Comprehensive Forecasting Analysis of Insurance Trends: A Comparative Study of ARIMA, Neural Networks, Dynamic Regressions, Prophet, and Combination Models and Volatility Forecast Using Insurance Dataset

Reza Heidari

Marzieh Yavari

Contents

Abstract

This study employs a comprehensive analysis of US insurance data through advanced time series forecasting models, including Autoregressive Integrated Moving Average (ARIMA), Neural Networks (NN), and Prophet. Utilizing forecast data from the fpp3 library, the research aims to identify the most effective model for predicting insurance trends. Additionally, a volatility forecast is performed to understand the dynamics of these variables. Through a rigorous comparative analysis, the study evaluates and compares the performance of these models using some metrics. The results indicate that the combination method outperforms other approaches, exhibiting lower Root Mean Squared Error (RMSE). Furthermore, the study observes a significant dependence of quote volatility on TV advertisements.

Introduction

Who cares about the interplay between television advertising and monthly quotations for an insurance company? The answer lies in the core of strategic decision-making within the insurance industry. This project delves into the monthly data spanning January 2002 to April 2005, aiming to decipher patterns and dependencies, with television advertising spending as a key predictor. The question addressed here is not merely about numbers; it's about empowering the insurance company with predictive capabilities that influence marketing strategies and resource allocation.

Understanding the nuanced relationship between television advertising and monthly quotations is imperative for the insurance sector. Accurate forecasting in this context is not just an academic pursuit; it directly impacts how companies allocate their marketing budget and strategize for the future. In an ever-evolving market, the ability to predict quotations based on historical data becomes a strategic advantage, enabling proactive responses to changing dynamics.

This project is more than an exploration of time series analysis—it's a practical application that brings theoretical methodologies to the forefront of decision-making. The complexity introduced by the temporal nature of the dataset, coupled with the inclusion of television advertising expenditure, makes this study highly relevant to real-world scenarios. As we venture into the application of various forecasting models, our aim is clear: to bridge the gap between theory and practice, offering tangible insights that matter to industry professionals and decision-makers in the insurance landscape.

Data

We sourced our dataset from the fpp3 library, encompassing monthly data spanning April 2002 to April 2005. The dataset includes variables such as 'Month' representing the timeframe, 'Quotes' indicating monthly insurance sales, and 'TV adverts' denoting the advertising costs for the insurance company. In this section, we present the results of exploratory data analysis (EDA) using critical plots, providing an extensive summary of essential trends, and relationships within the dataset.

STL Decomposition

A crucial observation is the apparent similarity between Quotes and the remainder, suggesting that unobserved variables may have significantly influenced the rate of Quotes in this dataset. The trend component indicates a notable increase in quote sales in the range from the end of 2002 to the first months of 2003 and after that, there is a considerable decrease (Fig.1).

