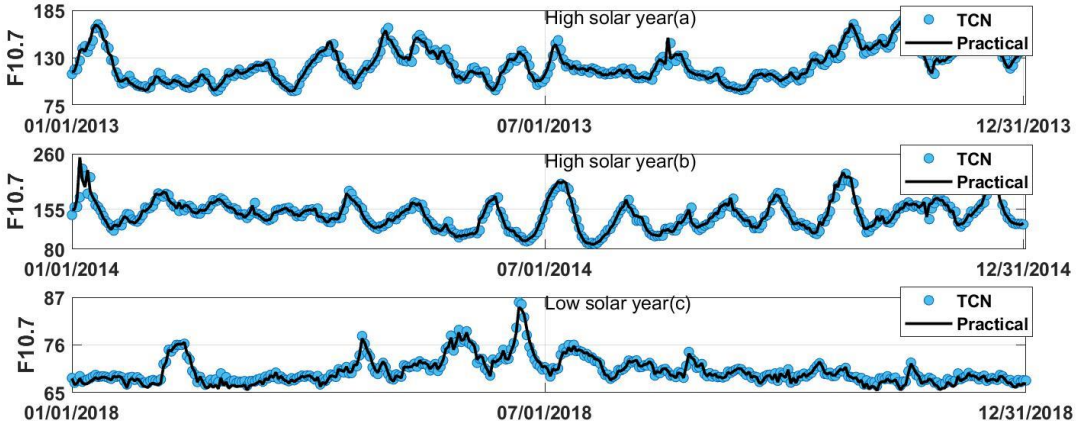


Figure 5:shows the predicted effects for solar activity high years in the Panel(a)-(b) and solar activity low year in the panel(c) for 1-day ahead in solar cycle 24.

In order to evaluate the performance of the model, we compare the forecast results of the TCN model with the SWPC forecast results and the AR model (Du et al., 2020) for1-3 days ahead. In addition, the predictions are compared with the BP neural network model (Wang et al., 2009) and LSTM (Zhang et al.,2020) for 3-days ahead.

Figure 6 shows the prediction results of the SWPC compared to the TCN model for 1-day ahead in the Panel (a), 2-days

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

ahead in the Panel (b) and 3-days ahead in the Panel(c). The blue bars represent the predicted outcome parameters for SWPC and the yellow bars represent the predicted outcome parameters for the TCN model. Figure 6 shows the TCN model's predictions are generally better than the forecasts of the SWPC. Compared F10.7 values for 1-3 days ahead, the TCN model's prediction for 1-day ahead is 0.13 sfu higher than the SWPC forecast only in 2012. While in other years, the TCN model consistently outperformed the SWPC forecast. Particularly for 2 and 3 days ahead predictions, the TCN model's performance is significantly better than the SWPC forecast. The RMSE of TCN is 5.22sfu for 1-day ahead, while the RMSE of the SWPC is 5.61 sfu in 2011. The RMSE of TCN is 0.39 sfu lower than SWPC, representing a relative decrease of 7%. For 2-days ahead prediction, the RMSE of TCN is 5.02 sfu, while the SWPC of RMSE is 9.17 sfu in 2011. The RMSE of TCN is approximately 4.15 sfu lower than SWPC, representing a relative decrease of 83%. For 3-days ahead prediction in 2011, the RMSE of TCN is 6.61 sfu, while the RMSE of the SWPC is 11.46 sfu. The RMSE of TCN is approximately 4.85 sfu lower than SWPC, representing a relative decrease of 73%.The main reason of the TCN model outperforms the SWPC forecast results, in predicting the F10.7 values for 2 and 3 days ahead, is that the TCN model effectively captures the long-term dependencies in the time series data by its structure of convolutional layers and residual connections. The structure of the TCN model could solve the non-linearities in the F10.7 sequence more effectively, to improved stability and prediction accuracy(Bai et al.2017).

