Pakistan Journal of Statistics and Operation Research

Using the Single-Exponential-Smoothing Time Series Model under the Additive Holt-Winters Algorithm with Decomposition and Residual Analysis to Forecast the Reinsurance-Revenues Dataset

Murtadha Mansour Abdullah

* Corresponding Author

University of Wasit. College of Administration and Economics, Iraq; mabdullah@uowasit.edu.iq

Abstract
Time series analysis plays a pivotal role in the strategic planning and risk management of reinsurance companies. It is an indispensable tool for gaining insights into the future utilization of reinsurance revenues. To effectively safeguard against substantial financial losses stemming from anticipated claims, reinsurance businesses must have a thorough understanding of the expected values of these claims. The ability to estimate the potential value of future claims is paramount, as it empowers reinsurance companies to proactively prepare and allocate resources, ensuring that they are well-equipped to cover likely future claims. Our research incorporates an innovative approach to estimate reinsurance revenues, leveraging the power of time series analysis. By applying the proposed paradigm to an original time series dataset, we aim to showcase its practical value and effectiveness in predicting future revenue trends. To assess the accuracy of these predictions, we employ the Box-Ljung statistical test, a statistical test commonly used in time series analysis. The corresponding p-value generated from this test provides a quantitative measure of the ability to analyze, capture and explain the underlying patterns in the data, thereby aiding reinsurance companies in providing an informed decisions and managing their financial risks effectively. In summary, the integration of time series analysis, single exponential smoothing (SEXS), and advanced forecasting techniques forms a critical foundation for enhancing the predictive capabilities of reinsurance businesses and ensuring their financial stability in the face of uncertain future claims.

Key Words: Cullen-Frey graph; Holt-Winters' Additive Algorithm; Box-Ljung Test; Reinsurance Revenues; Residuals Analysis; Forecasting; Single Exponential Smoothing.

1. Introduction

Reinsurance is often known as the “insurance of stop-loss” or the “insurance for insurers”, is a protective measure that organizations adopt to mitigate their exposure to substantial claim occurrences. It involves an arrangement where insurers change and transfer a certain part of the risk portfolios to others, reducing the likelihood of shouldering substantial liabilities arising from insurance claims. The entity that diversifies its insurance portfolio and offloads some of the risk is referred to as the ceding party. Conversely, the reinsurer is the entity that agrees to assume a share of the potential liability from the primary insurer in exchange for a portion of the insurance premium. In essence, the reinsurer provides insurance to the primary insurer, helping them manage their risk exposure and ensuring they have the financial capacity to cover large or unexpected losses. This arrangement allows the primary insurer to underwrite more policies than they otherwise could on their own, thus spreading risk across multiple parties within the insurance industry (see Mohamed et al. (2022 and 2024)).

Through reinsurance, insurers can maintain profitability by recovering a portion or the entirety of funds paid to claimants. This strategic approach not only diminishes the net liability associated with individual risks but also provides a safeguard against catastrophic losses stemming from significant or widespread events. Additionally, it enables companies seeking reinsurance, known as ceding companies, to enhance their capacity for undertaking risks across various levels. In light of these crucial aspects and current global trends, we present a model for projecting the revenues of reinsurance companies. This model employs straightforward methodologies and stands out for its high predictive accuracy, setting it apart from many expensive models with challenging requirements. Moreover, reinsurance may serve secondary purposes beyond its primary role in risk management, including reducing the capital requirements of the ceding firm, mitigating tax burdens, or fulfilling various other strategic objectives.

The topic of forecasting the revenues of reinsurance companies is one of the important topics in the field of actuarial statistics. In statistical literature, there are many methods, techniques, and models for statistical prediction. It is difficult to say that a certain technique, method, or model is the best ever. This issue of preference among techniques, methods, and models is determined by many factors and reasons, including required accuracy, cost, and speed. In this paper, we will take great care to balance all these factors in order to provide reliable and accurate forecasts to an acceptable and sufficient extent.

The renowned Box and Jenkins, often referred to as Bx-Jn, introduced the autoregressive-integrated-moving-average (ARIMA) models with the primary objective of harnessing time series data to make accurate predictions. Box and Jenkins' groundbreaking work, as cited in created by Box and Jenkins (1970), laid the foundation for advanced time series forecasting methodologies. The essence of the Bx-Jn technique lies in its ability to forecast data by drawing upon the historical information encapsulated within a specific time series. Over time, as data points accumulate in a structured manner, this method becomes adept at making predictions by analyzing the evolving trends and patterns inherent in the time series data. Central to this methodology is the assessment of variations and differences among the various data points collected over time. By employing a combination of moving average models, seasonal variations, and autoregressive models, the Box-Jenkins approach diligently seeks out meaningful patterns within the time series data. These patterns, once identified, serve as valuable insights into the future, enabling more accurate and informed predictions.

The Box-Jenkins method primarily finds its application in conjunction with ARIMA models, as these models provide a robust framework for capturing and modeling the complex interplay of autoregressive and moving average components present in time series data. For further details on time series analysis with forecasting and control, readers are encouraged to explore the comprehensive work by Box et al. (2015). This source offers a deeper understanding of the principles and methodologies behind time series analysis, ultimately enhancing the ability to predict future trends and control outcomes in a variety of domains. Contrarily, studies utilizing ARIMA models have been widely published in the actuarial literature; for more information on forecasting vehicle insurance claims using econometrics and ARIMA models, see Cummins and Griepentrog (1985). See Jang et al. (1991) for an analysis of certain medical insurance plans for employees based on these models. See Venezian and Leng (2016) for a few spectral and ARIMA analysis applications. To explore the fluctuations in unemployment reinsurance revenues utilizing the ARIMA model, refer to Mohammadi and Rich's study (2023). Additionally, Hafiz et al. (2021) and Kumar et al. (2020) employed the ARIMA model to predict insurance penetration rates in Nigeria and motor insurance claim amounts, respectively.

Numerous authors have allocated their research on moving average (MA), autoregressive (AR), and ARIMA models in various contexts. Jakaša et al. (2011) explored electricity price forecasting, Sahu et al. (2015) for production forecasting, Iqbal et al. (2016) worked on predicting wheat output, Darekar and Reddy (2017) conducted forecasts for oilseed prices in India, Nath et al. (2019) predicted India's wheat production, and Palakuru et al. (2019) expanded upon Shrahili's (2021) research. Mohamed et al. (2022) introduced a novel synthetic autoregressive model for analyzing left-skewed reinsurance revenue datasets, building upon Shrahili's work. This approach involves a comprehensive assessment of reinsurance revenue data using various realistic ARIMA models to identify the most suitable one. The selection of the appropriate model is crucial for efficient insurance claims processing. Statistical evaluation is utilized to determine the significance of the model's parameters, with a preference for models with fewer critical parameters. The initial step in constructing a customized Box-Jenkins model for analyzing time series data related to reinsurance revenues entails assessing the stationarity of the time series and identifying any noteworthy seasonal patterns that should be replicated. Once the Box-Jenkins model is established, the autoregressive component is selected. Subsequently, reinsurance revenues are simulated using the synthetic autoregressive model, and its effectiveness is assessed through a series of simulated tests.

Therefore, different models can be suggested for forecasting, such as the SEXS model which was first presented by Brown (1959), Holt (1957), and Winters (1960). The SEXS is a straightforward and easy-to-understand forecasting technique. It requires minimal mathematical calculations and can be implemented quickly, making it accessible to a wide range of users, including those without extensive statistical knowledge. The SEXS method allows for adaptability to various time series patterns. It can handle data with trends, seasonality, or random fluctuations, making it versatile for different forecasting scenarios. The SEXS places more weight on recent observations, which means that it can quickly respond to sudden changes or shifts in the data. This makes it particularly useful for short-term forecasting or when there are rapid fluctuations in the time series. It is computationally efficient, even for large datasets. It does not require the storage or processing of a large number of historical observations, making it suitable for applications with limited computational resources. For further elaboration on this topic, additional insights can be found by referring to the works of Holt and Winters. Delving into these seminal texts can provide a deeper understanding of the subject matter, offering nuanced perspectives and comprehensive analyses. Brown's seminal contribution in 1959, Holt's seminal work in 1957, and Winters' seminal piece in 1960 serve as foundational pillars in literature, each offering unique perspectives and valuable insights. Therefore, consulting these references is highly recommended for those seeking a thorough exploration of the topic at hand.

The weights employed in SEXS based predictions diminish exponentially as time progresses, resulting in weighted averages that give more prominence to earlier observations. Put differently, more recent observations carry a higher weight in this method. This characteristic allows for the rapid and accurate generation of forecasts across a range of time series data, making it particularly advantageous for industrial applications. The Holt-Winters (H-WNS) method is a comprehensive forecasting technique that offers two distinct variations, each tailored to accommodate various seasonal dynamics. One of these variations adopts a multiplicative approach, which proves effective when seasonal patterns fluctuate in proportion to the overall magnitude of the dataset. In contrast, the additive approach is preferred when seasonal variations maintain a relatively constant magnitude throughout the dataset's duration, representing the seasonal component in absolute terms relative to the series' scale. In the additive model, the level equation plays a pivotal role by adjusting the dataset to account for seasonality, effectively removing its influence. Consequently, this adjustment results in seasonal components averaging out to approximately zero within each year, contributing to a more accurate depiction of underlying trends and patterns. Exploring these two methodologies within the Holt-Winters framework provides a robust foundation for understanding and implementing forecasting techniques in diverse scenarios.

In our research, we adopt the additive Holt-Winters (H-WNS) algorithm for modeling and forecasting historical insurance data, and we specifically consider the Single Exponential Smoothing (SEXS) model as part of our analysis. For the aim of assessing the forecasting performance of the SEXS model, we employ the mean of squared errors (MSEs) as a standard evaluation metric. However, the reliability of any forecasting model hinges significantly on the integrity of its residuals, and we conduct a comprehensive numerical and visual assessment of these residuals for both the SEXS and H-WNS models.

It's important to note that "exponential smoothing" is a comprehensive approach used to enhance the analysis of time series data. In contrast to exponential functions that involve weights diminishing exponentially over time, explicit moving averages utilize weights that increase linearly as time progresses. This approach allows for straightforward decision-making, particularly when considering factors like seasonality, making it applicable across various scenarios.

Exponential smoothing stands as a prevalent method for the analysis of time-series data, widely embraced across various disciplines. It's noteworthy to highlight the parallels between exponential smoothing techniques, such as SEXS and moving averages, with established filtering methodologies. Specifically, the comparison can be drawn between SEXS and first-order infinite-impulse response (IIR) filters, as both utilize equal weighting factors, emphasizing their shared approach in assigning significance to past observations. Similarly, the moving average method aligns with finite impulse response (FIR) filters, given its reliance on a finite window of past data points with uniform weights. This comparative analysis not only underscores the versatility and applicability of exponential smoothing techniques but also provides insights into their underlying principles, aiding in a deeper understanding of their functionality in time-series analysis.

Holt's contribution in 1957 expanded the capabilities of simple exponential smoothing by introducing a method for forecasting datasets with a trend component. This method incorporates a forecasting equation alongside two smoothing equations: one for the level and one for the trend. A significant aspect of Holt's linear approach is its ability to generate predictions showcasing a consistent trend—either growth or decline—that extends indefinitely into the future. However, empirical findings indicate that these methods often exhibit a tendency to overpredict, particularly when dealing with longer forecast horizons, as discussed in Holt (1957) and Hyndman et al. (2018). This underscores the importance of considering potential overprediction when employing such forecasting techniques.

The rest of Sections of this work is presented as follows: The SEXS model and its corresponding additive H-WNS' method are presented in Section sec2. In Section sec3, forecasting reinsurance revenue under the additive H-WNS' method is covered. Finally, Section sec4 offers some remarks.

2. The additive H-WNS' method

Exponential smoothing stands as a valuable technique in the realm of time series analysis, specifically tailored for generating concise short-term forecasts based on historical data. Its applicability shines when dealing with time series data that aligns with an additive model characterized by a consistent level and an absence of apparent seasonality. In such cases, simple exponential smoothing emerges as a straightforward and effective method for crafting short-term predictions. This approach particularly excels in forecasting data that lacks a clear seasonal pattern or discernible trend.

At the heart of this technique is the fundamental SEXS (Simple Exponential Smoothing) model, which serves as a means to estimate the current level within the dataset. The key parameter of this model plays a pivotal role in determining its accuracy when forecasting the current level at a specific time point. Notably, this parameter can assume values within the range of 0 to 1. A value closer to 0 signifies that recent data points carry relatively less weight in shaping future predictions, while a value nearer to 1 suggests that recent data points are accorded greater significance. This strategy of assigning varying degrees of importance to recent observations in contrast to older ones serves as a foundational concept underlying the principles of exponential smoothing.

In essence, exponential smoothing offers a valuable tool for short-term forecasting, particularly beneficial when dealing with data characterized by stability, lack of conspicuous seasonality, or apparent trends. The ability to adapt the parameter to prioritize recent data underscores its versatility and utility in generating accurate predictions, making it a valuable asset in the arsenal of forecasting techniques for time series data analysis.