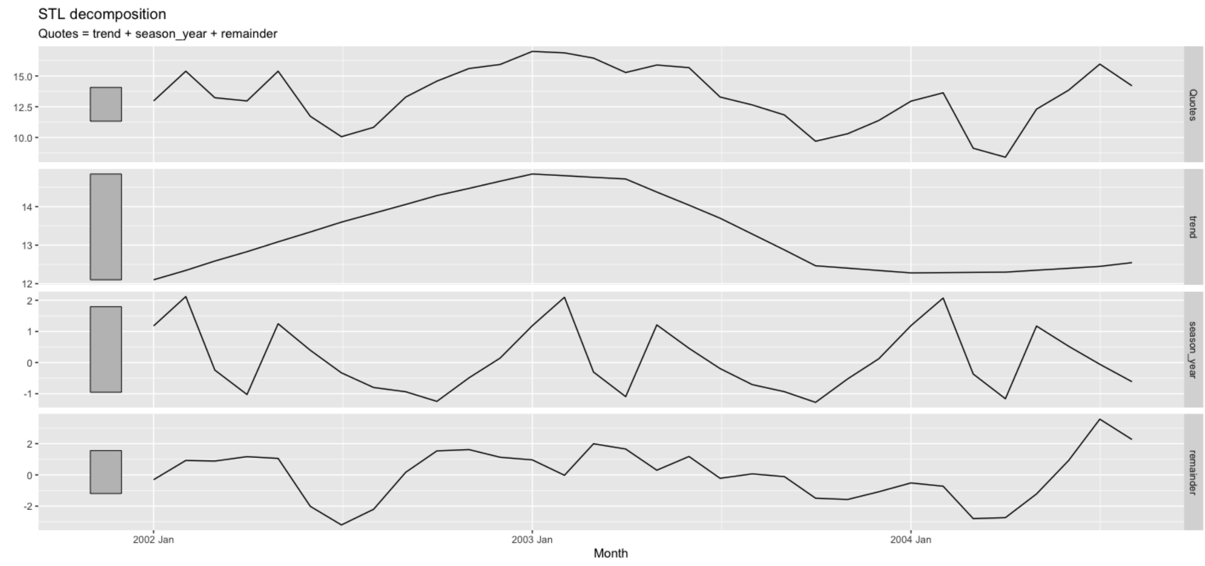

Figure 1 - STL Decomposition

Line plot

The line plot demonstrates a consistent alignment between the TV adverts line and the Quotes line, implying a positive relationship. The synchronous fluctuations in both lines suggest a direct impact of TV adverts on insurance company Quotes. The parallel nature of the ups and downs in the plot further supports the notion of a consistent and positive influence of TV adverts on the observed fluctuations in Quotes during this time

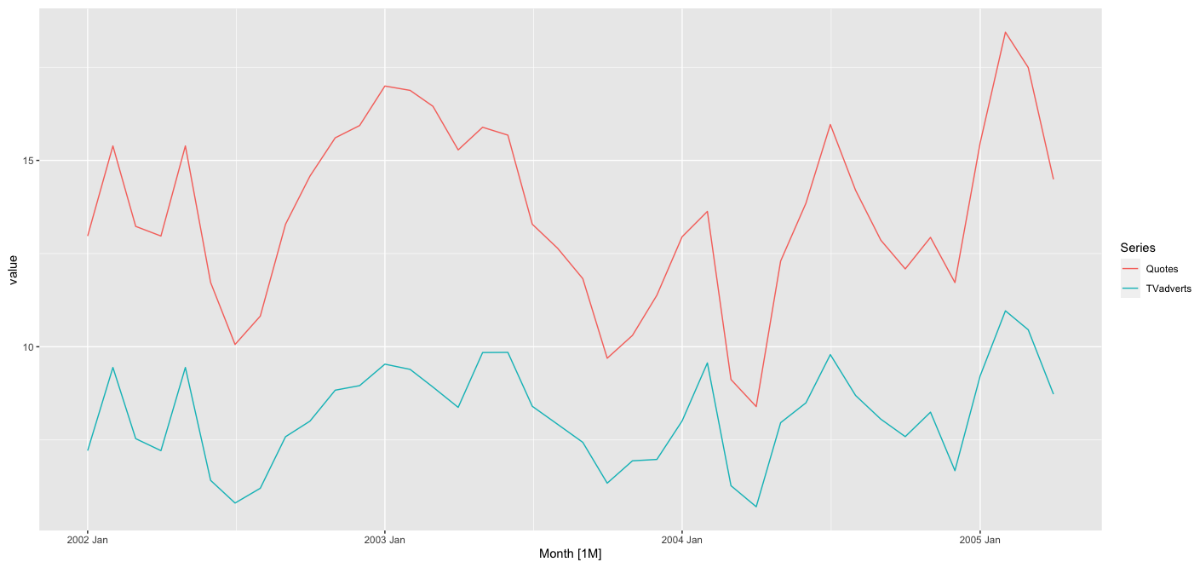

Figure 2 - Comparing the trends of quotes and TV adverts

ACF plot

In the Autocorrelation Function (ACF), at lag 1, there is a strong positive correlation, indicating a significant linear relationship with the immediately preceding observation. At lags 10 and 11, there are notable spikes, that can suggest periodic patterns every 10 and 11 time points.

