

Response letter - Manuscript ID 249381018

Dear Professor Zhang

Please find enclosed a revised version of our paper submitted for peer review for the *Journal of Operational Research Society*. Your comments have been very much appreciated and we have tried to suitably modify/amend our paper for those comments to be reflected in the revised version.

We quote below the detailed referees' comments that are followed by our response. In all cases, we point out, where necessary, the corresponding amendments in the new version of the paper and highlighted them in teal (greenish-blue color).

Editor comment

Your manuscript entitled "Forecasting interrupted time series", which you submitted to Journal of the Operational Research Society, has been reviewed. The referee comments are included at the bottom of this letter. The referee(s) would like to see some revisions made to your manuscript before publication. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. When you revise your manuscript please highlight the changes you make in the manuscript by using the track changes mode in MS Word or by using bold or coloured text.

Response: Thank you for giving us an opportunity to revise the manuscript. We have thoroughly reviewed the feedback provided by the referees, and made some changes as a result. We have highlighted all changes with coloured text in both the response letter and the manuscript.

We also wanted to provide further clarification to the editor regarding the comment provided by Reviewer 4.

- Reviewer 4 (comment 2) suggested including mathematical equations for each strategy. While we could add these equations in Section 2, we believe it would not improve the readability or accessibility of the paper. However, if the editor prefers a more mathematical description of each strategy, we are willing to include the equations. Since we are not conducting a mathematical analysis of the strategies, we have opted for a verbal description supported by code implementations.
- Reviewer 4 (comment 4) asked to add descriptive statistics. We do not believe the addition of descriptive statistics will add anything useful to the manuscript. We argue that the usefulness of descriptive statistics for time series analysis is limited. We have provided the raw data, and we have provided time plots to help visualize the data. Each data set is described in Section 3.
- Reviewer 4 (comment 5) recommended including three papers from the literature. However, these papers are not relevant to our current study. The suggested papers focus on forecasting COVID-19 variables, which is not the focus of our paper.

Reviewer 1

General comment: *The paper deals with the prediction of time series when a disruption happens in the data. Different prediction strategies are proposed, simulated and performances are compared. The methodology is well explained, all data sources and codes are available for replication and improvements. Results are clearly exposed, criticized and recommendations are provided. In my opinion, the paper is very good in its present form and I recommend publication.*

Response: Thank you for your positive feedback on our paper. We are pleased to hear that you found the methodology well-explained and the results clearly presented. We appreciate your recommendation for publication.

Minor comments

Comment 1. one typo, line 49, page 4 : “remaining”

Response: This has been corrected

Comment 2. ref Trapejo (2023) is incomplete

Response: This paper is currently available online. We have included the designation ‘In Press’ in the reference to reflect its status.

Reviewer 2

Comment 1: The topic is very interesting and very timely, in the wake of the Covid -19 related disruptions in supply chains, changes in consumer preferences, and inflationary environment. The time series patterns and causal relationships that were seen before the pandemic did not apply and whether the patterns observed before the pandemic will continue to be relevant remains to be evaluated on a case by case basis.

Response: We are delighted to hear that you appreciate the timeliness and importance of the paper. We agree that the topic is particularly relevant given the COVID-19-related disruptions in supply chains and various other areas. We also agree that evaluating whether pre-pandemic time series patterns and causal relationships remain applicable is a critical challenge that requires careful case-by-case analysis. That is why, in this paper, we propose various strategies rather than a single approach to address different situations and cases effectively.

We have added the following paragraph in the revised manuscript to address this comment:

The handling of time series interruptions is a complex and multifaceted issue, and the choice of the most appropriate strategy depends on the specific characteristics of the data and the nature of the interruption. Therefore, in this paper, we propose multiple strategies rather than a single approach to effectively address different situations and cases.

Comment 2: I like that the authors contrast the topic with others that it may be confused with. But need to check the claim that it is indeed the first to address this problem.

Response: Thank you for your comment. The idea for this paper emerged from our close contact with practitioners, who highlighted the challenges they face in forecasting time series data affected by COVID-19, particularly in handling observations impacted by the disruption. Following these discussions, we carried out a search to find out if any existing papers addressed this issue, but we could not find any published research in this area. To our knowledge, this is the first study to describe and compare general strategies for forecasting interrupted time series, such as COVID-19-affected data.

