This work presents a very interesting analysis of Ap index and Dst, which are a measure of geomagnetic activity or geomagnetic storms, in connection to solar activity indices and solar wind parameters. The Nonmetric Multidimensional Scaling (NMDS) statistical method is applied to the whole set of parameters, and then a machine learning method, the Long Short-Term Memory (LSTM) Networks, is applied to forecast Ap and Dst annual mean values along solar cycle 25 (from 2019 and ~2029).

I consider this work acceptable for publication, though it need some clarifications based on the comments I outline below. I also detail some minor errors at the end.

Major comments:

(1) I think that it is important to mention the time scale of the data series analyzed in the introduction, or somewhere at the beginning of the work, since I think that it is not usual to analyze prediction methods in interannual time scales for geomagnetic activity indices, since the importance of their forecast for space weather purposes is in general in much shorter timescales, as hourly or daily, which the timesacle of geomagnetic storms and solar disturbances.

Answer: We added the explain of the timescale.

(2) Line 19: I don't agree in that "heightened solar activity", which I think the authors refer to high solar activity level, is synonym of "solar storm". You can have high solar activity levels with no solar storms, and also solar storms during low solar activity levels.

Answer: We modified this statement.

(3) Line 76: Where you mention "The horizontal axis (Dimension 1) appears to represent the overall level of solar activity, with higher values corresponding to increased activity. The vertical axis (Dimension 2) seemingly captures the nature of solar wind disturbances, with positive values associated with enhanced plasma flow speed and proton temperature, and negative values linked to reduced solar wind pressure and geomagnetic storm intensity."

Why Dimension 1 "appears to represent"? Is this not for sure? And as I understand, the interpretation of negative values that you give would not agree with Dst located in negative values, since the more negative Dst is, it indicates a stronger storm. Maybe I am missing something here.

Answer: We considered that the NMDS method will have different results and interpretation angles when selecting different scales of the data, so we do not use very positive statements in the wording

(4) Line 80: What is mentioned in all this paragraph is something that can be also deduced from the correlation between Dst and Ap with with each of the indices that you analyze. That is, the direct association between Ap and all solar indices and solar wind parameters, and an inverse correlation with Dst but lower than the values with Ap. I list the values of the squared correlation coefficient in the following Table, based on annual mean values of each parameter:

	Ap	**Dst**
B	0.79	0.73
T	0.82	0.53
V	0.54	0.30

A/P	0.63	0.60
Pressure	0.75	0.61
Rz	0.33	0.40
F10.7	0.32	0.41
Sunspot Area	0.40	0.47
Dst	**0.77**	**1.00**
Ap	**1.00**	**0.77**

Answer: You are correct that correlation analysis reveals the strength of the relationship between Ap and Dst with other solar parameters. However, Non-metric Multidimensional Scaling (NMDS) aims to arrange objects (in this case, different solar wind parameters and geomagnetic indices) in a low-dimensional space (typically two or three dimensions) so that the distances between the objects reflect their dissimilarities in the original high-dimensional space as closely as possible. The key here is "dissimilarity." Unlike Principal Component Analysis (PCA), NMDS does not directly use the original data; instead, it uses a dissimilarity or distance matrix as input. Furthermore, it does not assume linear relationships within the data. Therefore, it does not directly provide specific correlation coefficients or linear relationship information.

(5) In Figure 1, Dst appears far from Ap, even though they have 77% of common variance, as can be notices in the Table above. The only difference is that they vary in counterphase. Why are they so apart? Maybe I do not understand the methods correctly.

Answer: As stated in the previous response, the goal of NMDS is to preserve the similarity relationships between samples in multidimensional data through dimensionality reduction, rather than directly providing specific correlation coefficients or linear relationship information. The distance between two points only represents the degree of similarity in their distances after dimensionality reduction. It does not directly indicate that they are linearly correlated, nor can it directly determine the positive or negative correlation. The proximity of points in the NMDS plot reflects overall similarity, not necessarily linear correlation. A large distance suggests dissimilarity, while a small distance suggests similarity, but the nature of that relationship (linear, non-linear, positive, negative) cannot be inferred from the NMDS plot alone. Further analysis would be required to determine the specific nature of the relationships between variables.

(6) Line 42: I would start a new paragraph with the sentence, "Analysis of Ap index forecasting performance ..." And what is the source of the data in this phrase? I mean, the correlation coefficient's' values mentioned. I understand they come from Paouris et al. (2021) work, but I think that here they lack context. Maybe you can add more explanation.

Answer: We have modified it.

(7) Lines 134-136: The normalization within each solar cycle, is it necessary? Have you repeated your analysis with a normalization of the whole period, or without normalization at all?

In doing this you are loosing an important aspect of geomagnetic activity that is its intensity. I see however that Ap along solar cycle 24 is significantly lower than in previous cycles (as is noticed in the figure I attached), and Dst is closer to zero along this cycle, accordingly.

I tried to reproduce your data and plot them with the Ap and Dst data prior to normalization, and they really look different. I think that you should discuss more on this procedure that you applied to the data. Or at least the consequences they have. One of them is that the value you obtain of the peak has nothing to do with the true value expected at the peak you detect. This should be highlighted, unless I am not understanding the analysis correctly. I am attaching the figure I made in the case of Ap where you can see clearly what I am mentioning here.