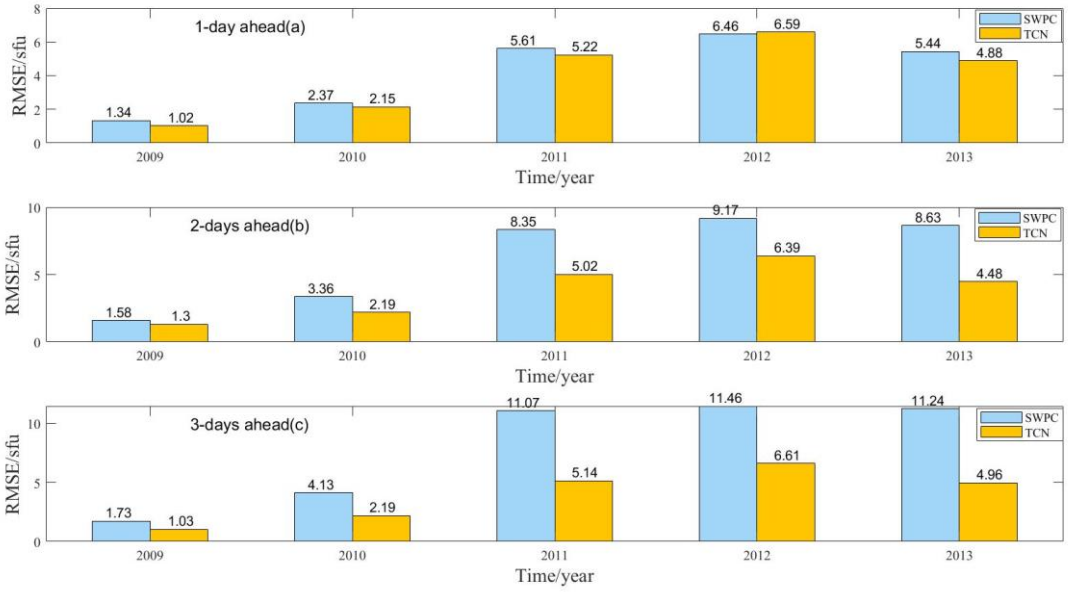


Figure 6:Comparison of the prediction performance between SWPC and TCN

Figure 7 shows the prediction results of the AR model compared to the TCN model for 1-day ahead in the Panel (a), 2-days ahead in the Panel (b) and 3-days ahead in the Panel(c). The blue bars represent the predicted outcome parameters for AR and the yellow bars represent the predicted outcome parameters for the TCN model. As can be seen in Fig.7, the TCN model outperforms the AR model overall in forecasting for 1-3 days ahead. The TCN model only has forecasts that are 0.04sfu and 0.03sfu larger than the AR model pattern for 1-day ahead in 2016 and 2019, respectively. In addition, the TCN model

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

outperforms the AR model in forecasting for both 2 and 3 days ahead. The RMSE of TCN is only 6.61 sfu for predicting outcomes for 3-days ahead in 2011, while the RMSE of AR model is 10.43 sfu. The stability and prediction accuracy of the TCN model in predicting F10.7 is again verified.