The additive H-WNS algorithm, also known as the triple exponential smoothing method, is an extension of the SEXS technique. It is specifically designed to handle time series data sets that exhibit trends and seasonality. The algorithm takes into account three components: level, trend, and seasonality, and applies smoothing techniques to forecast future values. Overall, the additive H-WNS algorithm is a powerful tool for forecasting time series data sets that exhibit trend and seasonality. It enables accurate predictions by considering the level, trend, and seasonal components, making it valuable in various industries and domains where such patterns are present. Generally, the usage of the Additive H-WNS algorithm in forecasting can be summarized as follows:

1) The additive H-WNS algorithm is effective in capturing and forecasting trends in time series data. It considers historical trends and projects it into the future, providing valuable insights for predicting future values. This is particularly useful in scenarios where the data exhibits a consistent upward or downward trend over time.
2) The algorithm can handle time series data with seasonal patterns by capturing and forecasting the seasonal component. It analyzes historical seasonal fluctuations and applies smoothing techniques to project future seasonal patterns. This is beneficial when dealing with data that exhibits regular seasonal variations, such as sales data with monthly or quarterly patterns.
3) In addition to trend and seasonality, the additive H-WNS algorithm considers the level component of a time series. It estimates the average level of the data and incorporates it into the forecasting process. This is useful when there is a need to forecast the overall level of the data, independent of trend or seasonality.
4) The algorithm can handle time series data sets that exhibit multiple seasonal periods, such as daily, weekly, and yearly patterns. It can capture and forecast the interactions between these different seasonal components, providing accurate forecasts for each period. This is advantageous in industries where multiple seasonal patterns are present, such as retail or tourism.
5) The additive H-WNS algorithm can be used for out-of-sample forecasting, where historical data is available up to a certain point, and future values need to be projected. By incorporating the trend, seasonality, and level

components, the algorithm can generate forecasts beyond the available historical data, allowing for effective planning and decision-making.

6) The algorithm is widely used in demand forecasting, particularly in industries with seasonality and trend, such as retail, manufacturing, and supply chain management. It helps businesses predict future demand patterns, adjust inventory levels, optimize production, and plan for seasonal variations.
7) The additive H-WNS algorithm aids in capacity planning by providing insights into future resource requirements based on historical trend and seasonality. This allows organizations to allocate resources efficiently, anticipate demand fluctuations, and avoid underutilization or overutilization of resources.

The SEXS model guarantees that all anticipated values are equal to the series' most recent value, where

$$\hat{\mathcal{Z}}(\Xi + \boldsymbol{h})|\Xi = \mathcal{Z}_{\Xi}|\boldsymbol{h} = 1,2,3 \ldots,$$

where all future predictions are determines using a simple average of the observed data, namely

$$\hat{\mathcal{Z}}(\Xi + \boldsymbol{h})|\Xi = \Xi^{-1}\sum_{\Xi=0}^{\Xi}\mathcal{Z}_{\Xi}|\boldsymbol{h} = 1,2,3 \ldots$$

The underlying principle behind averaging methods in time series analysis hinges on the assumption that each observation in the dataset should be treated with equal importance and, consequently, assigned similar weights during the forecasting process. While this approach ensures fairness in considering all data points, it may not always align with the practical realities of many real-life scenarios.

In practice, we often strive to strike a balance between these two extremes. Instead of treating all observations uniformly, it is frequently more reasonable to give higher weight to recent findings compared to earlier ones. This nuanced approach recognizes that recent data points are often more indicative of the current state of affairs and are likely to have a more significant impact on future outcomes.

This concept dovetails seamlessly with the fundamental principle of exponential smoothing. Exponential smoothing acknowledges the importance of adapting the weight assigned to each observation based on its recency. Rather than rigidly applying equal weights to all data points, exponential smoothing embraces the idea that recent observations should carry more influence in shaping forecasts. This adaptability allows for a more dynamic and responsive forecasting technique that aligns better with the evolving nature of time series data, making it a versatile and effective tool for short-term predictions in various domains.

Due to the model of SEXS, the forecasting at a certain time $\Xi + 1$ can be considered to be equal to a weighted-average (WAV) between the most-recent observation $\mathcal{Z}_{\Xi}$ and the previous forecast $\hat{\mathcal{Z}}(\Xi)|\Xi - 1$, i.e.

$$\hat{\mathcal{Z}}(\Xi)\big|\Xi - 1 = \pi\mathcal{Z}(\Xi) + (1 + \pi)\hat{\mathcal{Z}}(\Xi)\big|\Xi - 1, \Xi = 1,2, \ldots,$$

where $0 \le \pi \le 1$ refers to the parameter of smoothing. Therefore, the new fitted values can then be re-expressed by

$$\hat{\mathcal{Z}}(\Xi + 1)\big|\Xi = \pi\mathcal{Z}(\Xi) + (1 + \pi)\hat{\mathcal{Z}}(\Xi)\big|\Xi + 1, \Xi = 1,2, \ldots,$$

We can use I_0 to represent the first fitting value (which must be guessed) at time one because the process must start someplace. Then,

$$\hat{\mathcal{Z}}(2)|1 = \pi\mathcal{Z}(1) + (1 + \pi)I_0,$$
$$\hat{\mathcal{Z}}(3)|2 = \pi\mathcal{Z}(2) + (1 + \pi)\hat{\mathcal{Z}}(2)|1,$$
$$\hat{\mathcal{Z}}(\Xi)|\Xi - 1 = \pi\mathcal{Z}(\Xi - 1) + (1 + \pi)\hat{\mathcal{Z}}(\Xi - 1)|\Xi - 2,$$
$$\hat{\mathcal{Z}}(\Xi + 1)|\Xi = \alpha\mathcal{Z}(\Xi) + (1 + \pi)\hat{\mathcal{Z}}(\Xi)|\Xi - 1,$$

Any exponential smoothing strategy requires the choice of the parameter of smoothing and the initial value I_0. In particular, we need to choose the values of and I_0 for simple exponential smoothing. If we are aware of these figures, we can use the data to produce all forecasts. For the techniques that follow, a variety of smoothing settings and beginning elements are commonly available. As a result, we determine the parameter values that are unknown and the beginning positions that minimize the of squared errors (SSEs). In contrast to the realm of regression analysis, where we have well-defined mathematical formulas that provide us with precise values for the regression coefficients, ultimately minimizing the SSEs, a measure of the model's fit to the data), the scenario we encounter here entails a more complex challenge. This challenge revolves around a non-linear minimization problem, necessitating the utilization of dedicated optimization tools to arrive at a solution. In regression, the relationships between variables can often be expressed in linear equations, leading to straightforward analytical solutions for the coefficients that minimize errors. However, when dealing with non-linear problems, such as those encountered in certain optimization tasks, the relationships between variables take on more intricate, non-linear forms. Consequently, there is no direct formulaic solution to determine the optimal parameter values. Instead, we must turn to specialized optimization techniques, algorithms, or computational methods to navigate the multidimensional parameter space and find the values that minimize the objective function, which could be a cost function or another performance metric. This transition from linear regression to non-linear optimization underscores the complexity of the problem at hand and highlights the need for computational tools that can efficiently explore and identify the best solutions within this non-linear landscape.

The H-WNS approach encompasses several distinct iterations, each tailored to handle specific types of seasonal components within time series data. One of these iterations employs the additive method, which is most suitable when seasonal fluctuations remain relatively constant over the entire duration of the series. In contrast, the multiplicative approach is chosen when seasonal variations alter proportionately with changes in the overall level of the series.

To adapt the data for the additive technique, the level equation plays a pivotal role in removing the seasonal component. This component is expressed in absolute terms, relative to the scale of the observed series. Consequently, by applying the additive method, the seasonal component is effectively "adjusted out" of the data, resulting in a series where the seasonal fluctuations tend to average out to approximately zero within each year.

The additive H-WNS damped method incorporates a damped trend and is intended to function in unison with the multiplicative and additive H-WNS approaches. It works particularly well for producing accurate predictions for time series data that show seasonality patterns. Although the decomposition approach is used in this study to remove the seasonal pattern from the reinsurance income data, it is important to keep in mind that, depending on the particulars of the data, the additive H-WNS damped method may be the better option in a different project or environment.

In the realm of actuarial sciences and insurance, the Single Exponential Smoothing (SEXS) model finds its place in a toolkit used for risk modeling. It plays a crucial role in predicting future risk factors, such as shifts in mortality rates, morbidity rates, or accident frequencies. These forecasts are of paramount importance for assessing the financial implications of various risks, estimating reserves, and determining appropriate insurance rates. In essence, SEXS serves as a valuable tool for insurers and actuaries in making informed decisions and managing financial stability in the ever-evolving landscape of insurance and risk management. Here are the steps involved in implementing the additive H-WNS algorithm:

- Establish the starting points for the seasonal, trend, and level components. This entails calculating the starting seasonal indices (S), initial trend (T), and initial level (L). While the initial seasonal indices can be produced by averaging the numbers for each season, the initial level and trend can be approximated by straightforward techniques such as regression or averaging.
- Smooth the time series data by adjusting the L, T, and S components. This process entails employing exponential smoothing methods to refine the estimates for each component.
- Use the updated L, T, and the S components to forecast future values.
- Update the estimates for the level, trend, and S components for the next iteration. Use the updated estimates obtained from the smoothing step to calculate the updated level, trend, and seasonal indices for the next time period.
- Repeat steps 2 to 4 for the desired number of iterations or until convergence is achieved. Convergence is typically determined by a predefined criterion, such as a small change in the estimated values.

The Box-Ljung test detects whether a time series contains an autocorrelation. H_0 (The data do not exhibit autocorrelation.) proposes an independent distribution for the residuals. The alternative hypothesis holds that the residuals are not independently distributed, but serially associated. The prediction accuracy is evaluated using the Box-Ljung test and the corresponding p-value. (see Ljung and Box (1978)).

3. Forecasting

Actuarial professionals use SEXS to forecast future claims volumes or claim amounts. By analyzing historical data on claims, the model can help estimate the expected number and severity of claims, allowing insurance companies to anticipate their financial obligations and set appropriate reserves. SEXS can be employed to analyze and forecast the development of losses over time. Actuaries can track the progression of reported losses, estimate ultimate claim amounts, and predict future loss development patterns. This information is crucial for determining adequate premium rates and assessing the financial stability of insurance portfolios.

To effectively apply the concept of seasonality in time series analysis, two distinct methodologies have been developed, each meticulously designed to address the unique characteristics of seasonal patterns inherent in the data. These methodologies offer tailored approaches to accommodate the varying nature of seasonality, ensuring that the chosen method aligns precisely with the specific seasonal fluctuations present in the data.

The selection between these methodologies' hinges on a critical consideration: the nature of the seasonal fluctuations within the dataset. The first methodology, known as the multiplicative technique, becomes the preferred choice when seasonal variations exhibit a proportional relationship with the overall level of the series. This proportionality signifies that the magnitude of seasonal changes varies in direct proportion to the underlying trend. In the multiplicative approach, the seasonal component is expressed as a factor relative to the series' scale. To mitigate the influence of seasonality, the level equation is employed to adjust the series. This adjustment involves subtracting the seasonal factor from the series. Consequently, within each year, this seasonal component tends to average out to approximately zero, permitting a clearer view of the underlying trend and patterns.

Conversely, when opting for the additive algorithm, the seasonal component is represented in absolute terms, making it more interpretable within the context of the observed series. In this scenario, the level equation is again utilized to account for seasonality, but instead of scaling it down proportionally, it subtracts the absolute seasonal component. In this approach, the seasonal component maintains its absolute values, which can be positive or negative, rather than being expressed as a proportion of the series' scale. While the aim remains consistent with the multiplicative method—to neutralize the impact of seasonal fluctuations within the data—this approach preserves the absolute magnitude of these fluctuations, allowing for more explicit interpretation.

As time series data can exhibit a wide array of patterns, it is often necessary to dissect a time series into various components, each revealing a distinct underlying pattern. This decomposition typically involves merging the trend and cycle components into a single entity referred to as the trend-cycle component, which simplifies the analysis. Consequently, a time series is partitioned into three primary segments: the trend-cycle component, a seasonal component, and a remainder component, which encapsulates the residual data within the time series. Conventional decomposition methods generally assume a yearly recurrence for the seasonal component, a valid assumption for many shorter time series but not necessarily applicable to longer ones. A case in point is the shifting patterns in electricity consumption, notably influenced by the widespread adoption of air conditioning. While current consumption patterns peak during the summer, a few decades ago, in numerous regions, high demand occurred both in the winter (for heating) and summer (for air cooling). Traditional decomposition techniques struggle to capture these dynamic seasonal variations over time.

This article delves into an in-depth analysis of reinsurance revenue within the American insurance industry, focusing on the financial performance of a reinsurance company. We have at our disposal a valuable monthly time series dataset that spans from February 2015 to April 2020, providing a comprehensive window into the dynamics of reinsurance revenue over this period (see Hamed et al. (2022), Hamedani et al. (2023), Hashempour et al. (2023), Salem et al. (2023), Mohamed et al. (2024), Tashkandy et al. (2023) and Yousof et al. (2023a,b,c) for more details). To initiate our analysis, our first step involves a meticulous examination of the reinsurance revenue dataset. We employ a combination of numerical and visual techniques to gain insights from the real-life data. One of the crucial aspects of this exploration is the utilization of various graphical methods to assess how well theoretical distributions, such as the beta, logistic, uniform, normal, exponential, lognormal, and Weibull distributions, align with the empirical data. Among these graphical tools, the skewness-kurtosis graph, also known as the Cullen-Frey graph (CFG), serves as an excellent representation of distribution characteristics. However, it primarily focuses on comparing distributions in terms of squared skewness and kurtosis (see Abdullah et al. (2023a, b), Alkhayyat et al. (2023), Elbatal et al. (2024), Minkah et al. (2023) for more other relevant datasets).

