Dear anonymous Referee2:

I am very happy to receive your recommendation and very grateful for your advice. We have followed your comments to revise this manuscript. Then, due to the stupid organization and poor English making readers understand the difficulty, we have made efforts to revise and hope that you will be satisfied. In the resubmitted paper, new text is emphasised as red text. The Referee Comments is abbreviated to "RC", and the Authors' Response is abbreviated to "AR".

The following are the responses of each critical comments:

RC1:
In my first review, I suggested that the authors split their dataset into three sub-sets: training, validation and test (instead of their simple 2-way split, that could lead to hyperparameter overfitting). In their response (AR:6) and in the revised manuscript, the authors claim they have followed the suggestion and are now using solar cycle 23 for validation and solar cycle 24 for testing (the rest of the solar cycles being used for training, see Figure 1). This implies they have re-run their pipeline, done a novel hyperparameter search using the new validation set, and trained at least a new model with this split. It appears however that their hyperparameters are still exactly the same (see Table 2), which is surprising, but not impossible. However, the numerical values of their results did not vary at all, down to the hundredth, so that Table 3 and Figures 7 and 8 are exactly the same as in the first version. This is extremely implausible given the stochastic nature of neural network training and the remarkable fact that they claim having removed an entire solar cycle from their training set. I therefore assume that the authors simply changed Figure 1 and their text in response to my comment, but did not actually do the work necessary to really take it into account. Thus, their results most likely come from a model and a pipeline that do not correspond to what they describe, which is a serious misconception.
AC1:
I followed your suggestion and divided the data into training, validation, and test sets. In this study, the training set comprises data from 1957 to 1995, the validation set includes data from 1996 to 2008, and the test set consists of data from 2009 to 2019. These three datasets are mutually exclusive. While bidirectional splitting is a common method for dividing the training and validation sets, it may lead to overfitting of hyperparameters in certain cases. Therefore, I also chose to implement the leave-one-out method for cross-validation to thoroughly evaluate the model's performance under different training-validation set combinations and select more suitable hyperparameters. Regularization techniques have been employed to prevent model overfitting. However, due to my oversight, I failed to update the test results with those from the retrained model, causing your confusion. I sincerely apologize for this mistake, and I have now made the necessary amendments in my manuscript. You can view the revised Table 2 and the specific changes in Figures 7 and 8 in my manuscript.

RC2:
In Section 3, the authors compare the performance of their TCN model to the SWPC model. They make the following statement (L240 – 244):
"The main reason 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 improve stability and prediction accuracy (Bai et al.2017)."

However, when asked to give details about the SWPC model during the first revision, the authors replied (AR16-(1)):

"SWPC is the US Weather Prediction Centre, I'm sorry I don't know what model they used to predict F10.7."

Therefore, how can the authors explain the reasons that make their model better than the SWPC model if they do not even know the nature of the SWPC model? As far as we know, the SWPC could also use RNNs or even a TCN. The authors' assertion reflects the conclusion they would like to reach without being based on real arguments, which is a serious methodological error.

AC2:

This sentence is trying to express the reason why the TCN model forecast may be better than the SWPC official website forecast. However, since English is not my native language, I have removed this sentence from the text because it is may misleading.

RC3:

As mentioned in the general comments, the document is still rather poorly written. This doesn't just mean that there are grammatical and syntactical errors, but also that the way in which the sentences are constructed and the arguments of the authors in general sometimes seem rather unclear and confused to me. I believe this should be addressed in depth before a novel submission.

AC3:

I am very sorry,I'll try to check the grammar.I also have made efforts to revise and hope that you will be satisfied.