Comment 3: The paper is well written, but in terms of contributions, it falls short of the expectations for an original research paper. Improvement avenues are proposing a viable new method that has advantages over these existing methods, or a large-scale empirical research with carefully curated set of actual data representing different interruption patterns from different domains.

Response: Thank you for your feedback. We appreciate your comments on the writing quality and your recommendations for improving the paper's contributions. We believe that the originality of our research comes from addressing the actual challenges that many forecasters and modellers experience when their time series data is disturbed by recommending multiple viable and reliable strategies suitable for different cases. We recognise that presenting a new technique may be an exciting direction, but it is not the aim of this paper. So we have included it as a potential topic for future research in this area. Also, we included two empirical datasets from two different application areas in the paper. Moreover, the whole paper is reproducible, and the datasets and R codes are accessible on GitHub, allowing strategies to be tailored to any dataset in any application.

We have added the following paragraph in the revised manuscript.

This study proposes multiple viable strategies for forecasting time series data interrupted by different type of events. Future research could focus on exploring new strategies for handling such disrupted time series, or enhancing probabilistic forecasting methods that are designed specifically for such data. Large-scale empirical research on further exploring which strategy may work best in different scenarios also provide useful insights.

Reviewer 3

General comment: I would appreciate the editor for providing me an opportunity to review the article "Forecasting interrupted time series". In this paper, the authors investigate several strategies for dealing with interruptions in time series forecasting and make some comparisons among these methods. Though the paper is overall interesting, but the empirical design seems to be too simple. I hope these comments may be helpful for the authors.

Response: Thank you for taking the time to review our article. We appreciate your feedback and your acknowledgment of the interesting aspects of our work.

Comment 1. The authors regard the description and comparison of general strategies for forecasting interrupted time series as a major contribution, which in my opinion is quite limited.

Response: We believe that our research tackles fundamental challenges that many forecasters and modellers face when working with time series data that has been disrupted by events such as COVID-19, strikes, etc. This challenge was highlighted throughout our conversations with academics and practitioners in the field, who emphasised the issues they face when forecasting such time series data. Therefore, we think addressing this practical challenge is where our work stands out. Further, we have not found any other papers addressing this issue. Consequently, we feel that we are addressing an important and widespread problem.

Additionally, our work may open avenues for further research and reflection in this area. Also, we made sure that the entire work was reproducible, with datasets and R codes available on GitHub, allowing other academics and practitioners to tailor the strategies to their own needs.

Comment 2. The patterns of interrupted events may be very different, which makes the suitable approach varies. In this case, it may be better to focus on a certain event (e.g., COVID-19 pandemic) and make comprehensive comparisons.

Response: We agree that the patterns of interrupted events may vary, which would affect the suitability of proposed approaches. That is why we have proposed different strategies to consider that might be appropriate in different cases, each with its own pros and cons. Focusing on one event would greatly limit the scope of the paper, making it far less useful.

In future work, a large scale empirical study covering several types of disruption could be conducted to provide deeper insights and any possible link between strategies and types of disruption events. We have included the following paragraph in the conclusion section of the revised manuscript:

This study proposes multiple viable strategies for forecasting time series data interrupted by different type of events. Future research could focus on exploring new strategies for handling such disrupted time series, or enhancing probabilistic forecasting methods that are designed specifically for such data. Large-scale empirical research on further exploring which strategy may work best in different scenarios also provide useful insights.

Comment 3. The authors provide some visualizations on the forecasting. This should be appreciated since it is quite straightforward. However, the authors may not expect the readers to judge the optimal strategy from these figures. Instead, the authors should consider some out-of-sample validation and may introduce some metrics or measures to evaluate the performance different strategies.

Response: The visualizations are intended to help the reader understand the effect and intent of different strategies. In determining which strategy to apply in each case, it is important to think through what is being assumed, and what data are being used. We think the visualizations help the reader do that. Also, it does not always make sense to do a quantitative comparison of forecast accuracy. We have attempted to address this issue by adding the following paragraph in the discussion section.

Moreover, in determining which strategy to apply in each case, it is important to think through what is being assumed, and what data are being used. This is a qualitative and subjective process, and cannot simply be tackled with a quantitative comparison of forecast accuracy. The strategies we discuss have different ways of handling the disrupted period, and many of them do not seek to provide accurate forecasts during the disruption. For example, the "set to missing" strategy ignores the disrupted period, while the "estimate what might have been" strategy replaces the observations during the disruption with interpolations. In both cases, it does not make sense to evaluate the performance of the method during the disruption. Instead, the analyst needs to decide what approach makes the most sense for the application at hand.