In addition to CFG, we employ the total time on test (TTS) graph to scrutinize the initial shape of the empirical hazard or failure rate function. This aids in understanding the patterns of event occurrences over time. Furthermore, we utilize the normal-quantile-quantile (NQQ) graph to assess the normality of the reinsurance revenue data, enabling us to gauge whether it adheres to a Gaussian distribution. In our arsenal of graphical tools, we also consider the "Kernel estimation" and "nonparametric Kernel density estimation (N-KDE) method." These techniques allow us to examine the intrinsic shape of our dataset, providing valuable insights into its underlying structure. Moreover, we employ the "box-graph" to identify extreme values within the reinsurance revenue dataset, helping us pinpoint outliers or significant fluctuations. To gain a deeper understanding of the temporal relationships within the data, we employ the autocorrelation function (AUCF). This function reveals how the correlation between any two data points changes as their temporal separation varies. By studying the theoretical AUCF, we can delve into the randomness and memory characteristics of the time series, especially at lag 1. However, it's important to note that this analysis primarily focuses on the distribution of peaks and troughs over time and doesn't provide insights into the frequency content of the underlying process. Our comprehensive analysis employs a range of numerical and graphical tools to scrutinize the reinsurance revenue data. We assess distribution fits, temporal patterns, normality, data shape, extreme values, and temporal correlations, all of which contribute to a holistic understanding of the dynamics within the American insurance industry's reinsurance revenue dataset.

Figure 1 shows the box-graph, CFG, NQQ graph, scattergram (fragmentary diagram), fitted fragmentary diagram, AUCF (under lag 1), partial AUCF (under lag $k = 1$), TTS graph, and N-KDE graph, respectively, for the original reinsurance revenues dataset. For the converted reinsurance revenues dataset, Figure 2 shows the box-graph, CFG, NQQ graph, fragmentary diagram, fitted fragmentary diagram, AUCF (under lag 1), partial AUCF (under lag 1), TTS graph, and N-KDE graph, in that order.

Figure 1 and Figure 2 provide valuable insights into the characteristics of both the original reinsurance revenues dataset and the converted reinsurance revenues dataset, offering a comprehensive assessment of their behaviors and statistical properties. In Figure 1, the top left and top right plots illustrate that neither the original reinsurance revenues dataset nor the converted reinsurance revenues dataset exhibit extreme revenue values. This absence of extreme values signifies a degree of stability in the data, suggesting that there are no outlier observations that significantly deviate from the overall pattern.

Furthermore, the top middle graph in Figure 1 highlights that neither dataset conforms to any of the theoretical distributions and beta distributions. This finding suggests that the observed revenue data does not align with these hypothesized distribution patterns, indicating a departure from classical distributional assumptions. Moving to the middle left and middle right plots in Figure 1, the original reinsurance revenues dataset is depicted as being randomly dispersed without any discernible pattern. This random dispersion implies that there is no evident structure or trend in the data, further emphasizing its stochastic nature. Additionally, the middle right plots in both Figure 1 and Figure 2 illustrate that the autocorrelation functions (AUCFs) under lag 1 exhibit exponential decay, indicating a diminishing level of correlation between adjacent data points. Similarly, Figure 2 reaffirms the observations made in Figure 1. The top middle graph in Figure 2 reiterates that neither the converted reinsurance revenues dataset nor the original reinsurance revenues dataset adheres to any of the theoretical distributions, reinforcing the departure from classical distributional assumptions.

The middle left and middle right plots in Figure 2 mirror those in Figure 1, indicating that the converted reinsurance revenues dataset is also randomly dispersed without any apparent pattern, emphasizing the stochastic nature of the data. Furthermore, the middle right plots in both Figure 1 and Figure 2 consistently demonstrate the exponential diminishment of AUCFs under lag 1, corroborating the decline in correlation between adjacent data points. In summary, both the original reinsurance revenues dataset and the converted reinsurance revenues dataset exhibit stability in terms of extreme values and depart from classical distributional assumptions. They display random dispersion patterns and demonstrate a diminishing correlation between adjacent data points, as evidenced by the AUCFs. These findings provide valuable insights into the statistical properties and behaviors of the revenue datasets, contributing to a deeper understanding of their underlying characteristics.

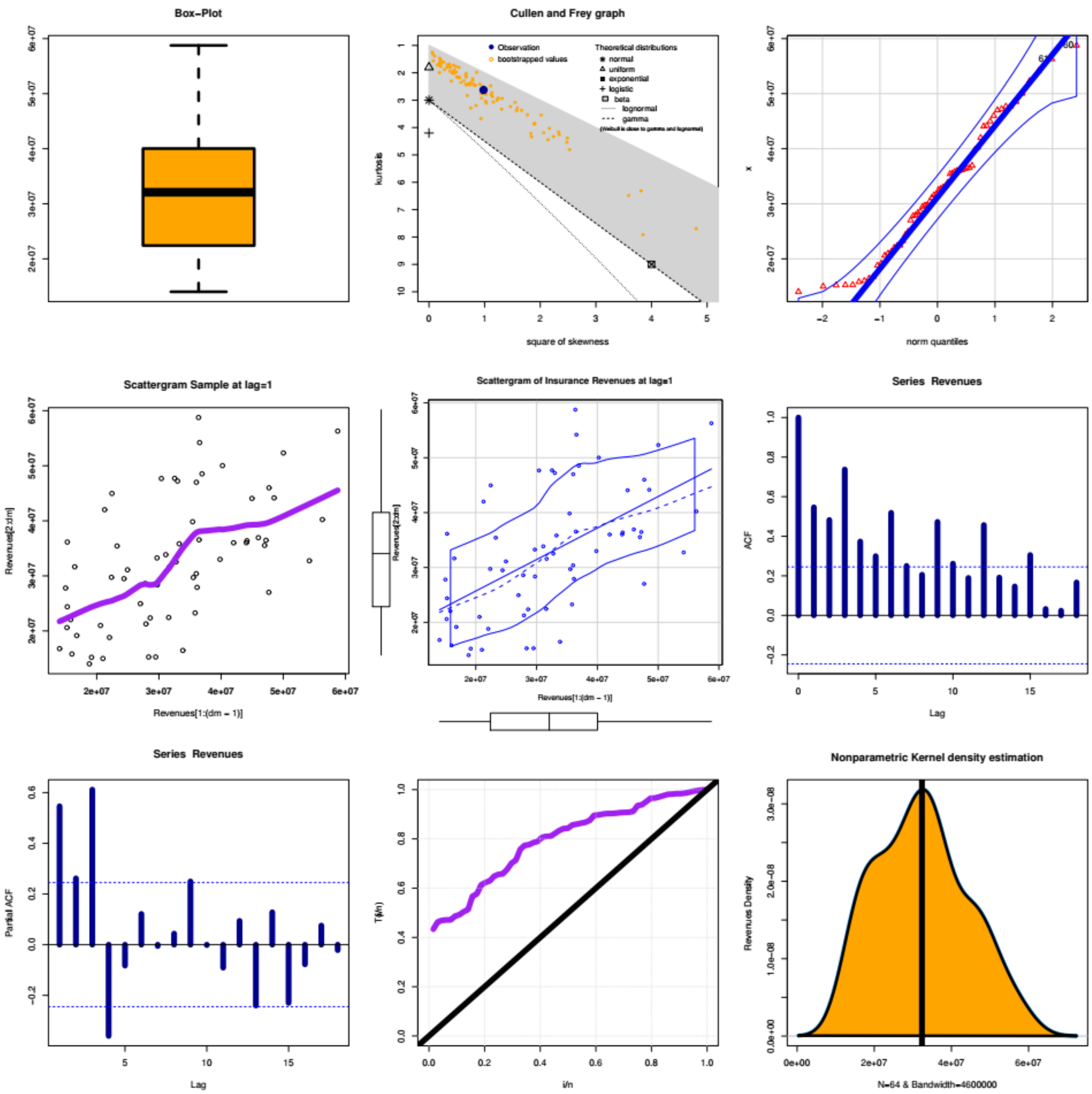

Figure 1: Graphical description for the original revenues data.

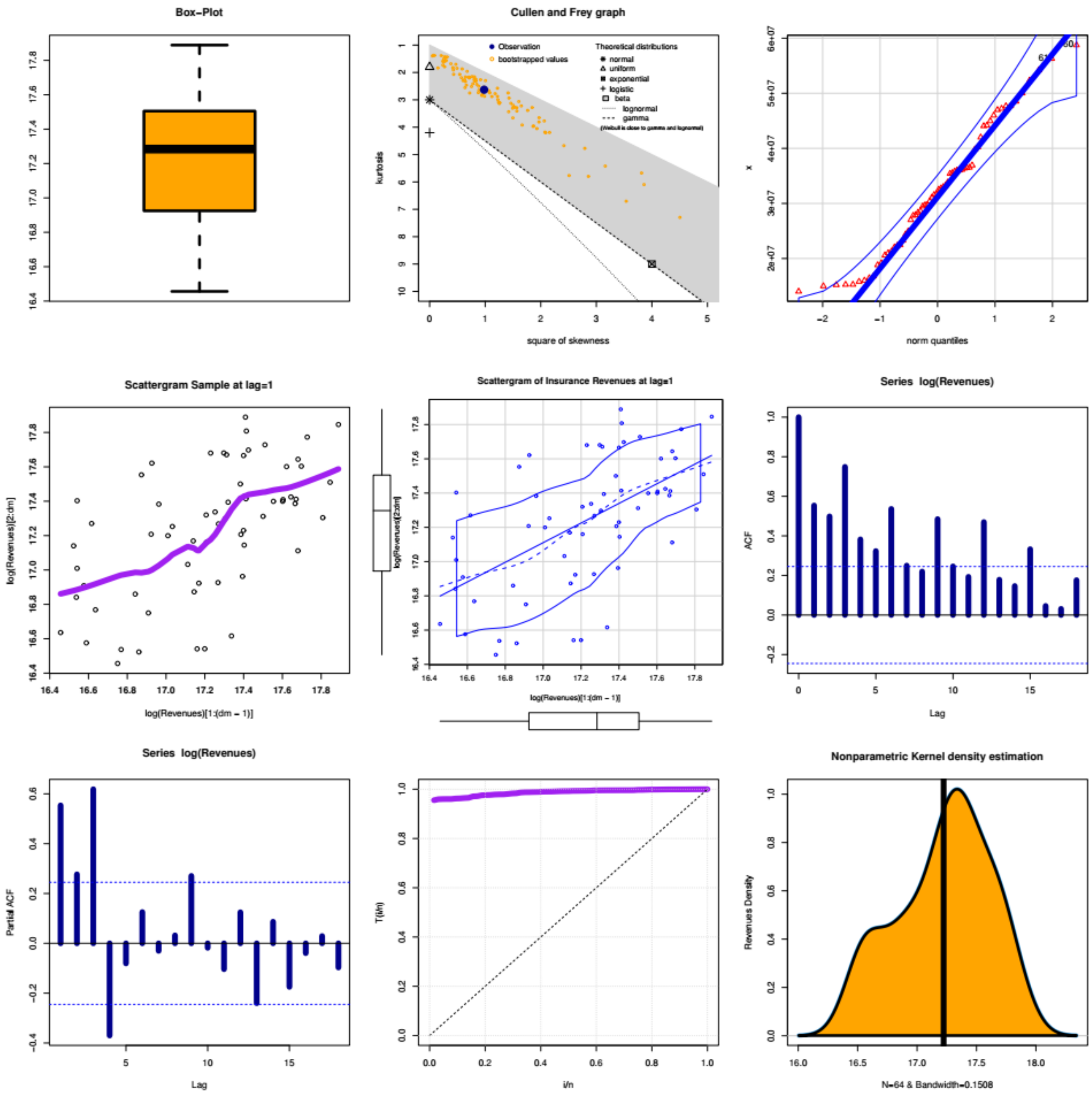

Figure 2: Graphical description for the converted revenues data

The partial AUCFs (under lag 1) are vanishing with frequency pattern, according to Figures 1 and 2 (bottom left plots). Additionally, Figure 1 The first value of lag can be proved to be significant (see the middle right plots), but none of the other coefficients of autocorrelation or partial ones for the other delays are. The hazard rate function (hzrf) for the initial reinsurance revenues dataset is monotonically growing, as shown in Figure 1 (bottom center graph). The hzrf for the converted reinsurance revenues dataset is similarly monotonically growing, according to Figure 2 (bottom middle graph). The density functions for the original and converted reinsurance revenues dataset are bimodal, as shown in Figures 1 and 2 (plots in the bottom right corner).

In Figure 3, we can observe the initial plots representing the reinsurance revenues dataset, with the original dataset displayed on the right panel and the transformed dataset on the left panel. These visualizations reveal a distinct seasonal pattern present in both datasets.

