Dear Editors and Reviewers,

Thank you for your comments and the opportunity to revise our manuscript. This document contains our point-by-point responses (in blue) to your comments (in black). We believe the manuscript has improved significantly and we look forward to additional feedback. The major changes to the manuscript are as follows:

*Major Changes to Manuscript:*

1. To target a wider audience, we have re-positioned the paper as proposing k-nTS+: a machine learning-based feature selection method paired with the k-nTS swapping mechanism for time series value replacement. We illustrate the methodology with an application to replace time series values for data privacy reasons. The machine learning feature selection method is flexible and can be used to select features for predicting any categorical or continuous outcome.
2. To increase the generalizability of our results, we have included all of the M3 competition data in the empirical application. We also applied our methodology to a subset of the M5 competition data which is representative of a real-world hierarchical forecasting scenario. Our new results show that…

We have uploaded all code to Github for reproducibility and posted the revised draft on a pre-print server.

In addition, we corrected the following mistakes in the code:

* An indexing error which caused only *k*-1 nearest neighbor time series to be used for swapping instead of *k* nearest neighbor time series.

**Associate Editor Comments:**      
  
The manuscript concentrates on a timely and relevant topic. The approach is original and the point is well supported. The reviewers and I see that the manuscript potentially brings some interesting novel ideas that could be worth of publication. However, at this stage, there seems to be quite a lot of work to be done for the paper to get there.

Thank you. We appreciate your interest in our work. As suggested, we have made significant changes to make the paper worthy of publication.

Besides the point of the reviewers, I would like to insist on the fact the presentation of the work needs a serious upgrade. For instance, A figure like Figure 1 cannot end up in the final version of the paper. I would encourage the authors to be careful in the way to design and produce that figure. Similarly, in general, the paper looks like a draft, which makes it a bit tricky at stages (e.g., when I checked the validity of equations...). Using a more "profession" text editor (possibly Latex) could help a lot in improving the readability and make the work easier for the reviewers and I. Finally, for the maths, it would be good to be really thorough, e.g., one does not need to use "\*" for multiplication, the flow of information around eq. (2) is quite confusing (possibly introduce relevant notation first, for diag, 1, among others), etc.. Similarly, algorithms could be better presented.

Thank you for the feedback. We have re-written the entire draft in LaTeX and overhauled the presentation of our figures and algorithms. We have thoroughly reviewed the math and writing to improve readability.

**Reviewer 1 Comments:**

The authors propose a matrix-based privacy method called k-nearest time series + (k-nTS+) swapping that preserves time series features to maintain forecast accuracy. The proposed privacy method has been applied to a forecasting competition data set and proven its advantages through a series of empirical studies. Overall, the paper is well-structured and written, while its contribution is clearly explained and justified.

Below you may find some comments that could help further improve the current work.  
  
1. Page 5: When first introducing the k-nTS+ swapping method in Figure 2, the authors should provide more details on how it works. The framework now is primitive.

Thank you for this suggestion, we have increased the amount of detail in Figure X on page XX to better illustrate our proposed k-nTS+ method.

2. Section 2.3:  The literature review of time series features for forecast accuracy could profit from including the relevant works such as:  
  
   Kang Y, Cao W, Petropoulos F, et al. Forecast with forecasts: Diversity matters[J]. European Journal of Operational Research, 2022, 301(1): 180-190.  
  
   Li L, Kang Y, Petropoulos F, et al. Feature-based intermittent demand forecast combinations: accuracy and inventory implications[J]. International Journal of Production Research, 2022: 1-16.  
  
   Montero-Manso P, Athanasopoulos G, Hyndman R J, et al. FFORMA: Feature-based forecast model averaging[J]. International Journal of Forecasting, 2020, 36(1): 86-92.

We appreciate this suggestion and have added these sources to our literature review on page XX.

3. Section 4.1: Why only use the monthly micro dataset from M3 competition? I recommend using all M3 competition data and discussing the performance of the proposed k-nTS+ method for the data with different frequencies. More recent M4 Competition data is also a better option.

Thank you for this suggestion. As mentioned in *Major Changes to Manuscript (2)*, we now include the full M3 competition data set and discuss the performance of the proposed method across different data frequencies. These results, included on pages XX – XX, show that…. We also include results from a subset of the M5 competition to test the proposed method in a hierarchical forecasting scenario. We find that…. These results are included on page XX.

4. Section 4.2: The authors should clarify the details of the feature selection, e.g., why select such features for k-nTS and add new features for k-nTS+. The process seems subjective. The authors should give all the alternative features and explain the reason.

Thank you for pointing this out. On page XX, we now differentiate between our manual feature selection based on the literature (sources here XX) and our proposed feature selection method. The proposed feature selection method is flexible and can accommodate any number of single-valued features, given sufficient runtime from the data owner (relates to your point #5 below). For our purposes, we select a large set of features that (1) occur commonly in the literature (Sources XX,XX), and (2) can be conveniently computed using the *tsfeatures* package in R. However, data owners can include any features that they are interested in. We discuss these details on page XX. On average across the different M3 frequencies, our proposed feature selection method for k-nTS+ improves forecast accuracy by XX% compared to manual feature selection.

5. A brief discussion of the computational cost is useful for other researchers.

We agree and have included this discussion on page XX.

6. Is there any reason for the error measures used in sections 4.4 and 4.5 to differ? If MSE does not provide significantly different results than MAE, personally I would prefer a consistent measure to be used for both sections.

Thank you, we now use MAE to assess both the accuracy of forecasts in Section 4.4 and the accuracy of the random forest predictions of the forecast MAE in Section 4.5.

7. Figure 6: Each diagram should be numbered differently, such as A.1, A.2, A.3, B.1, B.2, B.3.

Thank you, we have adjusted Figure XX accordingly.