Answer: As you pointed out, the 27-day averaged Ap and Dst data exhibit significant fluctuations. This study focuses on the long-term variations of Ap and Dst over an entire solar cycle. Therefore, we applied Gaussian smoothing to the data. Initially, we attempted to perform prediction calculations directly on the smoothed data, but the results were unsatisfactory. After analysis and discussion, we decided to normalize the data within each solar cycle. This approach ensures a more consistent range of variation across all cycles without affecting the timing of occurrences of specific values such as maxima and minima. We acknowledge that this pre-processing method may alter the true maximum and minimum values. However, considering that our primary objective is to predict the timing of these occurrences, we believe that this pre-processing method is more suitable for our research purpose. The focus on predicting the timing, rather than the exact amplitude, of events justifies the use of normalized data.

(8) Line 168: In the sentence "Furthermore, the LSTM+ model successfully identifies the timing of the AP maximum and the DST minimum, both projected to occur in September 2026." How do you know that the prediction is successful when September 2026 has not occurred yet?

And, can you explain further the following sentence "This prediction aligns with historical observations across multiple solar cycles". Or the explanation is the following paragraph? If this is the case, I think that the sentence should go then in the same paragraph.

Answer: We apologize for the misunderstanding. We have removed the word "successfully" from the text. We intend to convey that our prediction indicates that the maximum of Ap and the minimum of Dst are both likely to occur around the date obtained from the forecasting results. Furthermore, we will include additional details and supporting evidence in the revised manuscript to clarify this prediction. We will elaborate on the methodology and analysis that led to this prediction in the revised manuscript.

(9) In Figure 3, I notice that the gray lines (that is the observed values) have a departure from the predicted values larger than that observed in Figure 2 (for Ap and for Dst) along SC 24. Of course that years 2019-2023 that are seen in Figure 3 are not included in Figure 2 (the two panels), but I expected a better agreement for the period with data as in your previous cases. What happened here? Please check.

Answer: In response to the reviewer's comments, we have recalculated the prediction results for

Cycle 25.

(10) I guess that in the LSTM+ method you included all the solar and solar wind indeces. What is the purpose then of the NMDS analysis? Again maybe I am not understanding well all the methodology.

在 LSTM+预测中，我们是用 Ap 的 21 至 23 周期的值预测了 24 周期的值，和真实值进行了比较，选择了预测精度最高的 LSTM+参数，然后对 Ap 在 25 周期的值进行了预测计算。对 Dst 的计算也是相同的方法。在 NMDS 中我们是用了太阳和太阳风参数以及地磁参数进行的计算，目的是想通过这个方法分析地磁参数与太阳和太阳风参数之间的相似性，阐述地磁参数受到太阳和太阳风参数的影响性。

(11) When you mention " However, the development of geomagnetic storms, as indicated by DST, appeared to be influenced by a confluence of factors rather than being solely attributable to individual solar parameters."

Don't NMSD consider all the parameters combined? If not, a simple regression analysis would have served to conclude this. In fact, the correlation coefficient of a linear regression of Ap and all the parameters you consider (except SFI since I was not able to find the concatenated series for the period) is 0.96. And in the case of Dst it is 0.83. That is 96% of the variance of Ap is explained by the set of the 8 out of the 9 solar activity and solar wind parameters you considered, and 83% of Dst variance.

How the NMSD method improves the understanding of a high correlation in this case?

正如我们上面解释的内容。**NMDS 的目标是通过降维将多维数据中样本之间的相似性关系尽量保留，而不是直接提供具体的相关系数或线性关系信息。关键在于“非相似性”，它不像主成分分析 (PCA) 那样直接使用原始数据，而是使用对象之间的差异或距离矩阵作为输入。它也不假设数据呈线性关系。不是直接提供具体的相关系数或线性关系信息。**

(12) In the conclusion, I do not really see how "These findings contribute significantly to our understanding of the intricate relationship between solar activity and geomagnetic fluctuations.". I see that they confirm an intricate relationship, but I do not see the understanding that this analysis adds.

本文是通过 NMDS 方法展示了太阳活动和地磁波动之间的相似性结果，目的并不是深入的讨论他们之间相互影响的因果关系。我们的用词不太恰当，因此我们将这句话改成 These findings contribute a new sight for our understanding of the intricate relationship between solar activity and geomagnetic fluctuations.

Minor comments:

Line 29: I think that instead of "diminishment", "decrease" is better.

We changed "diminishment" to "decrease"

Line 34: In "Zhang et al." the year is missing.

We modifited the content

Line 37: " stormss" should be " storms"

We modifited the content

Line 37: In "Nilam and Ram" the year is missing.

We modifited the content

Line 39: "(Nilam and Ram, 2022)" could be deleted here, since the sentence deals with their resutls, mentioned at the beginning.

We modifited the content

Line 39: In "Abduallah et al." the year of this reference us missing.

We modifited the content

Line 80: I think that " DST" should be "Dst". Check this in all the manuscript and check please if it is more correct to use it as "Dst" instead of "DST".

Yes, it should be use as Dst. We modified it in the manuscript.