RC4:

I find the critical discussion in section 3 and its reiteration in section 4 still weak. The authors mostly limit themselves to describing their figures and concluding that their model is very good. In my opinion, this lacks the necessary critical perspective. For example, in the first review, I asked the authors to comment on the astonishing fact that their model performs equally well with forecast horizons of 1, 2 and 3 days (after all, why stop at 3 days and not, say, 1 week?) In their reply, the authors simply repeat that their model is good for these 3 forecast horizons, unlike the reference models, without adding anything else to the critical discourse of the article. I find this disappointing and it is one of the boundaries between a simple technical report and a real scientific article.

AC4:

Short, medium and long-term forecasts of the F10.7 index are important operations for space weather agencies around the world, with forecasts for the next three days being the most important.The forecasts for the next three days are essential.(YE Qian, SONCG Qiao, XUE Bingscn,: F10.7 index forecasting method based on area statistics of solar active regions .Chin.J.Space Sci39(5):582-590.DO1:10.11728/ciss2019.05.582. ,2019.).So this paper also only examines short-term forecasts of 1-3 days, but of course I would like to work on 7-

day or even longer forecasts for the F10.7 index in the future.

Specific Comments

RC1):L.32: The authors state: "The correlation between F10.7 at the current and previous moments decreases as the time interval increases". If I am not mistaken, F10.7 is actually highly autocorrelated with a period of 27 days. This statement is therefore incorrect and should be revised. For the same reason, why did the authors use 20-day long sequences for their forecasts and not 27-day (or greater) long sequences, that would allow them to take advantage of this autocorrelation?
AR1):
F10.7 has a clear periodicity, e.g. 27 days, 11 years, etc., but the cycles are not simply repetitive but have similar but different fluctuations. The state: "The correlation between F10.7 at the current and previous moments decreases as the time interval increases". What I'm trying to convey is that over a cycle his connection is diminishing with time. I have revised this problem in my manuscript. You can see more detailed information in **lines 32-33.** Our use of 20 days instead of 27 days or more is a conclusion reached after many experiments.

RC2):Section 2.2: The text added by the authors is useful, but would be better if it was better worded, and if the authors added the references for each framework.
AC2):
I agree with you. I have reworded it. Knowledge of these frameworks is available online, and I don't think it's necessary to add references. You can see more detailed information in **lines 94-105.**

RC3): Section 2.3, L. 101: "Since its introduction, TCN has caused a huge response": this sentence is exaggerated and misleading. See e.g. Google Scholar research for the term "temporal convolutional network" (within quotes) which yields 9,030 results, while "convolutional neural network" yields 852,000 results and "recurrent neural network" yields 455,000 results. Since this sentence is in addition not useful to the paper, it should be removed.
AC3): This sentence has been removed.

RC4):Table 2: The authors state their "batch size" is None. According to their letter (AC12-(1)), they "don't set the value of batch size, [they] directly take an entire training set for training. So, the batch size is none." This is very surprising, as mini-batched training is one of the fundamental techniques in neural network training, and I cannot see any other contemporary example of a study performed without mini-batched training (see any deep learning reference textbook, e.g. "Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. Deep Learning. Illustrated edition. Cambridge, Massachusetts: The MIT Press." or "Chollet, Francois. 2021. Deep Learning with Python. Simon and Schuster"). I am not familiar with the tensorflow framework used by the authors, so I suggest they double check that setting this value to None really makes their model train with the whole train set at once, and if so, that they try to make this value vary, as it would probably further improve their results.

AC4):

The setting of "batch size is none" here does not mean that the batch size is zero, but rather that the batch size is dynamically changing and not a fixed value.

RC5):Section 2.4 and Figure 5: The authors say they use the relative error to measure the bias of their model. However, the formula they give is the formula for a normalized absolute error. This should probably be corrected, specially since in Figure 5 they use what seems to be a non-normalized error.

AC5): I have already modified this problem, and you can see the details in my manuscript.