8. Figure 7: The ordinate title should not have "average"? It is unclear what "time series features for each privacy method" means exactly.

Thank you for pointing this out. We have adjusted the titles for Figure XX to be more informative.

**Reviewer 2 Comments:**

The authors propose a method for preserving data privacy in time series data through a swapping technique. This approach focuses on maintaining forecast accuracy by swapping the data values only if the essential features of the time series, such as mean and autocorrelation function (ACF), are likely to remain unchanged. The proposed method assumes a centralized approach, where a single data owner possesses the time series data. In this scenario, a forecaster selects a forecasting model F, and the data owner performs data swapping to prevent a decline in F accuracy.  
  
While the idea sounds interesting, the paper requires further clarification and enhancements to address the following points:  
  
- Applications: The authors should provide further clarification on the potential applications of their proposed method.

Thank you. As per *Major Changes to the Manuscript (1)*, we outline potential applications of our proposed method swapping method, including imputing missing values, nowcasting, and anonymizing sensitive time series, on page XX. We also highlight that our machine learning-based feature selection method can select features for predicting any categorical or continuous target value.

- The authors should explain why the data owner cannot provide the forecaster with the original or degraded model weights. This comparison would help illustrate the advantages and limitations of the proposed approach. Additionally, discussing whether this method could be extended or applied to cases where data are owned by multiple data owners (decentralized scenario) would be beneficial.

Thank you for the suggestion. The main advantage of our methodology is that through sharing protected time series, forecasters gain access to the entire distribution of time series values, rather than just model parameters. In our empirical application, we do not assume one forecasting model, rather we create a protected data set that works well for several models of varying complexities.

We cite multiple sources (Citations XX) that focus on scenarios where a data owner provides the forecaster with model weights. But, …

* this can expose sensitive time series values (citation)
* the data owner must train and adjust models on behalf of the forecaster
* Specific use cases, *e.g.*, a forecasting competition, require time series data.

Our method can also be applied to decentralized scenarios:

* Data owners could each apply the method to their own time series before sharing/pooling data together
* A central (trusted) party could apply the method after the data from multiple owners have been pooled – this could increase willingness to share data by reducing privacy concerns (citation) (an alternative solution to incentivizing sharing through a data market).
* If data is limited, it might make sense to augment data owners’ series with synthetic time series with similar features to give them something to swap with (future research?)

We include this discussion on page XX in Section XX.

- A crucial aspect missing is the impact of the swapping process on forecasting accuracy in a real-world setting. For instance, if a forecaster aims to perform a one-timestep ahead forecast using a VAR (Vector Autoregression) model with a lag of 1, he would require the value of X[T] to predict X[T+1]. Therefore, it is essential to evaluate whether the swapping process can change the last point of the time series without significantly affecting the forecasting accuracy.

Thank you. We now include results specifically for a VAR model with a lag of 1 on page XX and discuss how swapping values directly affects model accuracy through the model equations.

As per *Major Changes to the Manuscript (2)*, we also include results from applying the proposed method to a subset of the M5 competition data to illustrate the real-world setting of protecting retail data. The results on page XX suggest that the methodology is suitable for protecting retail data, but requires adjustments to maintain the values of aggregate (less sensitive) time series while swapping the values of lower-level (sensitive) time series.

Figure 6 in the paper shows that the last point in the window is very close to the original one. Does this mean that a curious forecaster could reconstruct the data by running the model for some time?

Thank you for this comment. While the last value in the window appears to be very close to the original one, this will not always be the case. For example, the protected point in time period 59 is significantly lower than the original. However, time series reconstruction is a valid concern particularly when time series models are shared that utilize lags of sensitive time series values (Citation here).

One of the advantages of differential privacy is that differentially private data is immune to post-processing (Citation), so any forecast or attempted reconstruction of the original time series would also be differentially private. However, at acceptable levels of , the forecasts are unusable. By increasing a data owner would improve forecast accuracy while maintaining some bound on overall privacy risk, however, our empirical application shows that results in significantly higher identification disclosure risk than our proposed method k-nTS+. Essentially, data owners must tradeoff between bounding all privacy risks and creating protected data that produces usable forecasts.

Your question points to the fact that there are privacy risks, such as time series reconstruction or attribute disclosure (*i.e.*, are time series features themselves privacy risks?), that we do not address in depth in this paper. One potential solution to limiting attribute disclosure is to remove sensitive time series features from consideration in the swapping process. We assess this approach by excluding some features which could be considered sensitive (*e.g.*, the mean of the time series) and were initially included in the swapping process. We measure the resulting impact on forecast accuracy and find that… These results are included in the appendix on page XX.

Overall, the privacy risks associated with time series data are important areas for future research, and we discuss these points on page XX.

Detailed comments:  
- Acronyms meaning is missing. Some examples: SES, DES, LGBM, OOB, MSE, MAE, etc.

We have now provided the meanings for all acronyms.

- Equations should not be figures, e.g., (1) and (5).

We have re-written the paper in LaTeX and all equations are now properly formatted.

- Figures 2 and 8 should be of better quality.

We have revised Figures X and X to improve their quality.

- Figure 4: I suggest a unique plot with different color/shape lines representing the methods.

Thank you for this suggestion. We have revised Figure X accordingly.

- Notation needs to be introduced appropriately in many equations - see, e.g., (9).

We have added introductions to all notation where appropriate in the text.

- I would include the detailed proposal version in the main text.

We have included the detailed version of the algorithm in the main text on page XX.

- Text needs revision. Typo example: "matices", sometimes ':' is used instead of '='.

Thank you. We have carefully edited the text for typos and to improve readability.