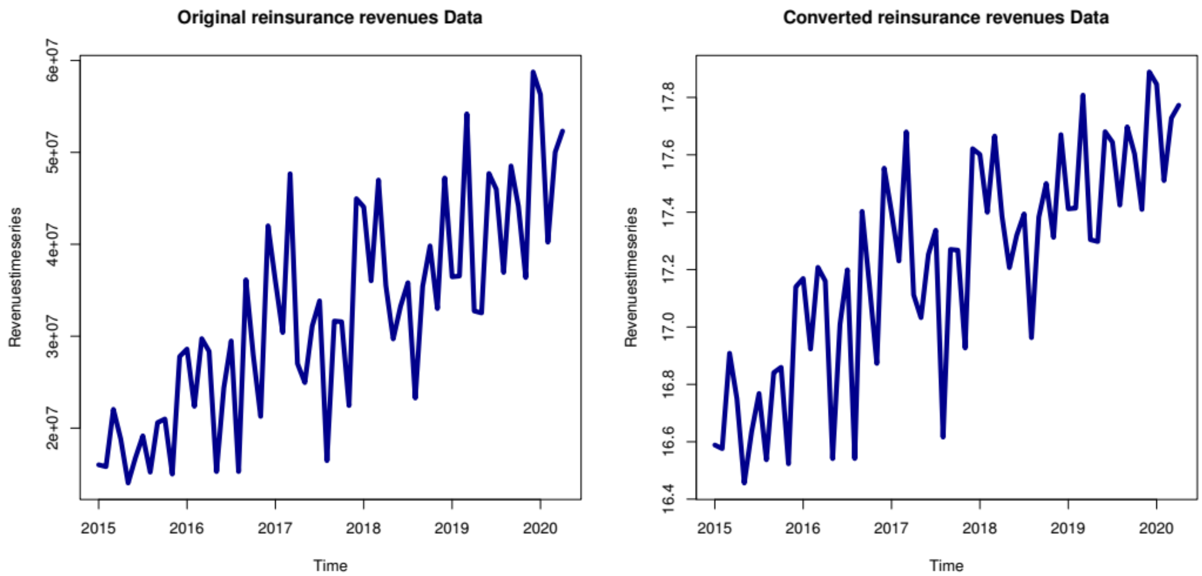

Figure 3: Initial time series plot for the original revenues (right plot) and converted revenues (left plot).

To effectively model and account for this seasonality, we employ a methodology known as seasonal adjustment within an additive model. The essence of this approach lies in estimating the seasonal component inherent in the data and subsequently eliminating it from the original time series. This adjustment process enables the additive model to accurately capture the underlying seasonal variations within the time series data. To achieve this, we leverage the "decompose ()" function, which provides an estimate of the seasonal component. This estimate is then used to separate the seasonal influence from the reinsurance revenue time series. As a result, we obtain a breakdown of the additive reinsurance revenue time series, effectively isolating the seasonal patterns and rendering the data more amenable to analysis and forecasting. This procedure enhances our ability to understand and model the inherent seasonality within the dataset, facilitating more accurate predictions and informed decision-making in the context of reinsurance revenue analysis.

Moving forward, Table 2 offers a complementary perspective by presenting the separated trend components for both the original and converted reinsurance revenues datasets. This table illuminates the underlying long-term trends and fluctuations within the datasets, offering valuable insights into the broader patterns and directional shifts in reinsurance revenue. Additionally, Table 3 complements the decomposition process by offering a comprehensive view of the separated random components for both the original and converted reinsurance revenues datasets. This table sheds light on the stochastic or irregular aspects of the data, aiding in the identification of unpredictable elements within the revenue datasets. To enhance the clarity of these findings and provide a visual representation of the decomposition process, we have incorporated decomposition charts in Figures 4 and 5. Figure 4 offers a detailed depiction of the decomposition plots for the initial reinsurance revenues dataset. These plots encompass observed revenues, trend components, seasonal variations, and random elements, offering a comprehensive view of how these components interact and contribute to the overall revenue pattern.

Figure 5, on the other hand, provides a parallel set of decomposition charts for the converted reinsurance revenues dataset. These charts offer a visual comparison between observed revenues and their corresponding trend, seasonal, and random components, allowing for a detailed examination of the decomposition process within this specific dataset. These tables and figures collectively serve as indispensable tools in unraveling the intricacies of the decomposition process applied to the reinsurance revenues datasets. They provide a comprehensive view of the seasonal, trend, and random components, enabling analysts to gain a deeper understanding of the underlying patterns and dynamics within the data.

Table 1: Seasonal components.

Time	Original data					
2015	4715247.02	-742028.390	12060973.98	-2212708.70	-7976948.32	-18955.4103
	1242600.57	-10733295.64	1872643.01	-214220.570	-7524548.14	9531240.59
2016	4715247.02	-742028.390	12060973.98	-2212708.70	-7976948.32	-18955.410
	1242600.57	-10733295.64	1872643.01	-214220.570	-7524548.14	9531240.59
2017	4715247.02	-742028.390	12060973.98	-2212708.70	-7976948.32	-18955.410
	1242600.57	-10733295.64	1872643.01	-214220.570	-7524548.14	9531240.59
2018	4715247.02	-742028.390	12060973.98	-2212708.70	-7976948.32	-18955.410
	1242600.57	-10733295.64	1872643.01	-214220.570	-7524548.14	9531240.59
2019	4715247.02	-742028.390	12060973.98	-2212708.70	-7976948.32	-18955.410
	1242600.57	-10733295.64	1872643.01	-214220.570	-7524548.14	9531240.59
2020	1242600.57	-10733295.64	1872643.01	-214220.570	-7524548.14	9531240.59
	1242600.57	-10733295.64	1872643.01	-214220.570	-----------	-----------
	Converted					
2015	0.175671531	0.003676427	0.334268406	-0.032238677	-0.270362010	0.008874969
	0.062403719	-0.408166948	0.079778219	0.020623802	-0.271303781	0.296774344
2016	0.175671531	0.003676427	0.334268406	-0.032238677	-0.270362010	0.008874969
	0.062403719	-0.408166948	0.079778219	0.020623802	-0.271303781	0.296774344
2017	0.175671531	0.003676427	0.334268406	-0.032238677	-0.270362010	0.008874969
	0.062403719	-0.408166948	0.079778219	0.020623802	-0.271303781	0.296774344
2018	0.175671531	0.003676427	0.334268406	-0.032238677	-0.270362010	0.008874969
	0.062403719	-0.408166948	0.079778219	0.020623802	-0.271303781	0.296774344
2019	0.175671531	0.003676427	0.334268406	-0.032238677	-0.270362010	0.008874969
	0.062403719	-0.408166948	0.079778219	0.020623802	-0.271303781	0.296774344
2020	0.175671531	0.003676427	0.334268406	-0.032238677	-0.270362010	0.008874969
	0.062403719	-0.408166948	0.079778219	0.020623802	-0.271303781	0.296774344

Table 2: Trend components.

Time	Original data					
2015	----------- ----------- ----------- ----------- ----------- ----------- ----------- ------					
	19044119	19842077	20435866	21153695	21602843	21971379
2016	22718310	23151217	23801309	24737116	25287155	26141365
	27042538	27685683	28767734	29459504	29807304	30490634
2017	30952217	31183336	31045570	31010826	31212133	31384062
	31842900	32413052	32618942	32945625	33498678	33786615
2018	33959008	34325328	34766470	35267834	36051673	36585360
	36362366	36067310	36389348	36573107	36574756	37294785
2019	38320378	39313320	40428645	41154962	41475990	42097259
	43404050	44383512	44362770	45004444	------	------
2020	----------- ----------- ----------- ----------- ----------- ----------- ----------- ------					
	Converted					
2015	----------- ----------- ----------- ----------- ----------- ----------- -----------					
	16.73961	16.77825	16.80518	16.83474	16.85536	16.87447
2016	16.90799	16.92613	16.94972	16.98501	17.01146	17.04325
	17.07006	17.09244	17.12489	17.14254	17.16100	17.19161

2017	17.20749	17.21635	17.21394	17.21354	17.22091	17.22597
	17.23722	17.25271	17.25918	17.27001	17.28870	17.29875
2018	17.30392	17.32072	17.33985	17.35423	17.37997	17.39806
	17.39221	17.38492	17.39146	17.39399	17.39438	17.41321
2019	17.43864	17.46830	17.50067	17.51808	17.52644	17.53960
	17.56680	17.58889	17.58955	17.60573	-----------	-----------
2020	-----------	-----------	-----------	-----------	-----------	-----------

Table 3: Random components.

Time	Original data					
2015	-----------	-----------	-----------	-----------	-----------	-----------
	-1124826.94	6096203.01	-1704569.52	53400.2300	915074.82	-3710811.510
2016	1168029.490	-42114.110	-6123674.50	5 826601.03	-2045603.08	-1736751.30
	1201378.940	-1682270.30	5500650.36	-1330139.91	-1010706.55	1992284.89
2017	339916.2300	-44532.610	4571586.42	-1784152.20	1713659.74	-263760.970
	763321.9800	-5225089.24	-2841492.68	-1159198.74	-3527758.98	1648270.59
2018	5393266.24	2436987.07	168546.8900	2481361.95	1624874.14	-3305339.08
	-1778431.44	-2065358.71	-2838501.39	3472678.96	3948937.70	395803.010
2019	-6575664.98	-2024793.36	1709088.16	-6198263.80	-967383.83	5631398.33
	1345491.190	3283448.97	2290846.95	-629806.82	-----------	------------------
2020	-----------	-----------	-----------	-----------	-----------	-----------
	Converted					
2015	-----------	-----------	-----------	-----------	-----------	-----------
	-0.033583719	0.167047781	-0.043969469	0.004330781	-0.060935802	-0.030989344
2016	0.085308885	-0.006703094	-0.076027990	0.207399927	-0.200045490	-0.042616219
	0.066971281	-0.142863469	0.198272615	-0.018487552	-0.016797469	0.065138573
2017	0.016067219	0.009820656	0.131774510	-0.069439240	0.081795344	0.017910031
	0.037785448	-0.228423052	-0.068710302	-0.022840885	-0.090751219	0.025890656
2018	0.121639719	0.075191073	-0.008551740	0.064079510	0.097034094	-0.087038719
	-0.060411219	-0.014129302	-0.088350719	0.085552031	0.188914615	-0.039609760
2019	-0.202585698	-0.057878510	-0.026764656	-0.181610073	0.041646177	0.132175031
	0.014775865	0.243905698	0.028295531	-0.023016719	-----------	----------- ------
2020	-----------	-----------	-----------	-----------	-----------	-----------

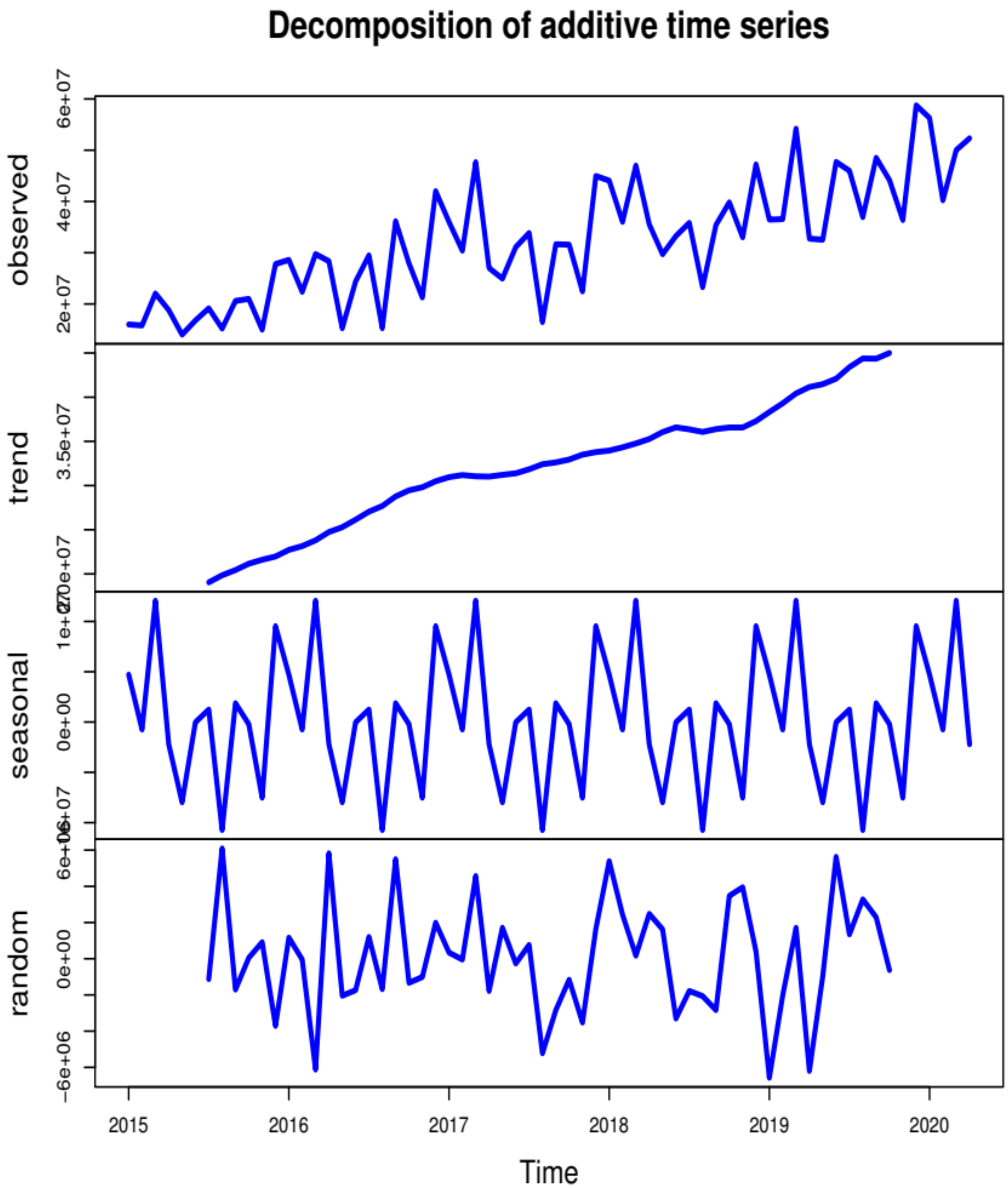

Figure 4: Decomposing the revenues dataset.