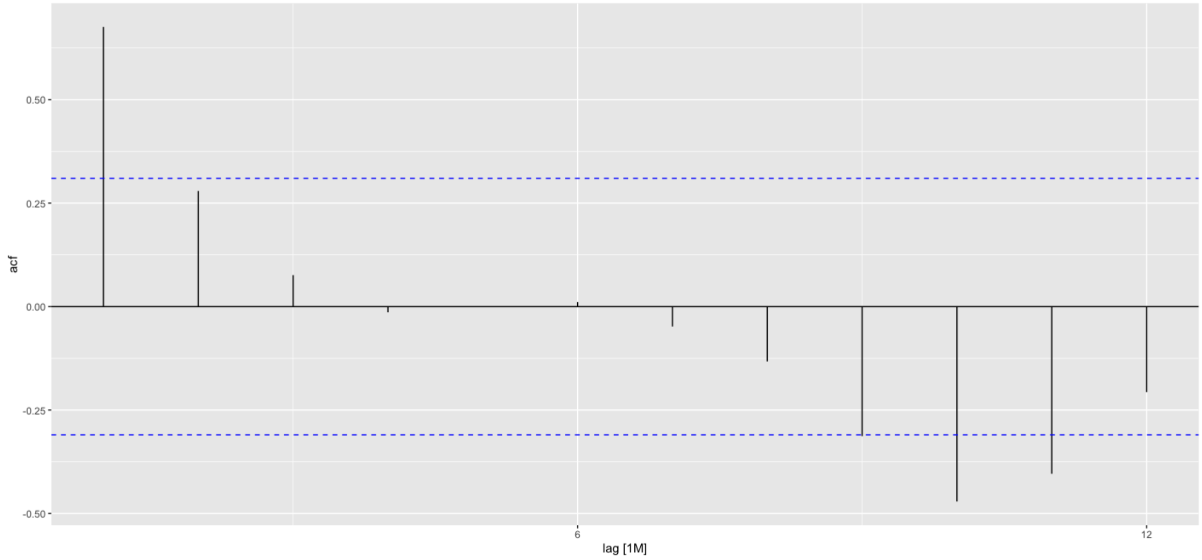

Figure 3 - Auto Correlation Function plot

Method

All the R codes for implementing and analysis of the dataset can be found in GitHub link: https://github.com/yavarim21/TimeSeriesPr.git

Modeling Methods

Our modeling strategy was characterized by the implementation of a diversity of forecasting models. The collection contains comparative assessments of benchmark methods and advanced forecasting techniques, including ARIMA, Dynamic Regression, Prophet, and Volatility Forecasting. ARIMA, a time series forecasting method, was applied to capture temporal patterns, while dynamic regression allowed us to incorporate external factors influencing insurance trends. Prophet, designed for forecasting time series with strong seasonal patterns, was employed for its adaptability to various datasets. We applied a combination model to check the performance of the combinations of different methods to improve accuracy. A volatility forecast was computed for this project to get insight into the changes in market conditions.[1]

Cross-Validation

To mitigate overfitting and ensure the generalizability of our models, we implemented a robust cross-validation framework. Specifically, we employed k-fold cross-validation, partitioning the dataset into k subsets iteratively. Each iteration involves training the model on k-1 subsets and validating it on the remaining subsets. This process was repeated k times, allowing us to validate the models on different subsets of the data and assess their performance across various scenarios.[1]

Evaluation Metrics

In this project, we exclusively utilized RMSE as the primary metric for model selection, focusing on its ability to quantify average prediction errors. The decision to rely solely on RMSE was driven by its alignment with our analysis goals and emphasis on accuracy assessment.

Results

Quotes Volatility Results

This part allows the insurance company to be adaptable to changes in market conditions. If there are sudden shifts in the relationship between TV adverts and the rate of quotes, the company can quickly adjust its strategies to capitalize on new opportunities or mitigate risks.

Sign Correlation

We obtained the amount of sign correlation based on the volatility of Quotes and the result is 0.8055941 suggesting the volatility follows an almost Beta distribution (4,4).

Volatility forecast

To do a volatility forecast, we applied the neural network method to Quote volatility. The neural network model results indicate an NNAR (2,2) configuration, designed to leverage information from the two most recent time steps and utilize two hidden neurons for learning complex relationships within the data. Forecast results based on the test dataset for 7 months ahead of Quotes data suggest an RMSE equal to 0.2036623.

The plot below offers insights into the NNAR model's effectiveness in capturing and predicting the evolving volatility of insurance company quotes over a 7-month horizon. The volatility forecast demonstrates a reasonable ability to capture changes, and the confidence intervals provide a precise coverage of potential changes over the specified period.