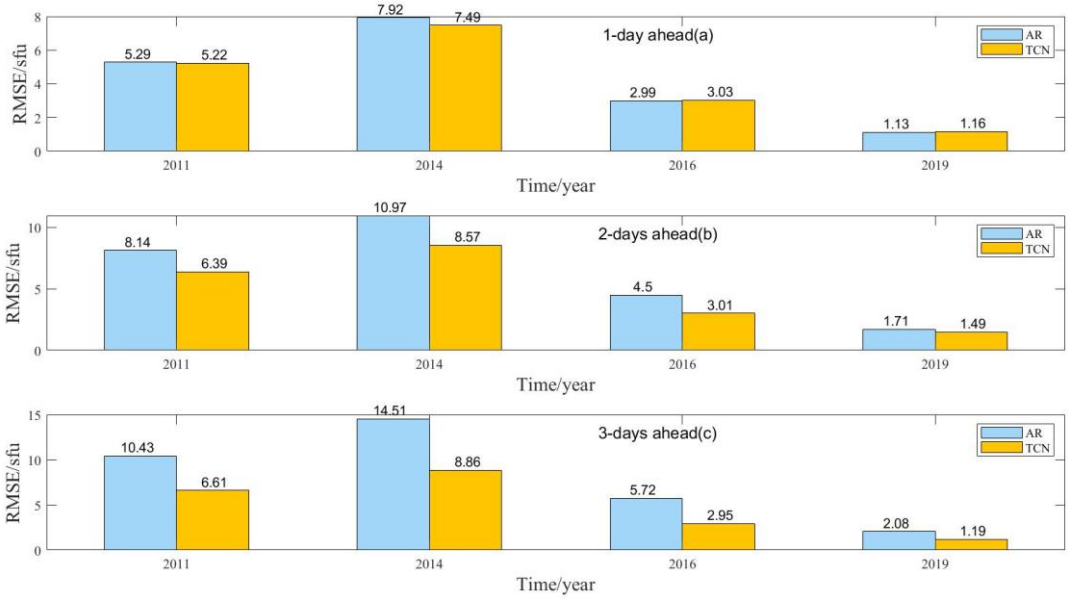


Figure 7: Comparison of the prediction performance between AR and TCN

A comparison of the TCN model with other commonly used neural network models ,like(BP neural network model (Wang et al., 2009) and LSTM model (Zhang et al., 2020) for 3-days ahead prediction is shown in Table 4. The RMSE of BP and LSTM models in predicting F10.7 in the high solar activity years of 2003-2004 are 12.28sfu and 6.09 sfu, respectively. However, the RMSE of TCN 3-days ahead is 5.73sfu in 2003-2004, which is better than those of other classical models. The TCN model predicts F10.7 better than the LSTM and BP models results during high solar activity years. There could be three reasons for such results. Firstly, the TCN model use a structure of convolutional layers and residual connections, which enables it to better capture long-term dependencies in time series data (Bai et al., 2017). In comparison, although the LSTM model can also handle long-term dependencies in sequential data, its gated unit structure may not fully capture the complex nonlinear relationships in the data (Zhang et al., 2020). On the other hand, the BP model is simpler and lacks specialized structures for handling time series data, which may result in an ineffective capture of temporal features (Wang et al., 2009). The residual connections in the TCN model can help mitigate the vanishing gradient problem and improve the stability of the model. This is particularly important for long-term prediction tasks, as the model needs to propagate gradients through multiple time steps. In contrast, the LSTM model may encounter issues of vanishing or exploding gradients in long-term prediction, leading to difficulties in training and unstable predictions (Zhang et al., 2022). The BP model, as a traditional feedforward neural network, may also face similar problems. The TCN model possesses higher flexibility and adaptability, being able to automatically learn appropriate feature representations based on the characteristics of the data. In comparison, the LSTM and BP models require

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

manual feature design and selection, which may not fully leverage the information in the data. The adaptive nature of the TCN model helps it better adapt to different time series data and improve the accuracy of predictions. Therefore, it is precisely because of the advantages mentioned above that TCN performs better in F10.7 prediction.

Table 4. Results of the TCN model's forecast performance 3-days ahead compared to other models

Models	RMSE	Year
BP	12.28	2003-2004
LSTM	6.09	2003-2004
TCN	5.73	2003-2004

4. Conclusion

The F10.7 solar flux is an important indicator of solar activity. Its applications in solar physics include serving as an indicator of solar activity level and predicting solar cycle characteristics.In view of the long observation time and certain periodicity of F10.7, this paper introduces for the first time the theory and technique related to TCN based on machine learning into the F10.7 sequence prediction of space weather.

Firstly, we analyze the ability of the TCN model to predict daily F10.7 during 24 solar cycle using training samples from 1957 to 2008.

Secondly, we compared the predictive performance of the TCN model with the SWPC forecast results and autoregressive (AR) model forecast results. The results show that the TCN model outperformed the SWPC and AR models in terms of prediction accuracy. The predictive accuracy of the TCN model do not significantly variation with the lead time of short-term forecasts (1-day, 2-days, and 3-days). This demonstrates the stability of the TCN model's predictions.

Thirdly, the TCN model has been compared to other classic models such as the BP neural network model and the LSTM model. The TCN model outperformed these models with lower root mean square error (RMSE) and higher correlation coefficient, indicating its superior predictive accuracy. This validates the effectiveness and reliability of the TCN model in predicting the F10.7 solar radio flux. The TCN model is capable of capturing sudden increases or decreases in F10.7, indicating extreme enhancements in solar activity. Therefore, the TCN model has significant implications in predicting F10.7, as it can help us better understand and forecast changes in solar activity.

Although the TCN method has proven to be a viable method for predicting F10.7, there is still room for further improvement in its predictive ability. Future work could attempt to introduce the variable of sunspot number into the model and use a more scientific approach to improve the generalization ability of the model.

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Data availability

The data of F10.7 used in this study are availabl from the National Oceanic and Atmospheric Administration at https://spaceweather.gc.ca/forecast-prevision/solar-solaire/solarflux/sx-5-en.php.

Author contributions

ZL is responsible for data acquisition, processing, data analysis, and drafting the manuscript. LYW, HZ, and GSP have made substantial and ongoing contributions to model development, interpretation, and manuscript writing. In addition to writing the article, they have also contributed to visualizing the observed results and providing explanations and discussions.

Competing interests

The authors declare no conflict of interests.

Acknowledgments

This work was supported by the National Key R&D Program of China (2021YFA0718600), the National Natural Science Foundation of China grant (No.41931073, 42074183). The author acknowledges the National Oceanic and Atmospheric Administration for providing the F10.7.

References

Bai, S.J., Kolter, J.Z., Koltun, V., et al. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[J]. ArXivPreprint, arXiv:1803.01271v2,2018.

Du, Z.: Forecasting the Daily 10.7 cm Solar Radio Flux Using an Autoregressive Model. Sol Phys 295, 125,doi: 10.1007/s11207-020-01689-x,2020.