RC6):Section 3, L.182: The authors state that "Zhang et al. (2020) showed that the variation in error follows the same trend as the sunspot number" and use this statement to explain the performance of the TCN model, which varies over the years. However, as I understand it, Zhang et al. (2020) conclusions apply only to the LSTM model developed in their own article. Thus, the authors should conclude that they come to the same conclusions as Zhang et al. (2020), rather than citing Zhang et al. (2020) to explain their results. This may simply be a writing problem (see critical comment 3), but it should not be overlooked.

AC6):

I agree with your advice. The sentence"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." is replaced by"Li et al. (2023) defined the years in which the mean value of F10.7 is greater than 110sfu as high solar activity, and the years in which the mean value is less than 110sfu as low solar activity. In this paper, the annual average of F10.7 from 2011 to 2015 is greater than 110sfu, so the years from 2011 to 2015 are called high solar activity years and the remaining years are called low solar activity years. Table 1 shows that solar activity has a periodic effect, and the prediction accuracy of the model is negatively correlated with the intensity of solar activity.The magnitude of error is related to the year of high and low solar activity." You can see more detailed information in **lines 198-203.**

RC7):Section 3, L.196 – 197 and Section 4, L.287 – 288: The authors repeat that their model "does not affect the final F10.7 forecasts due to specific properties of the data". This sentence is completely incomprehensible to me, since obviously their model, which produces the forecasts, does affect them. The authors repeat this sentence in both Section 3 and in the conclusions so it seems important. It should be rephrased so that it makes sense.

AC7):

Because the F10.7 data has a solar cycle effect. The prediction accuracy of the model is negatively correlated with the intensity of solar activity. The change in prediction accuracy of the model is related to the intensity of solar activity but has no significant relationship with the model. This proves that the TCN model is robust to some extent. I repeated the sentence. You can see more detailed information in **lines 198-203 and 295-296.**

You can see the detailed changes in the resubmitted manuscript. If you have any problems, please contact me immediately. I am very grateful for your comment. Thank you very much.

Best Regard
LuYao Wang
The 1st author of this manuscript

Dear anonymous Referee 3:

I am very happy to receive your recommendation and very grateful for your advice. We have followed your comments to revise this manuscript. Then, due to the stupid organization and poor English make readers' understand difficulty, we have made efforts to revise and hope that you could be satisfied. In the resubmitted paper, new text is emphasised as red text. The Referee Comments is abbreviated to "RC", and the Authors' Response is abbreviated to "AR".

The following are the responses to each major comment:

RC 1:
Figure 1. Caption inadequate. Need to explain the black, red, blue colors (in addition to the legend on the top right hand corner) and mention the "leave-one-out"method.
AR 1:
I agree with your advice. The sentence"Figure 1: The daily values of F10.7 index from 1957 to 2019." is replaced by"Figure 1: The daily values of F10.7 index from 1957 to 2019. Where the black line represents the training set(solar cycles 19-22), red represents the validation set(solar cycle 23), and blue represents the test set(solar cycle 24)." The sentence has been revised in **lines 92-93**. The use of the leave-one-out method in this paper is a way for us to consider whether choosing different datasets could lead to large differences in model performance. We leave iteratively one solar cycle out as a test dataset and rerun the model each time (e.g. keep solar cycle 23 as test dataset and train the model with the remaining solar cycles, then keep solar cycle 22 as test dataset and train the model with the remaining solar cycles, etc.).The specific application of the leave-one-out method in this paper as well as the results in **lines 205-215.**

RC 2:
No need of Table 1 and I recommend removing it. Describe the software and hard-ware details in words.These are all standard software and hardware information in the context of machine learning and therefore, no need to highlight in the form of a table.
AR 2:
I agree with your advice. Table 1 is removed and describes the software and hardware details in words. You can see more detailed information in **lines 94-105.**

RC3:
Figure 2: Explain the diagram.
AC3:
In time series forecasting tasks, the goal is to predict future sequences based on known sequences. The original dataset is lengthy, and during training, a continuous subsequence needs to be inputted. The output and input lengths of the temporal convolutional network (TCN) are equal, meaning the length of the output sequence generated by the TCN is equal to the specified input length. To meet the prediction requirements, the specific number of steps to forecast (referred to as output length) should not exceed input length, allowing for partial overlap between the input and output sequences. Figure 2 The blue part represents the original sequence,

the green part represents the input subsequence, and the orange part represents the overlap and the actual predicted lengths. You can see more detailed information in **lines 118-125.**

RC4:

Figures 3 and 4. These two figures are the same as the two panels in the figure in Bai et al., 2018.