Comment 4. If the authors focus on the interrupted time series for a certain event, some exogenous variables are encouraged for potential better performance.

Response: Thank you for your suggestion. We agree that incorporating exogenous variables related to specific events could enhance the performance of forecasting models for interrupted time series. We have acknowledged this by adding the following paragraph to the revised manuscript:

When forecasting interrupted time series data, exogenous variables associated with specific events should be identified and incorporated where possible, regardless of which approach is employed, since they may improve forecasting model performance.

Reviewer 4

General comment: The manuscript deals with forecasting interrupted time series for forecasting teams, especially in light of emergency events like COVID-19. The manuscript investigates several strategies for dealing with interruptions in time series forecasting, and their advantages and disadvantages. The authors evaluated two datasets to understand the effectiveness of these strategies. The manuscripts deal with important areas and insight into the strategies.

A few suggestions need to be included such as:

Comment 1: Different strategies are discussed in section 2 in a few words without giving a proper introduction of the strategies. It makes the reading of the manuscript difficult. The authors are required to give practical examples of the strategies to make it understandable to read.

Response: Thank you. We are not clear what is being requested here. We describe the strategies in Section 2, and we provide fully reproducible code showing how they are implemented. We also give practical examples of the strategies in Section 3. Please explain what more is required to be included, and we are willing to do that.

Comment 2. No mathematical support is provided as to how the strategies work. And what are the practical implementation of the strategies? Authors are required to provide detailed insight into the strategies i.e. which datasets these strategies are applied?

Response: Thank you for the comment. We have carefully thought through this request. While we could add equations in Section 2 for each strategy, we do not think it would add to the readability or accessibility of the paper. We are not undertaking a mathematical analysis of the strategies, so we have opted for a verbal description accompanied by code that implements each strategy. However, if the editor would prefer us to add equations providing a more mathematical description of each strategy, we can do so. To respond to other remarks:

- The practical implementation of the strategies is provided in Section 3 and in the supplementary code.
- The data sets are described in Section 3.
- Insights are provided in Sections 3 and 4.

Comment 3. The figures presented in the manuscripts require detailed explanation. Authors are required to explain the given figures in detail one by one.

Response: Thank you for your feedback. We have provide a detailed explanation of each figure in the manuscript, discussing them one by one to ensure clarity. We have now added some of the same information in the captions of each figure to aid readability.

Comment 4. Two datasets are used and the strategies are applied to them. However, the descriptive statistics of these data sets missing.. The authors are required to explain the datasets.

Response: Thank you for this comment. We do not believe the addition of descriptive statistics will add anything useful to the manuscript. We argue that the usefulness of descriptive statistics for time series analysis is limited. First, they cannot highlight patterns in the data, identify unusual observations, describe changes over time, or help to understand relationships between variables. Second, they could be misleading, since similar statistics can stem from very different datasets (please see Anscombe's Quartet).

We have provided the raw data, and we have provided time plots to help visualize the data. Each data set is described in Section 3.

Comment 5. Literature must reflect the relevant studies i.e.

- a) *OntoCOVID: Ontology for Semantic Modeling of COVID19 Statistical Data* Authors Shaukat Ali, Shah Khusro, Sajid Anwar, Abrar Ullah Publication date 2022 Book *Proceedings of International Conference on Information Technology and Applications* Pages 83-194 Publisher Springer, Singapore
- b) *COVID-19 patient count prediction using LSTM* M Iqbal, F Al-Obeidat, F Maqbool, S Razzaq, S Anwar, A Tubaishat, ... *IEEE Transactions on Computational Social Systems* 8 (4), 974-981
- c) *Countering malicious URLs in internet of things using a knowledge-based approach and a simulated expert* S Anwar, F Al-Obeidat, A Tubaishat, S Din, A Ahmad, FA Khan, G Jeon, ... *IEEE Internet of Things Journal* 7 (5), 4497-4504

Response: Thank you for your comment. We would like to clarify that our study does not focus on forecasting COVID-19-related data and variables, such as COVID19 patient count, or statistical data, as suggested. So none of these references appear to be relevant to our paper. Instead, our paper addresses interrupted time series data caused by events like pandemics, strikes, and other disruptions. We have included literature relevant to this topic. We hope this clarifies the scope and focus of our research.