Huang, C., Liu, D.-D., & Wang, J.-S.: Forecast daily indices of solar activity, F10.7, using support vector regression method. Research in Astronomy and Astrophysics, 9(6), 694–702, doi:10.1088/1674-4527/9/6/008,2009.

Henney, C. J., Toussaint, W. A., White, S. M., and Arge, C. N.: Forecasting F10.7 with solar magnetic flux transport modeling, Space Weather, 10, S02011, doi: 10.1029/2011SW000748,2012.

Lampropoulos, G., Mavromichalaki, H., & Tritakis, V.: Possible Estimation of the Solar Cycle Characteristic Parameters by the 10.7 cm Solar Radio Flux. Solar Physics, 291(3), 989–1002. doi:10.1007/s11207-016-0859-4,2016.

Liu, C., Zhao, X., Chen, T., & Li, H.: Predicting short-term F10.7$F_{10.7}$ with transport models. Astrophysics and Space Science, 363(12). doi:10.1007/s10509-018-3476-x ,2018.

Luo, J., Zhu, H., Jiang, Y., Yang, J., & Huang, Y.: The 10.7-cm radio flux multistep forecasting based on empirical mode decomposition and back propagation neural network. IEEJ Transactions on Electrical and Electronic Engineering, doi:10.1002/tee.23092,2020.

https://doi.org/10.5194/egusphere-2023-1801
Preprint. Discussion started: 15 August 2023

Mordvinov, A.V .: 1986, Prediction of monthly indices of solar activity F10.7 on the basis of a multiplicative autoregression model. Soln. Dannye, Byull. 12, 67.,1986.

Ortikov, M. Y., Shemelov, V. A., Shishigin, I. V., & Troitsky, B. V. : Ionospheric index of solar activity based on the data of measurements of the spacecraft signals characteristics. Journal of Atmospheric and Solar-Terrestrial Physics, 65(16-18), 1425–1430. doi:10.1016/j.jastp.2003.09.005,2003.

Swarup, G, Kakinuma, T, Covington, A E, Harvey, G A, Mullaly, R F, and Rome, J.: HIGH-RESOLUTION STUDIES OF TEN SOLR ACTIVE REGIONS AT WAVELENGTHS OF 3-21 CM. Country unknown/Code not available: N. p, doi:10.1086/147601,1963.

Tapping, K. F.: The 10.7 cm solar radio flux (F10.7), Space Weather, 11, 394– 406, doi:10.1002/swe.20064, 2013.

Tapping, K. F., & DeTracey, B.: The origin of the 10.7 cm flux. Solar Physics, 127(2), 321–332. doi:10.1007/bf00152171, 1990.

Warren, H. P., Emmert, J. T., and Crump, N. A.: Linear forecasting of the F10.7 proxy for solar activity, Space Weather, 15, 1039– 1051, doi:10.1002/2017SW001637, 2017.

Worden, J., & Harvey, J.: An Evolving Synoptic Magnetic Flux map and Implications for the Distribution of Photospheric Magnetic Flux. Solar Physics, 195(2), 247–268. doi:10.1023/a:1005272502885,2000.

Si-qing, L., Qiu-zhen, Z., Jing, W., & Xian-kang, D.: Modeling Research of the 27-day Forecast of 10.7cm Solar Radio Flux (I). Chinese Astronomy and Astrophysics, 34(3), 305–315. doi:10.1016/j.chinastron.2010.07.006,2010.

XIAO Chao, CHENG Guosheng, ZHANG Hua, RONG Zhaojin, SHEN Chao, ZHANG Bo, HU Hui.: Using Back Propagation Neural Network Method to Forecast Daily Indices of Solar Activity F10.7[J]. Chinese Journal of Space Science,37(1): 1-7, doi:10.11728/cjss2017.01.001,2017.

Yaya, Philippe & Hecker, Louis & Dudok de Wit, Thierry & Fèvre, Clémence & Bruinsma, Sean.: Solar radio proxies for improved satellite orbit prediction. Journal of Space Weather and Space Climate. 7. A35. doi:10.1051/swsc/2017032,2017.

Zhang, W.; Zhao, X.; Feng, X.; Liu, C.; Xiang, N.; Li, Z.; Lu, W.: Predicting the Daily 10.7-cm Solar Radio Flux Using the Long Short-Term Memory Method. Universe 2022, 8, 30, doi: 10.3390/universe8010030,2022.

Zhang, H., Xu, H. R., Peng, G. S., Qian, Y. D., Zhang, X. X., Yang, G. L., et al.: A prediction model of relativistic electrons at geostationary orbit using the EMD-LSTM network and geomagnetic indices. Space Weather, 20, e2022SW003126, doi: 10.1002/essoar.10511184.1, 2022.