I strongly recommend removing the two figures from the manuscript as they are not newly developed for the current work/manuscript. After removing the figures, you could simply refer to the Bai et al., 2018 figure in the text. Explain the difference, if any, in the architecture you have chosen for the current work/manuscript.

AC4:

To let readers understand more about the TCN network, Figure 3 and Figure 4 are new diagrams drawn in this article, Figure 3 and Figure 4 are titles in parentheses just want to express that if you want to know more about the content of the belief can refer to bai's article, it may be that I expressed the error, I have made changes in the article. You can see more detailed information in **lines 141-147 and 154-155.**

RC5:

Figure 5: Panel numbering is not clear.Suggestions: "Panel (a)", "Panel (b)" and "Panel (c)" or simply, "(a)", "(b)" and "(c)" at the top right hand corner.

AC5:

I agree with your advice. I have revised my manuscript. You can see more detailed information in **Figure 5.**

RC6:

Figure 6. If you extend the Y-axis slightly, the plot will have a cleaner and clearer look. I strongly recommend this.

AC6:

 I agree with your advice. I have revised my manuscript. You can see more detailed information in **Figure 6.**

RC7:

Make similar changes as the one on the line above for Figures 7 and 8. For figure 7, you may want to choose the top left corner for giving the panel names. For Figure 8, you need to extend the Y-axis because the numbers over the bars are falling outside of the top axis. Then you may use the top left corner for placing the panel names.

AC7:

I agree with your advice. I have revised my manuscript. You can see more detailed information **in Figures 7 and 8.**

SOME SPECIFIC CORRECTIONS SUGGESTED

RC1): "... RNN (...) ...." Change this to, "... recurrent neural network (RNN) ... "
AC1): Line 56: "... RNN (..) ...." has been changed to "... recurrent neural network (RNN) ."

RC2): Line 59: Expand "TCN".I see that this is expanded in the abstract but I suggest giving the full form with the acronym in parenthesis where this is first introduced in the main text as well. Make this change to other acronyms not explained in the main text, if there are any.Please check.
AC2):
The"TCN" has already been mentioned in line 10, and where a proper noun is mentioned before, it will be used as an abbreviation afterwards. But it doesn't appear in the text. So I've added TCN's explanation here as well. You can see more detailed information in **line 59**.

RC3): Line 62: This is a repetition - use only RNN because it is already defined on Line 56.
AC3): The phrase "recurrent neural network" was removed.

RC4): Line 101: Remove the second sentence.
AC4): The second sentence was removed.

RC5): Line 155: Use "ReLU", which I guess is the standard.
AC5): Line 171:The word "Relu" is replaced by "ReLU"

RC6): Line 191: Change "rest" to "remaining".
AC6): Lines 207-208:The word "rest" is replaced by "remaining"

RC7): Line 192: See the comment above for Line 191 - apply the same.
AC7): The same as RC 6)

RC8): Line 193: Replace "relatively poor" by "larger".
Line 193: Remove "relatively" before the word "better".
(You do not need to use "relatively" if you are using comparatives such as "better".
A similar usage is seen in many other places - you may want to correct them as well.)
AC8):
Line 209:The phrase "relatively poor" is replaced by "larger" and the word "relatively" is removed.

RC9): Line 194: See the comment above for Line 193.
AC9): The same as RC 8)

RC10): Line 196: Same as before - remove "relatively" before "poorer".
AC10): The same as RC 8)

RC11): Line 205: delete "we see that".

