

Peer Review

Review of: "Spatiotemporal Forecasting in Climate Data Using EOFs and Machine Learning Models: A Case Study in Chile"

Arnobio Palacios Gutiérrez¹

1. Universidad Complutense de Madrid, Spain

This study presents a hybrid methodology capable of regionalising climate prediction in Chile. The methodology integrates machine learning (ML) methods for time series prediction with statistical techniques. The results and conclusions give relevance to the study, but for its publication, a general review and refinement of the paper should be done. Here are some suggestions and comments that may be useful to enhance its impact and innovation:

Abstract: The abstract seems to me to be not very tidy and needs to be rewritten. It is not easy to intuit clearly the proposed procedure followed in this study. Generally, the methods and techniques are mentioned, indicating some use, but there is no clear thread or sequence to the procedure followed. In addition, the summary should be reduced, and if possible, the symbolic denotations should be included from the introduction. I also need to see something relevant in relation to the results, which could be indicated at the end of the abstract to show the relevance of the study.

Introduction: The introduction does not set out the innovative aspects of its methodological development. This makes it difficult to highlight the particularity of the study. It might be useful to consider the following:

- The introductory section lacks an exhaustive review of the literature and does not allow for a clear understanding of the limitations of the knowledge that this study aims to cover.
- Currently, the introduction focuses on the techniques and methods used in the study, indicating uses, sometimes in a generalised way, and does not allow for a clear thread of the methodology itself in this section, which is then repeated in the methodology section itself.

- On several occasions, the authors indicate: “This methodology facilitates the transition from a high-dimensional multivariate spatiotemporal data forecasting problem (in this case, entailing forecasting using 6355 time series corresponding to grid points) to a low-dimensional univariate time series forecasting problem (in this case, up to a couple of dozens of forecasts), significantly reducing computational complexity while yielding forecasts of reasonable utility.”, however, they do not clearly explain how this transition is facilitated. Would this help emphasise the importance and motivation of the study? If this contributes to highlighting its novelty, the authors should articulate it clearly alongside its relevance within a broader climatological context.

Methodology: In my view, the methodology should be reviewed and improved. The techniques and methods are generally presented in a dense manner, which could be simplified further to aid the reader’s comprehension without losing the flow of the discussion, which at times is disrupted. Rather than merely stating generalised definitions of the techniques used, the authors should establish a methodological framework that clearly defines the purpose of employing these methods and techniques, indicating how they complement each other in achieving the study’s objective. Additionally, some paragraphs included in the methodology section would be better placed in the results section.

Results, Discussion, and Conclusions: The results are extensive, and the regionalisation of forecasts in Chile, obtained through clustering, allows for observations that may be of interest to climate research. I believe the discussion and conclusions sections should be separated, while the discussion should instead be merged with the results. This would involve conducting a comparison with relevant previously published studies, as the current approach appears somewhat limited when combined with the conclusions. Structuring it this way would give greater significance to the results. The conclusions are clear and effectively highlight the relevance of the study.

Others:

- When introducing an acronym or abbreviation for a term, the acronym should continue to be used thereafter. In some instances, “Empirical Orthogonal Functions (EOF)” is repeated after its initial introduction, instead of using “EOF” alone. This aspect should be reviewed throughout the document for all abbreviated terms.
- In the results section, only one subsection is indicated. The authors may consider including additional subsections within the results, which, together with the discussion, could enhance the article’s readability and comprehension.

- Some paragraphs in the results section contain information that could be more effectively placed in the methodology section to improve its clarity.

Declarations

Potential competing interests: No potential competing interests to declare.