

Additions are shown in red, modifications in blue

Reviewer: 3

Comparative Performance of ARIMA and Stepwise Regression for Forest Fire Prediction in Yunnan Province

Comments to Authors:

Abstract

1. The title of the study tells us about the comparison of the dual tools used in the study, meanwhile, the abstract tells us about the different story. See Lines 14-18.

We have revised the abstract to better align with the study title, ensuring consistency in communicating our comparison of ARIMA and Stepwise Regression tools for forest fire prediction.

2. Lines 12-13: What is the research problem? Occurrences of Forest Fires or Lack of Prediction Tools? In my opinion, there are a bunch of modern applications available that can easily predict future scenarios. If the problem is forest fire occurrences, then certainly there are natural and manmade phenomena that speed the fire occurrence process.

Modern applications do exist, but our study aims that by comparing these two models, we can fully understand the dynamics of forest fires from a time series perspective as well as from a multivariate analysis perspective. This comparison is not just a technical comparison, but also an innovative attempt to explore the effectiveness and limitations of the different models in dealing with the same problem, which can promote model optimisation, and by comparing these two models, we can By comparing these two models, a more comprehensive and diversified perspective can be provided for the prediction of forest fires in Yunnan Province.

3. Why did authors compare or experience ARIMA and Stepwise Regression?

ARIMA and Stepwise Regression were compared to evaluate their predictive performances and applicability to forest fire data, capitalizing on their unique statistical strengths.

4. The authors did not talk about the data collection and feeding process to generate results in the abstract. Correction is required.

We have now included a brief on data collection and processing in the abstract to clarify how the results were generated.

5. What is the significance of the study?

The significance of this study lies in the comparison of prediction accuracy and the analysis of meteorological factors that have a strong influence on the number of fires and fire area through stepwise regression. It aims to improve the accuracy and reliability of forest fire prediction.

Introduction

1. This section lacks important facts and figures about forest fire occurrences in the world and hence in the study area. Show that your study is need-based. This is a massive correction to upgrade the quality of the introduction section.

We will incorporate current data on global and regional forest fire occurrences to demonstrate the necessity and timeliness of our study.

2. Clarify that selected methods are the best option to predict the forest fire. Case studies should be cited to link with the upcoming section of methodology. The link is missing.

The superiority of the selected methods will be justified with relevant case studies, establishing a clear connection to the methodology section.

3. The citation numbers 11-15 are outdated. I wonder why the authors included these to clarify fire risk prediction models. The situation is changed as described in these articles from 11-15. The authors are advised to search for the latest citations to validate their arguments.

We will update the citations 11-15 with the latest literature to ensure our discussion on fire risk prediction models is current.

4. I also urge authors to relook at their citation portion as given in different sections of the manuscript. Use fresh citations. These should not be older than a decade.

The citation portfolio will be refreshed with recent literature, adhering to the guideline of using citations from within the last decade.

Study area and research methodology

1. Not a single citation was found between lines 76—83. The authors have referred to important parameters in these lines, however, the authors have referenced certain literature. This gives a bad impression.

We will add citations between lines 76-83 to substantiate references to the key parameters temperature, mean annual rainfall, wind speed and relative humidity.

2. Published data were taken from 1991-2021 taken from the different agencies. Is it good to analyze published data? As a statistician, published data should not be further analyzed. If so, researchers should take expert opinion to validate the choice of secondary data usage. Therefore, validation and authentication are required.

Our rationale for using published data was its relevance and accessibility, which ensured the robustness of the study, and also the consistency of the data for the purpose of carrying out our modelling studies, and we took into account expert opinion when selecting secondary data.

3. Subsection 3.2. No citation is given for the formulas (1) to (4). Why do you discuss these here? What was the purpose and limitations?

Citations for formulas (1) to (4) will be provided, clarifying their relevance and limitations within our methodology.

4. See lines 98-103. How was the data obtained from the government agencies? In what year, data were collected?

Details regarding the acquisition of government agency data, including the year of collection, will be clarified.

5. For missing data, why long-term averages were used? This may generate outliers and vanish the aim of the analytical process.

Long-term averages, calculated on the basis of many years of observational data, are a better reflection of the typical climatic conditions in the region and are considered to be a reliable proxy for climatic characteristics. While long-term averages are valid alternatives, they may not fully reflect recent trends in climate change.

6 Line 160-162. This is a biased statement. The evidence is also not seen by referring to Figure 3. 1998-2000 the massive outlier is visible. Clarify discussing Figure 3.

We will re-examine the statement in lines 160-162 and clarify the discussion around Figure 3, addressing the outlier in 1998-2000.

7. What is hm²? Looks like a unit of area. Clarify. Also, define all such units to make the manuscript readable.

The unit "hm²" refers to a hectare (10,000 square meters), and we will define all units clearly for readability.

Temporal Characteristics of Forest Fires

1. This section is overstretched and looks hazy. One cannot figure out why these steps were performed by the authors. Are these performed in the quest to achieve study objectives? It is not clear in this section.

This section will be streamlined to clearly articulate the steps performed in relation to achieving the study objectives.

2. Multiple tests and steps were performed in this section without justification. Their link with the reviewed literature and methodology section is also missing.

We will add the rationale for performing multiple tests and steps, linking them to the literature and methods reviewed

3. Draw a flow chart to summarize the required steps to achieve objectives that need to be utilized in this section.

A flow chart summarizing the steps toward achieving the study objectives will be included for clarity.

4. The authors need to correct this section. Do not go out of scope. Compare this section with others to establish a link.

This section will be corrected to maintain scope and establish coherence with other sections.

5. Many of the outliers were observed in the Figures.

The presence of outliers in figures will be addressed and discussed.

Conclusion and Recommendations

1. This section merged two important sections. It is advised to split. Recommendations are made in the light of key findings of the study. Whereas, the conclusion section sums up the study in a professional way.

The sections on conclusions and recommendations will be separated for clarity, with recommendations stemming from key findings and conclusions summarizing the study professionally.

2. Highlight the limitations of the study. As some of the data inputs are found missing.

Justifications for the multiple tests and steps performed will be added, linking them to the reviewed literature and methodology.

The limitation of this study is that it is limited to the analysis of historical data and does not take into account the impact of possible future climate change on fire

patterns. Also, the predictive accuracy of the model is limited by the quality and availability of the data.

3. Significance and research contribution should be added as separate subheadings to clarify the importance of the study.

Significance and research contributions will be articulated under separate subheadings to emphasize the study's importance.

4. Why concerned agencies, reviewers, or students review this study? Justify.

This study has important applications for relevant government agencies, researchers and students. It provides a scientific basis for decision makers to help them make more informed choices in resource allocation and risk management. In addition, the methodology and findings of this study can be used as teaching cases in courses related to environmental science and forest management.