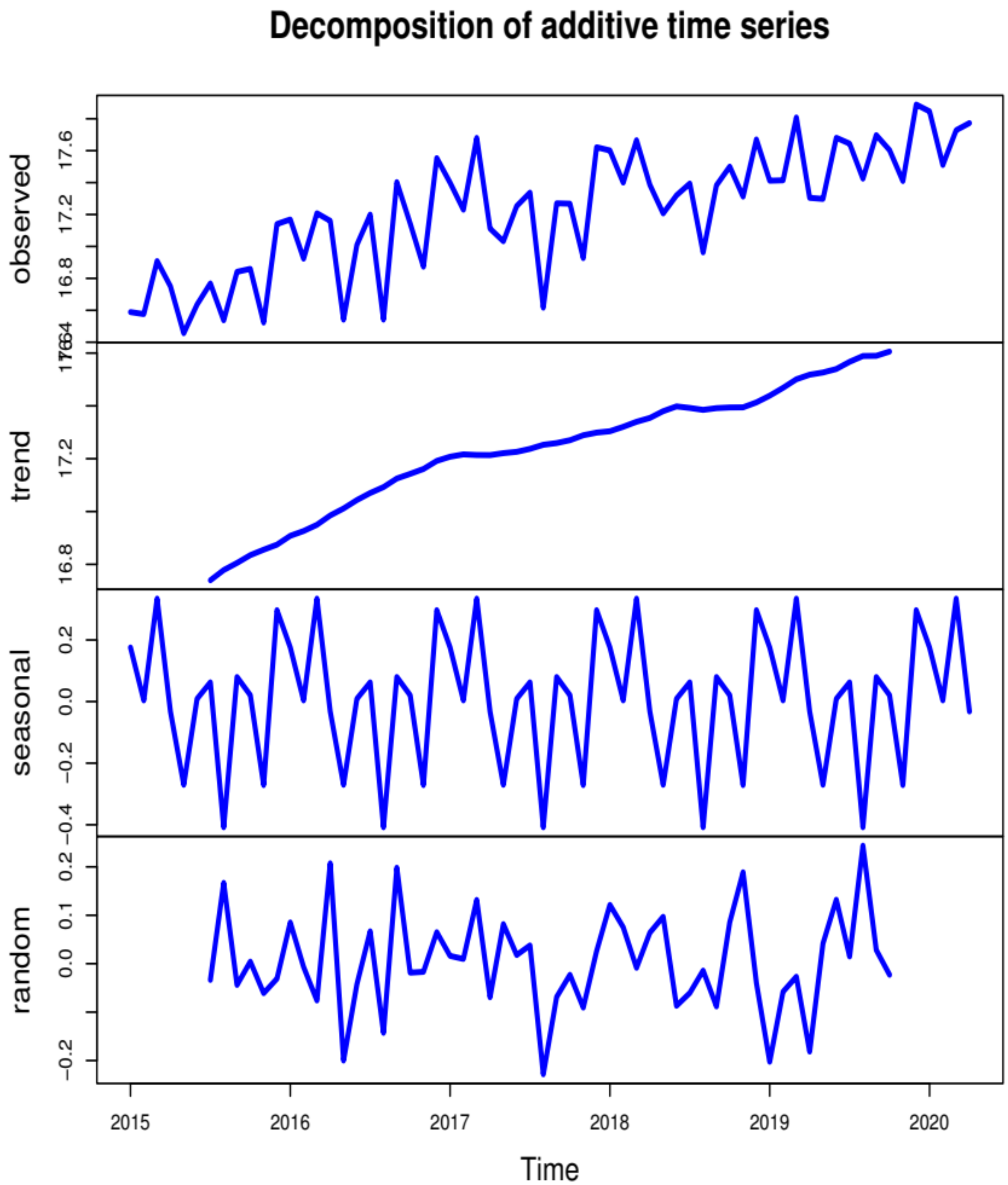

Figure 5: Decomposing the converted revenues dataset.

The decomposition process provides valuable insights into the various components that make up the reinsurance revenue datasets, shedding light on their individual characteristics and interactions. To present this decomposition effectively, we have organized the results into three distinct tables and complemented them with informative plots.

 DOI: http://dx.doi.org/10.18187/pjsor.v20i2.4409

Table 1 serves as a comprehensive reference, listing the separated components of the seasonal values for both the original reinsurance revenues dataset and the converted reinsurance revenues dataset. This tabulated data allows for a detailed examination of the seasonal patterns within each dataset, facilitating comparative analysis.

In Table 2, we delve into the trend components of both the original and converted reinsurance revenues datasets. This table offers a breakdown of the underlying trends, providing valuable insights into the long-term patterns and fluctuations in reinsurance revenue over the given time frame.

Table 3 complements the decomposition process by presenting the random components after separating for both the original and converted reinsurance revenues datasets. This table unveils the stochastic or irregular elements within the data, helping us understand the unpredictability inherent in reinsurance revenues.

To enhance the clarity of these findings, we have incorporated visualization in the form of decomposition plots. Figure 4 offers a detailed depiction of the decomposition plots, showcasing the observed revenues alongside their trend, seasonal, and random components for the original reinsurance revenues dataset. This graphical representation aids in visually understanding the interplay of these components.

Similarly, Figure 5 provides a parallel set of decomposition plots for the converted reinsurance revenues dataset. These plots offer a visual comparison between the observed revenues and their corresponding trend, seasonal, and random components, facilitating a comprehensive understanding of the data's decomposition.

In Figure 6, we take a step further by presenting seasonally adjusted plots, trend-adjusted plots, and random-adjusted plots for both the original and converted reinsurance revenues datasets. These adjusted plots offer a refined view of the data, highlighting specific components and patterns, making it easier to discern and analyze.

To conclude, Table 4 offers a comprehensive assessment of the Single Exponential Smoothing (SEXS) model for the years 2015 and 2016. This assessment consolidates the model's performance metrics and findings for these specific years, providing a concise summary of its predictive capabilities.

Collectively, these tables, plots, and assessments offer a holistic view of the decomposition process and its implications for the reinsurance revenue datasets, enabling a thorough analysis of the underlying trends, seasonality, and randomness within the data.

Table 4: Assessing the SES model for 2015 and 2016.

	2015	
Jan	-----------	---------
Feb	16010072	16.58873
Mar	15962607	16.58566
Apr	17388895	16.66365
May	17723093	16.68454
Jun	16855391	16.62938
Jul	16838640	16.63096
Aug	17383238	16.66415
Sep	16872629	16.63348
Oct	17747293	16.68359
Nov	18508096	16.72611
Dec	17684202	16.67710
	2016	
Jan	20053545	16.78892
Feb	22057307	16.88069
Mar	22129920	16.89093
Apr	23913487	16.96747
May	24953694	17.01400
Jun	22682457	16.89981

Jul	23081707	16.92629
Aug	24583071	16.99225
Sep	22400004	16.88339
Oct	25621063	17.00884
Nov	26158822	17.04164
Dec	25013305	17.00089

Table 5: Assessing the SES model.

	2017	
Jan	28998507	17.13433
Feb	30641470	17.19829
Mar	30584111	17.20591
Apr	34591151	17.32037
May	32814969	17.27003
Jun	30971056	17.21264
Jul	31001598	17.22233
Aug	31669021	17.25011
Sep	28102596	17.09703
Oct	28934171	17.13886
Nov	29552558	17.16999
Dec	27886783	17.11123
	2018	
Jan	31890384	17.23442
Feb	34744849	17.32298
Mar	35043827	17.34148
Apr	37845555	17.41973
May	37304283	17.41160
Jun	35521655	17.36212
Jul	34991746	17.35192
Aug	35187430	17.36213
Sep	32393528	17.26567
Oct	33103787	17.29397
Nov	34680858	17.34376
Dec	34286645	17.33609

Table 6: Assessing the SES model (continued).

	2019	
Jan	37318805	17.41680
Feb	37117482	17.41558
Mar	36983636	17.41522
Apr	41019053	17.51010
May	39079280	17.46039
Jun	37544440	17.42111
Jul	39927297	17.48378
Aug	41348969	17.52246

Sep	40313969	17.49884
Oct	42239028	17.54683
Nov	42689424	17.56048
Dec	41209238	17.52400
	2020	
Jan	45322518	17.61211
Feb	47893027	17.66858
Mar	46095608	17.63029
Apr	47016039	17.65388
π	**0.2344118**	**0.2414504**
MSE	**868.3633**	**864.2555**

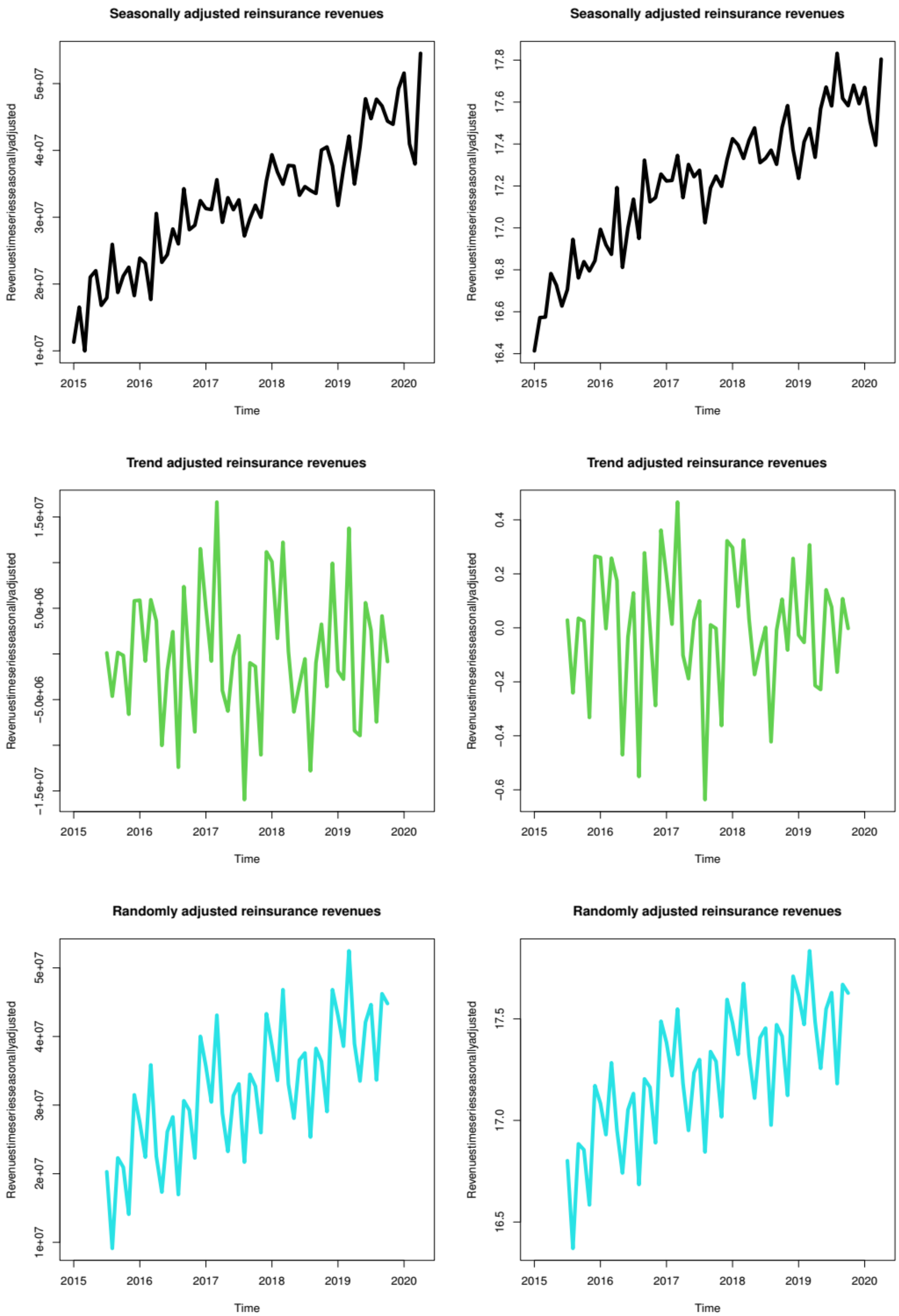

Figure 6: Adjusted seasonally graph under the main revenues (the right plots) and revenues after converting (the left plot).

Up to this point, we have explained the reinsurance revenue data and used the decomposition process to remove the seasonal components. We haven't yet thought about how well a particular model might describe our work, though. In this section, we look at ways to select the "optimal" model to describe the study system as well as techniques for assessing the additive H-WNS' method's forecasting performance. To do this, we will use the H-WNS filtering plots to re-predict all of the reinsurance revenue numbers as if we were unaware of their existence. So, by evaluating the discrepancy between actual observations and forecasts generated using the H-WNS approach for the same months, we may evaluate the H-WNS method. Figure 7 illustrates the H-WNS filtering plots for both the original (on the right) and converted (on the left) reinsurance revenue data. This visual representation serves to showcase the model's accuracy in predicting the values of reinsurance revenue.