AC11): "we see that"was deleted.

RC12): Lines 210-214: The figure caption contains too many repetitive words/phrases. This can berewritten in a simpler manner.For example, change the sentence to, "Figure 5: shows the frequency distribution of the difference between the observed values and the model predictions during 1996-2019(solar cycles 23 and 24) for 1-day ahead (Panel (a)), 2-day ahead (Panel (b)), and 3-day ahead (Panel (c)) observations and the corresponding model predictions."

AC12):
Lines 224-225:The sentence "Figure 5: shows the frequency distribution of the difference between the observed values and the model predictions during 1996-2019 (solar cycles 23 and 24). Panel (a) displays the frequency distribution of the difference between the 1-day ahead observations and the model predictions. Panel (b) displays the frequency distribution of the difference between the 2-days ahead observations and the model predictions. Panel (c) displays the frequency distribution of the difference between the 3-days ahead observations and the model predictions" is replaced by "Figure 5: shows the frequency distribution of the difference between the observed values and the model predictions during 2009-2019 (solar cycle 24) for 1-day ahead (Panel (a)), 2-days ahead (Panel (b)), and 3-days ahead (Panel (c))."

RC13): Line 211: insert space after "23" and "and". Also, between "9" and "(".
AC13): Line 224: I have revised in my manuscript.

RC14): Line 217: It must be Figure 6, please check.
AC14): Line 228:The "Figure 5" is replaced by "Figure 6".

RC15): Line 219: Again, this must be Figure 6.
AC15): Line 230:The same as RC 14)

RC16): Lines 224-225: Caption of Figure 6 is inadequate. Explain all the symbols and lines briefly and if more details are needed, write within brackets, e.g. "(see text for further details)" and provide the necessary details in the section where the figure is discussed.For example, though the panels show the legends for the blue dots and the lines, explain them briefly in the caption as well.

AC16):
Lines 235-237:The sentence "Figure 6: shows the predicted effects for solar activity high years in the Panel(a)-(b) and solar activity low years in the panel(c) for 1-day ahead in solar cycle 24." is replaced by "Figure 6: shows the predicted effects for solar activity high years in the Panel(a)-(b) and solar activity low years in the panel(c) for 1-day ahead in solar cycle 24. The black line represents observed values, while the blue dots represent predicted values."

RC17): Line 227: Check the spacing after the parenthesis.
AC17): Line 239:The spacing has been revised.

RC18): Line 232: insert "that" after "shows".
AC18): Line 244: The word "that" was inserted

RC19): Line 233: insert "with" after "compared".
AC19): Line 245: The word "with" was inserted

RC20): Line 292: Change "variation" to "vary".
AC20): Line 300: The word "variation" is replaced by "vary"

RC21): Lines 244-246: See comment above for Lines 210-214. I suggest simplifying the caption. For example, "Figure 7: Comparison of the prediction performance of SWPC and TCN for, Panel(a) 1-day ahead, Panel (b) for 2-day ahead, and Panel (c) for 3-day ahead . . . ."
AC21):
Lines 256-257:The sentence "Figure 7: Comparison of the prediction performance of SWPC and TCN. Panel (a) is a comparison of the prediction performance of SWPC and TCN 1-day ahead. Panel (b) shows the performance comparison between SWPC and TCN 2-days ahead. Panel (c) shows the performance comparison between SWPC and TCN 3-days ahead." is replaced by "Figure 7: Comparison of the prediction performance of SWPC and TCN for Panel(a) for 1-day ahead, Panel (b) for 2-days ahead, and Panel (c) for 3-days ahead during different years."

You can see the detailed changes in the resubmitted manuscript. If you have any problems, please contact me immediately. I am very grateful for your comment. Thank you very much.

Best Regard
LuYao Wang
The 1st author of this manuscript