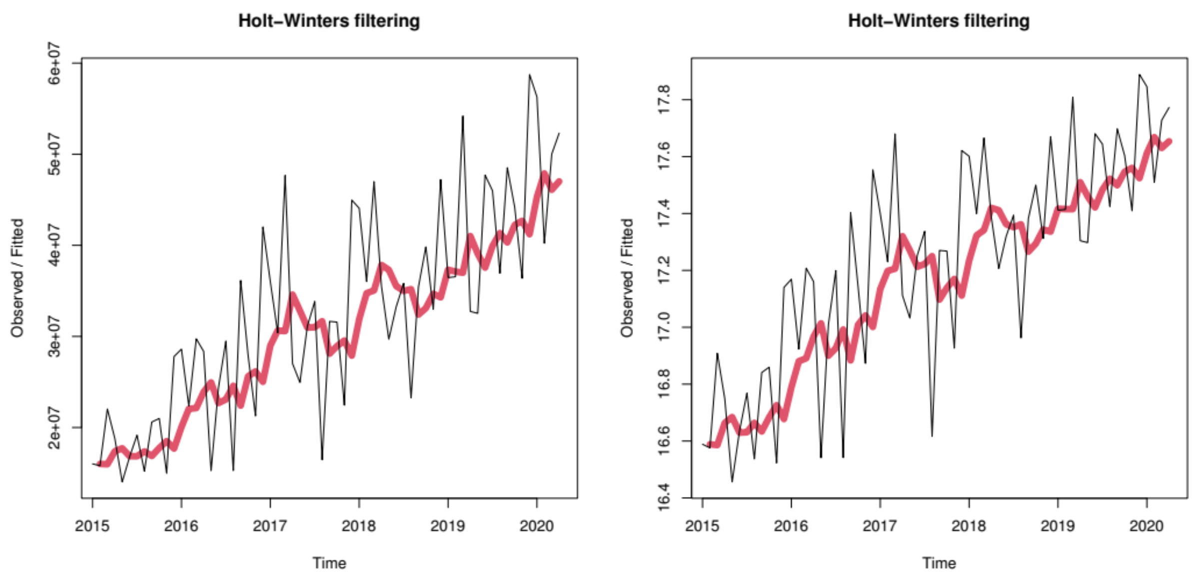

Figure 7: Filtering under H-WNS for the original (right panel) and converted (left panel) under the revenues data.

The SEXS model serves as a valuable tool for generating short-term forecasts, particularly when dealing with time series data that adhere to an additive model characterized by a constant level and an absence of seasonality. Within this context, the simple exponential smoothing method plays a pivotal role in estimating the current level at a given time point. To control the smoothing process, the model relies on a parameter with values ranging from zero to one. The parameter's value effectively determines the weight assigned to the most recent data when predicting future values. When the parameter approaches zero, it implies that recent data points are accorded relatively less significance in forecasting future values. To assess the performance of the SEXS model, we provide a detailed evaluation in Table 4, Table 5, and Table 6. These tables present the projected values generated by the model, facilitating a comprehensive examination of its forecasting capabilities. Table 5 specifically focuses on the assessment of the SEXS model for the years 2017 and 2018, while Table 6 extends this evaluation to cover the years 2019 and 2020. Within these tables, you will find critical information regarding the estimated parameter values for both the original and transformed data. For the original data, the estimated parameter stands at 0.2344118, while for the translated data, it is calculated to be 0.2414504.

Furthermore, the SSE serves as an essential metric for evaluating model performance. For the original data, the SSE amounts to 868.3633, while for the transformed data, it is slightly lower at 864.2555. This discrepancy suggests that employing the transformed data may be a more preferable choice, indicating improved model performance. In summary, the SEXS model proves to be a valuable tool for short-term forecasting in cases where time series data conform to an additive model with a constant level and no seasonality. Through the assessment of parameter values and SSE, we can determine the suitability of the model for different datasets, ultimately guiding us towards making more informed decisions regarding forecasting strategies.

Using the additive H-WNS algorithm, we may forecast using a SEXS predictive model. The forecast HoltWinters() function provides an annual forecast as well as 80% and 95% prediction intervals. The projected values for evaluating the SEXS model for two subsequent years are presented in Table 7. The H-WNS forecasting plots for the original (right) and converted (left) reinsurance revenue data up to the year of 2022 are also shown in Figure 8. Figures 9 and 10 show the AUCFs for the original and converted reinsurance revenues dataset, respectively. These AUCFs guarantee the predictive accuracy of the additive H-WNS model.

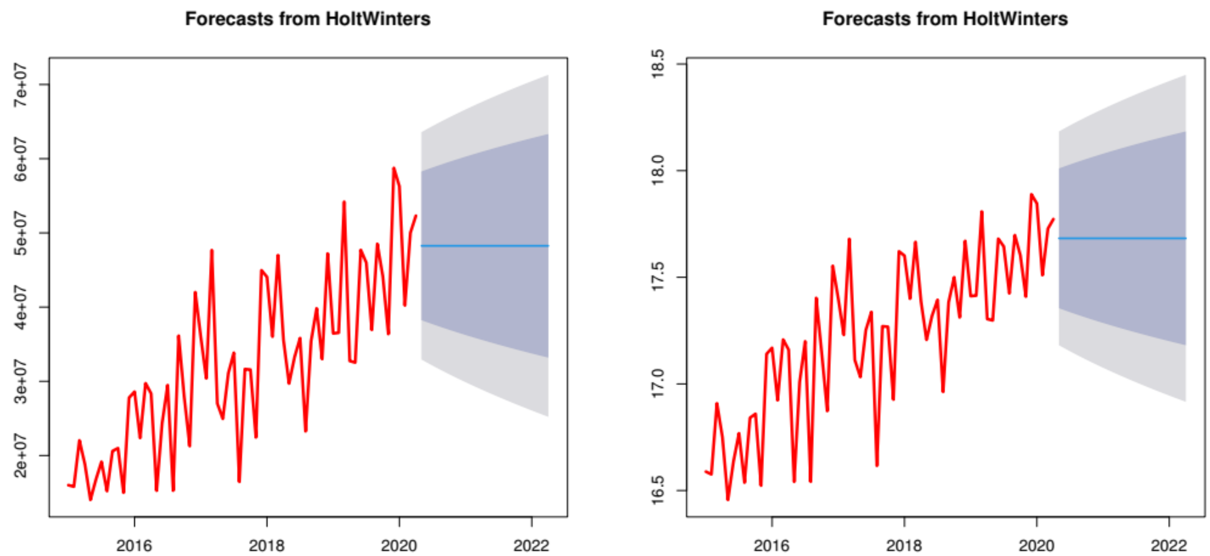

Figure 8: Forecasting under H-WNS for the original revenues (right) and converted revenues (left).

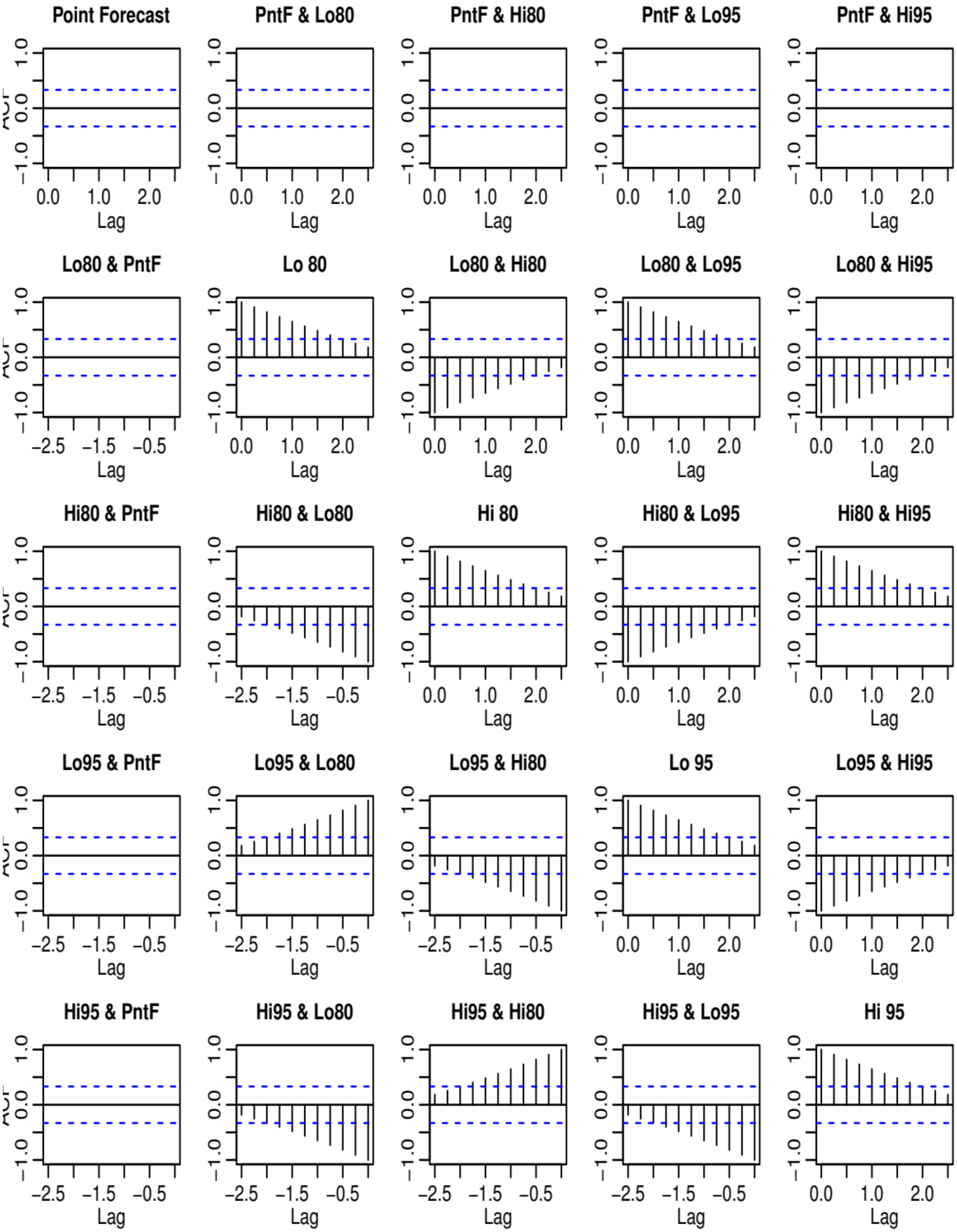

Figure 9: AUCFs analysis under original revenues.

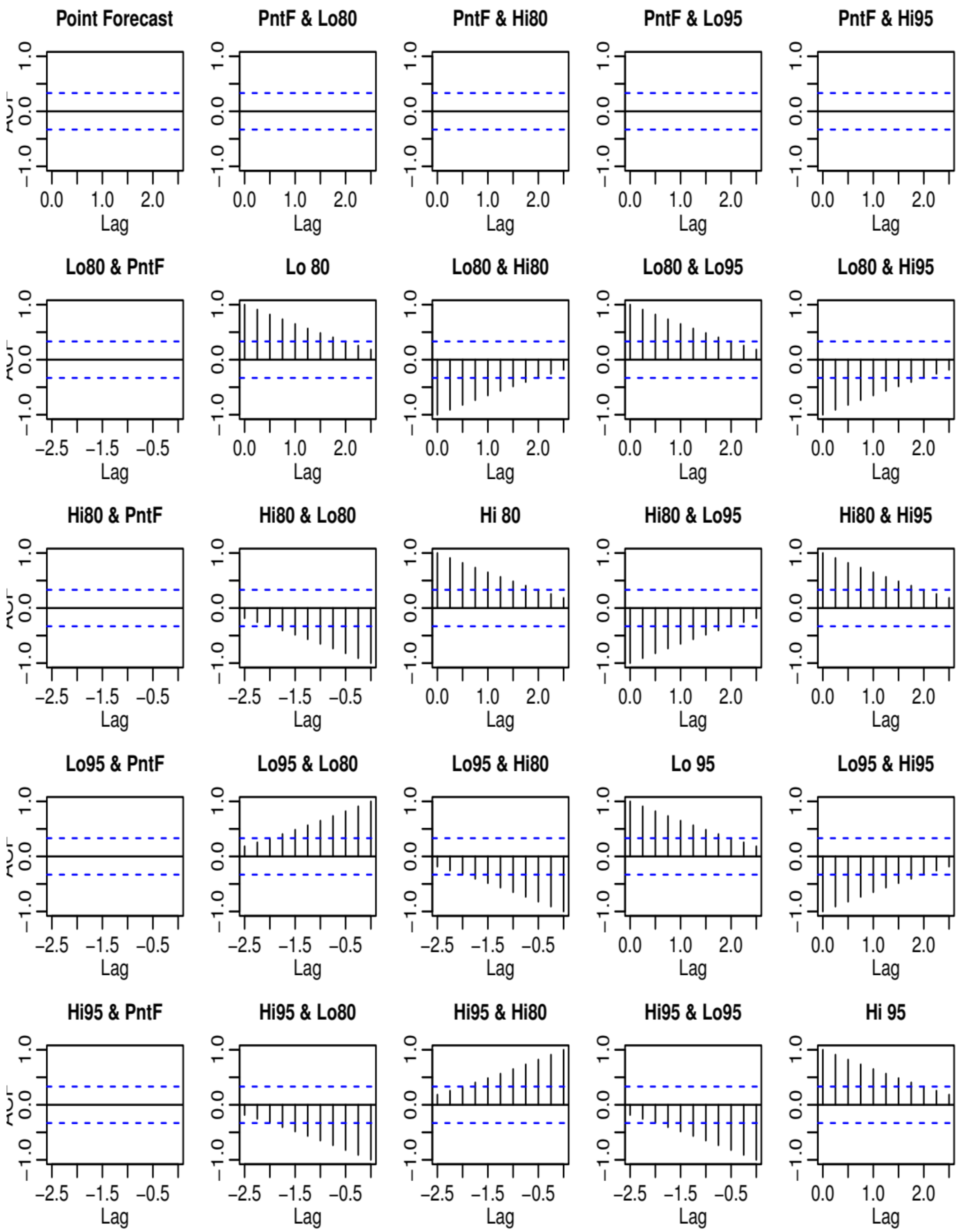

Figure 10: AUCFs under the converted revenues.

Ensuring the reliability and effectiveness of a predictive model is a paramount concern in the field of reinsurance revenue analysis. To achieve this, it is essential to conduct a rigorous assessment of the forecast errors generated by the model. This assessment serves as a critical step in evaluating the model's performance and ascertaining whether it has reached its optimal state, leaving minimal room for further refinement. One of the key aspects in assessing forecast errors is to determine whether they exhibit consistent patterns. This entails scrutinizing whether these errors have an average value close to zero and whether their variance remains stable over time. These characteristics are indicative of a well-calibrated and reliable predictive model.

A valuable technique employed in this evaluation is the creation of a temporal graph, as demonstrated in Figure 11. This graphical representation provides a dynamic and visual perspective on the behavior of forecast errors within the in-sample dataset across time. By closely examining this graph, analysts can gain insights into whether the variance of in-sample forecast errors exhibits a consistent and steady pattern throughout the observed period. Such stability in variance is a positive indicator, suggesting that the model's performance remains robust over time. In addition to graphical assessments, statistical measures play a pivotal role in gaining deeper insights into the nature of forecast errors. When analyzing the original reinsurance sales data, we observe specific statistical indicators. For instance, a p-value of 0.70634 and a Box-Ljung test statistic of 23.4914 provide valuable information. These findings suggest the presence of non-zero autocorrelations within the in-sample forecast errors at various lag intervals, particularly at lags 1 to 20. This autocorrelation analysis is essential for detecting potential temporal dependencies and patterns within the forecast errors, thereby offering guidance for refining and optimizing the predictive model.

Similarly, when examining the converted reinsurance revenues dataset, we encounter another set of statistical results, including a Box-Ljung test statistic of 22.6911 and a p-value of 0.73665. These findings echo the presence of non-zero autocorrelations within the in-sample forecast errors at similar lag intervals, further emphasizing the importance of considering temporal dependencies. In conclusion, the meticulous assessment of forecast errors is a critical element in the evaluation of a predictive model's performance and stability. This multifaceted analysis, which combines graphical representations and statistical evaluations, as demonstrated through the temporal graph and Box-Ljung tests, provides a comprehensive understanding of how forecast errors behave over time. These insights are instrumental in ensuring the reliability and accuracy of the predictive model, facilitating informed decision-making in the context of reinsurance revenue analysis and management.

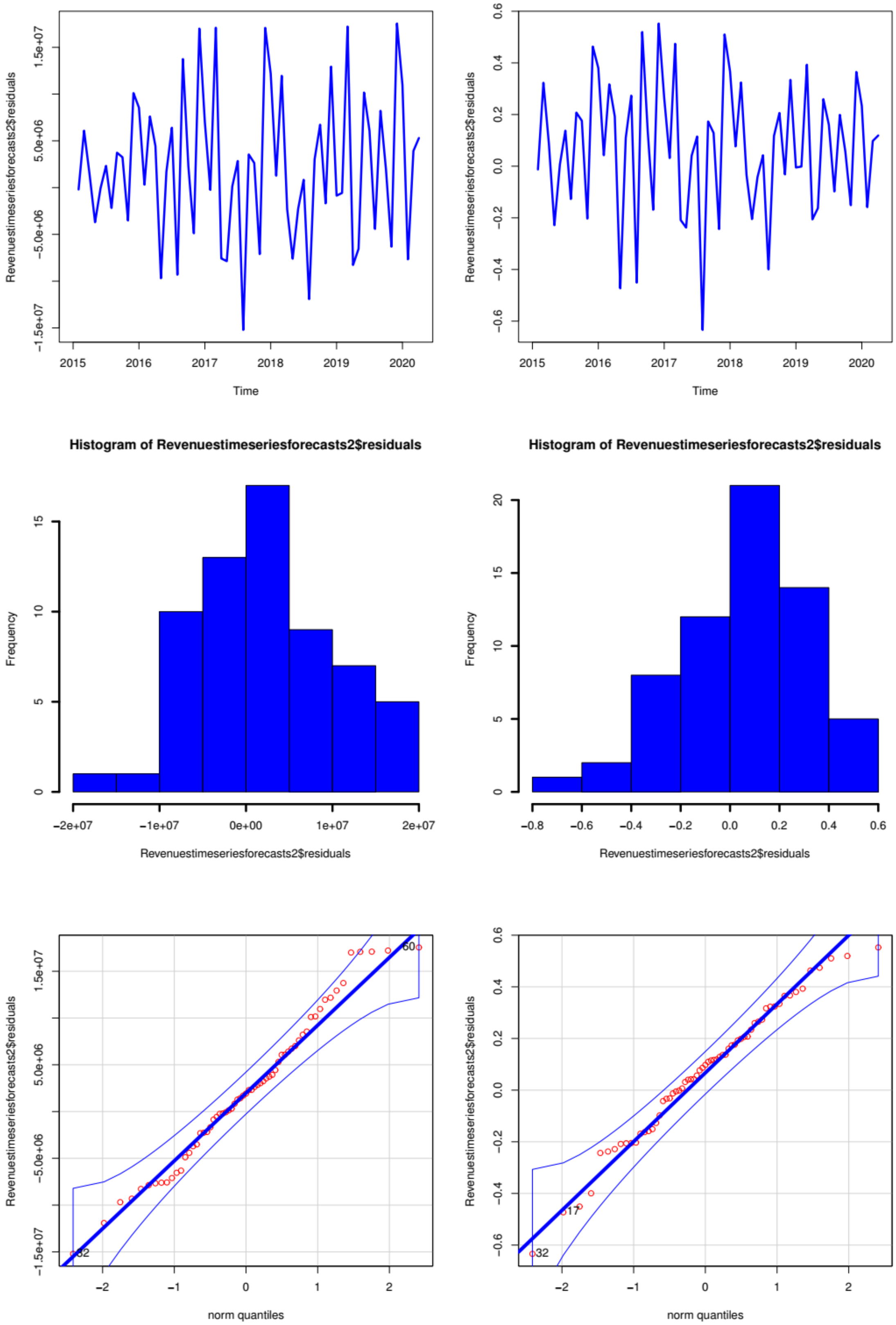

Figure 11: Plots for describing the residuals for the original (right) and converted (left) revenues data.

In essence, these assessments help us gauge the quality and reliability of the predictive model, particularly with regards to the persistence of autocorrelations in the forecast errors over various lag periods.

Table 7 provides detailed insights into predictive confidence intervals for the years 2020 through September 2022, spanning from May 2020 onwards. These intervals are delineated for two confidence levels: 80% and 95%, with lower and upper bounds denoted as LO-B and UP-B, respectively. A noteworthy observation from the data is the trend where the width of the 80% confidence intervals tends to exceed that of the 95% confidence intervals across all months. This discrepancy in interval width can be attributed to the nature of statistical forecasting. As the time range for future projections extends further into the future, the inherent uncertainty and variability in the data tend to amplify, consequently widening the confidence intervals. Therefore, it is not uncommon for the 80% confidence intervals to be wider than their 95% counterparts due to the increased forecast inaccuracy associated with longer time horizons.

Moreover, the preference for the 95% confidence intervals over the 80% intervals is justified by their higher level of confidence. The 95% confidence intervals offer a greater degree of certainty in capturing the true value of the forecasted variable within the specified range. This preference stems from the understanding that a higher confidence level provides a more conservative estimate of the forecast uncertainty, thereby offering a more reliable basis for decision-making. The observed increase in prediction error and widening of confidence intervals with the statistical forecast's time horizon corroborates with the general principles of forecasting. As the forecast horizon extends further into the future, the accuracy of predictions tends to diminish due to the accumulation of various sources of uncertainty and unpredictability. Therefore, it becomes imperative for analysts to consider and account for the increasing levels of uncertainty associated with longer-term forecasts when interpreting and utilizing the forecast results. So, the data presented in Table 7 highlights the dynamic interplay between forecast accuracy, confidence levels, and forecast horizon. By acknowledging and understanding these nuances, analysts can make informed decisions and adjustments to their forecasting methodologies to better account for and mitigate the inherent uncertainties associated with longer-term predictions.

Table 7: Predictive confidence intervals, Box-Ljung and its p-value.

2020		
	(Original)	**(Converted)**
	80% (LO-B, UP-B), 95% (LO-B, UP-B)	**80%(LO-B, UP-B), 95%(LO-B, UP-B)**
May	(38247556, 58271470), (32947546, 63571480)	(17.35505 18.01019), (17.18164 18.18359)
Jun	(37976161, 58542865), (32532484, 63986542)	(17.34563 18.01960), (17.16725 18.19799)
Jul	(37711747, 58807279), (32128098, 64390928)	(17.33648 18.02876), (17.15324 18.21199)
Aug	(37453802, 59065224), (31733604, 64785423)	(17.32756 18.03768), (17.13960 18.22564)
Sep	(37201871, 59317155), (31348309, 65170717)	(17.31885 18.04638), (17.12629 18.23895)
Oct	(36955555, 59563472), (30971600, 65547426)	(17.31035 18.05488), (17.11329 18.25194)
Nov	(36714492, 59804534), (30602927, 65916099)	(17.30204 18.06319), (17.10058 18.26465)
Dec	(36478360, 60040666), (30241795, 66277231)	(17.29391 18.07132), (17.08814 18.27709)
2021		
Jan	(36246870 60272156), (29887761 66631265)	(17.28595 18.07929), (17.07596 18.28927
Feb	(36019757 60499269), (29540421 66978605)	(17.27814 18.08709), (17.06402 18.30121
Mar	(35796782 60722244), (29199410 67319616)	(17.27048 18.09476), (17.05231 18.31293
Apr	(35577726 60941300), (28864394 67654632)	(17.26296 18.10228), (17.04080 18.32443)
May	(35362391 61156635), (28535067 67983959)	(17.25557 18.10966), (17.02951 18.33573)
Jun	(35150593 61368433), (28211149 68307877)	(1 17.24831 18.11693), (17.01840 18.3468)
Jul	(34942162 61576864), (27892382 68626644)	(17.24116 18.12407), (17.00747 18.35776)
Aug	(34736944 61782082), (27578528 68940498)	(17.23414 18.13110), (16.99672 18.36851)
Sep	(34534794 61984232), (27269367 69249659)	(17.22721 18.13802), (16.98614 18.37909)

Oct	(34335579 62183448), (26964693 69554333)	(17.22040 18.14484 16.97571 18.38952)
Nov	(34139173 62379853), (26664317 69854709)	(17.21368 18.15155), (16.96544 18.39979)
Dec	(33945463 62573563), (26368062 70150964)	(17.20706 18.15818), (16.95531 18.40992)
	2022	
Jan	(33754339 62764687), (26075764 70443262)	(17.20052 18.16471), (16.94532 18.41991)
Feb	(33565701 62953325), (25787267 70731759)	(17.19408 18.17115), (16.93546 18.42977)
Mar	(33379454 63139572), (25502427 71016599)	(17.18772 18.17752), (16.92574 18.43950)
Apr	(33195510 63323516), (25221108 71297918)	(17.18144 18.18380), (16.91613 18.44910)
May	(33013785 63505241), (24943184 71575842)	(17.17524 18.19000), (16.90664 18.45859)
Jun	(32834200 63684826), (24668533 71850493)	(17.16911 18.19613), (16.89727 18.46796v
Jul	(32656683 63862343), (24397044 72121982)	(17.16305 18.20218), (16.88801 18.47722)
Aug	(32481162 64037864), (24128608 72390418)	(17.15707 18.20817), (16.87886 18.48638)
Sep	(32307573 64211453), (23863126 72655900)	(17.15115 18.21408), (16.86981 18.49543)
	Box-Ljung **23.4914**	**22.6911**
	p-value **0.70634**	**0.73665**

A few observations about Table 7:

I. The confidence intervals seem to be converted from an original format to another format, possibly for easier interpretation or different statistical analysis.

II. The Box-Ljung statistics and p-values are likely related to some form of time series analysis or autocorrelation assessment. These statistics are commonly used to assess whether there is significant autocorrelation remaining in the residuals of a time series model.

III. The p-values provided indicate the significance of the autocorrelation. A higher p-value (close to 1) suggests that there is no significant autocorrelation remaining in the residuals, while a lower p-value (close to 0) suggests the presence of significant autocorrelation.

4. Concluding remarks discussion and future research points

SEXS is a widely employed forecasting technique within the realm of time series analysis, serving as a valuable tool for generating short-term predictions or forecasts based on historical data. Its fundamental premise lies in assigning weights to past observations, with a crucial twist – these weights decrease exponentially as we move further back in time, thereby giving greater importance to more recent data points. The forecast for the upcoming period is a calculated blend of the current observation and the previous forecast, with the assigned weights determining the significance of each component in shaping the prediction. This method finds its true efficacy when dealing with time series data that can be effectively modeled using an additive approach characterized by a constant level and an absence of seasonality. It excels at short-term forecasting, making it an invaluable choice for scenarios where the need is to anticipate trends and variations in the near future without being confounded by complex seasonal patterns or long-term trends. For reinsurance companies, having a grasp of future values for projected reinsurance revenues is imperative to preempt substantial financial losses stemming from potential claims. Time series data plays a pivotal role in this context, providing the foundation for informed decision-making and risk management. In our study, we delve into the application of the additive H-WNS' method for forecasting reinsurance revenues, a technique renowned for its ability to capture seasonal and trend components. To achieve this, a decomposition process is applied to remove the seasonality components, allowing for a clearer focus on the underlying trends and patterns. The effectiveness of this proposed approach is rigorously assessed through the application of additive H-WNS' filtering and comprehensive testing investigations. A novel time series dataset is employed to showcase the practical utility of our paradigm, demonstrating its prowess in predicting reinsurance revenues accurately. Furthermore, we employ the Box-Ljung test and its corresponding p-value as valuable tools for evaluating the predictive accuracy of our model, ensuring its robustness in real-life applications. Additionally, our analysis extends to residual analysis, which encompasses both point and interval predictions for reinsurance revenue data. This comprehensive assessment enhances our understanding of the model's performance and allows us to fine-tune its predictive capabilities. Overall, our study underscores the significance of time series analysis and forecasting in the context of reinsurance revenue management, offering a robust and effective framework for ensuring financial stability and informed decision-making in this critical industry.

The application of SEXS, particularly within the framework of the additive H-WNS' method, has gained significant prominence in the realm of time series analysis. This statistical technique has proven to be invaluable for generating short-term forecasts based on historical data, and its fundamental premise revolves around assigning exponentially decreasing weights to past observations. This approach gives precedence to recent data points, reflecting the idea that more recent observations are often more indicative of future trends and patterns. SEXS stands out as an effective choice in scenarios where the underlying time series data can be effectively modeled using an additive approach, characterized by a constant level and the absence of seasonality. Its strength lies in its ability to excel at short-term forecasting, making it particularly suitable for predicting trends and variations in the near future without being confounded by complex seasonal patterns or long-term trends. In the context of reinsurance companies, where the anticipation of future values for projected reinsurance revenues is pivotal, SEXS offers a practical and efficient solution.

Time series data plays a crucial role in the decision-making processes of reinsurance companies, as it forms the foundation for risk assessment and financial management. In the study at hand, we delve into the application of the additive H-WNS' method for forecasting reinsurance revenues, a technique renowned for its capability to capture both seasonal and trend components. The decomposition process employed in this method effectively removes the seasonality components, allowing for a clearer focus on the underlying trends and patterns that drive reinsurance revenues. To rigorously assess the effectiveness of our proposed approach, we apply additive H-WNS' filtering and conduct comprehensive testing investigations. These efforts aim to ensure that our model can accurately predict reinsurance revenues. A novel time series dataset is utilized to demonstrate the practical utility of our paradigm, showcasing its prowess in capturing and forecasting reinsurance revenue patterns. Furthermore, we employ statistical tools such as the Box-Ljung test and its corresponding p-value to evaluate the predictive accuracy of our model. These tests provide valuable insights into the robustness of our approach when applied in real-life scenarios, where the consequences of inaccurate forecasts can have significant financial implications. In addition to assessing the overall performance of our model, we conduct residual analysis, which encompasses both point and interval predictions for reinsurance revenue data. This comprehensive evaluation enhances our understanding of the model's strengths and weaknesses, allowing us to fine-tune its predictive capabilities and make informed decisions based on the uncertainty inherent in the forecasts.

In conclusion, our study highlights the critical role of time series analysis and forecasting in the context of reinsurance revenue management. It provides a robust and effective framework for ensuring financial stability and informed decision-making within this vital industry. By applying the additive H-WNS' method with decomposition and thorough testing, we contribute to the growing body of knowledge aimed at enhancing the accuracy and reliability of reinsurance revenue forecasts, ultimately assisting reinsurance companies in managing risk and achieving their financial goals.

Certainly, here are few potential future statistical points for research related this work:

- I. Conduct a comparative study to assess the forecasting accuracy of the Single-Exponential Smoothing model under the additive H-WNS Algorithm with Decomposition against other time series forecasting methods commonly used in the insurance and reinsurance industry, such as ARIMA, GARCH, or neural networks.
- II. Investigate the sensitivity of the model's forecasting performance to the choice of smoothing parameters and assess whether fine-tuning these parameters can lead to improved accuracy.
- III. Explore the applicability of the proposed model to different data frequencies (e.g., monthly, quarterly, annual) and evaluate its performance under various time intervals.
- IV. Develop techniques to handle outliers and anomalies in reinsurance revenue data, as these can significantly impact forecasting accuracy, and assess how robust the proposed model is in the presence of such data irregularities.
- V. Extend the research to evaluate the model's ability to provide reliable long-term forecasts for reinsurance revenue, which is crucial for strategic planning in the insurance industry.
- VI. Investigate the integration of external economic, political, or environmental factors into the forecasting model to improve its predictive power, as these factors can influence reinsurance revenue.

VII. Develop a framework for incorporating uncertainty and risk assessment into the reinsurance revenue forecasts, allowing insurers to make informed decisions while considering potential variability.
VIII. Adapt the model for real-time forecasting by continuously updating it with new data and assessing its ability to adapt to changing market conditions.
IX. Explore how the proposed model performs under extreme scenarios and conduct stress testing to assess its resilience in times of financial or economic crises.
X. Evaluate and compare different forecast evaluation metrics to determine the most appropriate ones for assessing the accuracy of reinsurance revenue forecasts and ensuring robust model performance.
XI. Develop methods for effectively communicating the uncertainty inherent in reinsurance revenue forecasts to decision-makers and stakeholders.
XII. Apply the model to real-life case studies and provide practical insights into its utility for reinsurance companies, including its impact on decision-making theory and risk management analysis.

These new statistical research points can provide advance the field of reinsurance revenue forecasting and contribute to more accurate and reliable predictions in the insurance and reinsurance industry.

References:

1. Abdullah, M. M., & Masmoudi, A. (2023a). Modeling Real-life Data Sets with a Novel G Family of Continuous Probability Distributions: Statistical Properties, and Copulas. Pakistan Journal of Statistics and Operation Research, 719-746.
2. Abdullah, M., M., Mohammed, W. S., & Farhana, A. H. (2023b). The Poisson-Topp-Leone Burr Type-X II Model: Various Uncensored Applications for Statistical Modeling and Some Copulas. Statistics, Optimization & Information Computing, 11(2), 276-298.
3. Alkhayyat, S. L., Mohamed, H. S., Butt, N. S., Yousof, H. M., & Ali, E. I. (2023). Modeling the Asymmetric Reinsurance Revenues Data using the Partially Autoregressive Time Series Model: Statistical Forecasting and Residuals Analysis. Pakistan Journal of Statistics and Operation Research, 425-446.
4. Box, G. and Jenkins, G. (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
5. Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
6. Brown, R. G. (1959). Statistical forecasting for inventory control. McGraw/Hill.
7. Cummins, J. D. and Griepentrog, G. L. (1985). Forecasting automobile insurance paid claim costs using econometric and ARIMA models. International Journal of Forecasting, 1(3), 203-215.
8. Darekar, A. and Reddy, A. (2017). Forecasting oilseeds prices in India: Case of groundnut. Forecasting Oilseeds Prices in India: Case of Groundnut (December 14, 2017). J. Oilseeds Res, 34(4), 235-240.
9. Elbatal, I., Diab, L. S., Ghorbal, A. B., Yousof, H. M., Elgarhy, M., & Ali, E. I. (2024). A new losses (revenues) probability model with entropy analysis, applications and case studies for value-at-risk modeling and mean of order-P analysis. AIMS Mathematics, 9(3), 7169-7211.
10. Hafiz, U. A., Salleh, F., Garba, M. and Rashid, N. (2021). Projecting insurance penetration rate in Nigeria: An ARIMA approach. REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS, 11(3), 63-75.
11. Hamedani, G. G., Goual, H., Emam, W., Tashkandy, Y., Ahmad Bhatti, F., Ibrahim, M., & Yousof, H. M. (2023). A new right-skewed one-parameter distribution with mathematical characterizations, distributional validation, and actuarial risk analysis, with applications. Symmetry, 15(7), 1297.
12. Hamed, M. S., Cordeiro, G. M., & Yousof, H. M. (2022). A new compound lomax model: properties, copulas, modeling and risk analysis utilizing the negatively skewed insurance claims data. Pakistan Journal of Statistics and Operation Research, 601-631.
13. Hashempour, M., Alizadeh, M., & Yousof, H. M. (2023). A New Lindley Extension: Estimation, Risk Assessment and Analysis Under Bimodal Right Skewed Precipitation Data. Annals of Data Science, 1-40.
14. Holt, C. E. (1957). Forecasting seasonals and trends by exponentially weighted averages (O.N.R. Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh USA.
15. Hyndman, R. J., and Athanasopoulos, G. (2018). Holt's linear trend method. Forecasting: principals and practice, 2nd Ed.
16. Iqbal, M. C., Jamshaid, T. M., & Rashid, A. Q. A. (2016). Forecasting of wheat production: a comparative study of Pakistan and India. International Journal of Advanced Research (IJAR), 4(12), 698-709.
17. Jakaša, T., Andročec, I. and Sprčić, P. (2011). Electricity price forecasting-ARIMA model approach. In 2011 8th International Conference on the European Energy Market (EEM) (pp. 222-225). IEEE.

18. Jang, K. P., Kam, S. and Park, J. Y. (1991). Trend and Forecast of the Medical Care Utilization Rate, the Medical Expense per Case and the Treatment Days per Cage in Medical Insurance Program for Employees by ARIMA Model. Journal of Preventive Medicine and Public Health, 24(3), 441-458.
19. Kumar, V. S., Satpathi, D. K., Kumar, P. P. and Haragopal, V. V. (2020). Forecasting motor insurance claim amount using ARIMA model. In AIP Conference Proceedings (Vol. 2246, No. 1, p. 020005). AIP Publishing LLC.
20. Ljung, G. and Box, G. (1978). On a Measure of Lack of Fit in Time Series Models, Biometrika, 65, 297-303.
21. Minkah, R., de Wet, T., Ghosh, A., & Yousof, H. M. (2023). Robust extreme quantile estimation for Pareto-type tails through an exponential regression model. Communications for Statistical Applications and Methods, 30(6), 531-550.
22. Mohamed, H. S., Cordeiro, G. M. and Yousof, H. M. (2022). The synthetic autoregressive model for the insurance claims payment data: modeling and future prediction. Statistics, Optimization & Information Computing, forthcoming.
23. Mohamed, H. S., Cordeiro, G. M., Minkah, R., Yousof, H. M., & Ibrahim, M. (2024). A size-of-loss model for the negatively skewed insurance claims data: applications, risk analysis using different methods and statistical forecasting. Journal of Applied Statistics, 51(2), 348-369.
24. Mohammadi, H. and Rich, D. P. (2013). Dynamics of unemployment insurance claims: an application of ARIMA-GARCH models. Atlantic Economic Journal, 41(4), 413-425.
25. Nath, B., Dhakre, D. S. and Bhattacharya, D. (2019). Forecasting wheat production in India: An ARIMA modelling approach. Journal of Pharmacognosy and Phytochemistry, 8(1), 2158-2165.
26. Palakuru, M., Yarrakula, K., Chaube, N. R., Sk, K. B. and Satyaji Rao, Y. R. (2019). Identification of paddy crop phenological parameters using dual polarized SCATSAT-1 (ISRO, India) scatterometer data. Environmental Science and Pollution Research, 26(2), 1565-1575.
27. Sahu, P. K., Mishra, P., Dhekale, B. S., Vishwajith, K. P. and Padmanaban, K. (2015). Modelling and forecasting of area, production, yield and total seeds of rice and wheat in SAARC countries and the world towards food security. American Journal of Applied Mathematics and Statistics, 3(1), 34-48.
28. Salem, M., Emam, W., Tashkandy, Y., Ibrahim, M., Ali, M. M., Goual, H., & Yousof, H. M. (2023). A new lomax extension: Properties, risk analysis, censored and complete goodness-of-fit validation testing under left-skewed insurance, reliability and medical data. Symmetry, 15(7), 1356.
29. Shrahili, M.; Elbatal, I. and Yousof, H. M. (2021). Asymmetric Density for Risk Claim-Size Data: Prediction and Bimodal Data Applications. Symmetry 2021, 13, 2357.
30. Tashkandy, Y., Emam, W., Ali, M. M., Yousof, H. M., & Ahmed, B. (2023). Quality control testing with experimental practical illustrations under the modified Lindley distribution using single, double, and multiple acceptance sampling plans. Mathematics, 11(9), 2184.
31. Venezian, E. C. and Leng, C. C. (2006). Application of spectral and ARIMA analysis to combined-ratio patterns. The Journal of Risk Finance.
32. Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management Science, 6(3), 324-342.
33. Yousof, H. M., Ansari, S. I., Tashkandy, Y., Emam, W., Ali, M. M., Ibrahim, M., & Alkhayyat, S. L. (2023a). Risk analysis and estimation of a bimodal heavy-tailed burr XII model in insurance data: exploring multiple methods and applications. Mathematics, 11(9), 2179.
34. Yousof, H. M., Chesneau, C., Hamedani, G., & Ibrahim, M. (2021). A new discrete distribution: Properties, characterizations, modeling real count data, Bayesian and non-Bayesian estimations. Statistica.
35. Yousof, H. M., Emam, W., Tashkandy, Y., Ali, M. M., Minkah, R., & Ibrahim, M. (2023b). A novel model for quantitative risk assessment under claim-size data with bimodal and symmetric data modeling. Mathematics, 11(6), 1284.
36. Yousof, H. M., Tashkandy, Y., Emam, W., Ali, M. M., & Ibrahim, M. (2023c). A New Reciprocal Weibull Extension for Modeling Extreme Values with Risk Analysis under Insurance Data. Mathematics, 11(4), 966.