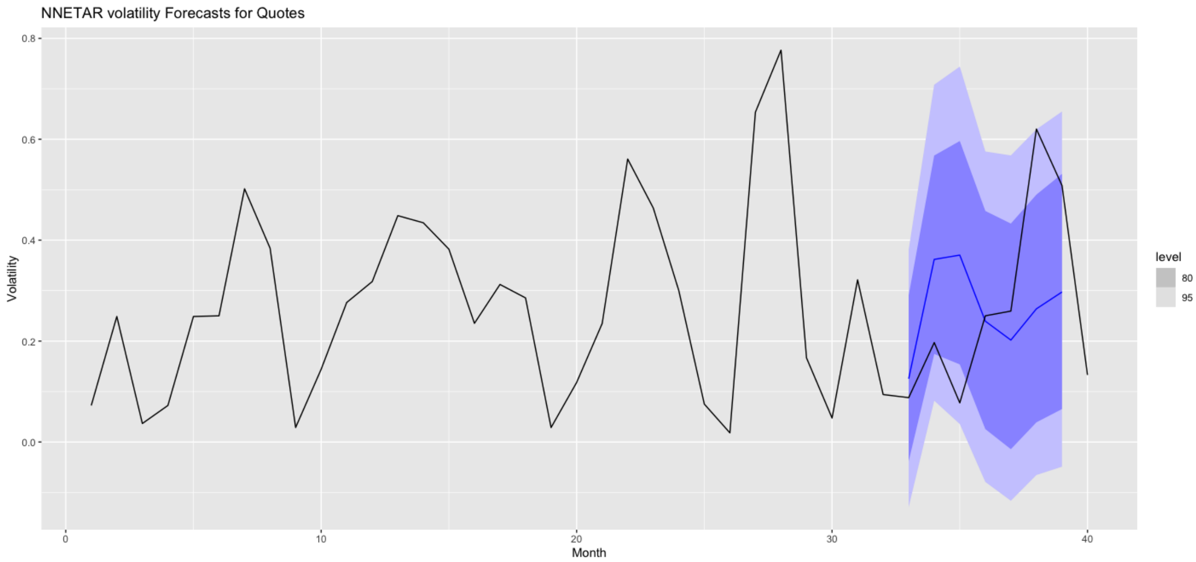

Figure 4 - Volatility Forecasts for Quotes

Models Comparison

In this section, we compared TSLM, ARIMA, NNAR, Mean, Naive, Seasonal Naive, Drift, EWMA, and Prophet models to identify best model for our forecasting needs.

The top-performing models, based on RMSE, were TSLM and ARIMA models. The TSLM model, which incorporates TVadverts as a predictor, exhibited an RMSE of 0.465, showcasing its ability to provide accurate predictions. Similarly, another TSLM model, which incorporates TVadverts and its lag as predictors, achieved an RMSE of 0.475.

Among the ARIMA models, ARIMA(0,1,1) errors, integrating TVadverts as a predictor, and ARIMA(0,1,1) errors, incorporating both TVadverts and its lag, showed promising performance with RMSE values of 0.935 and 0.943, respectively. These results suggest that considering TVadverts in the forecasting process enhances the predictive power of the models.

It's noteworthy that the Drift model, despite having a lower RMSE compared to some other models, exhibited a relatively modest performance with an RMSE of 2.287. This indicates that incorporating a drift term alone might not be sufficient for accurate predictions in this context.

In contrast, models like Prophet and NNET demonstrated higher RMSE values of 10.480 and 4.620, respectively, indicating less accurate predictions. These models may not be suitable for this specific dataset or might require further tuning to improve their performance. The forecast accuracy of the first four models with the lowest RMSE:

Model	RMSE
TSLM2 = TSLM(Quotes ~ TVadverts)	0.4651792
TSLM4 = TSLM(Quotes ~ TVadverts + lag(TVadverts))	0.4751055
ARIMA2 = ARIMA(Quotes ~ TVadverts)	0.9348257
ARIMA4 = ARIMA(Quotes ~ TVadverts + lag(TVadverts))	0.9428390

Table 1- RMSE Forecasting

In the pursuit of refining forecasting accuracy for the insurance dataset, we explored combinations of the 4 top-performing models that identified. The combination models were created by averaging predictions from pairs of the top four models, two ARIMA and two TSLM models, to leverage the strengths of each constituent model.

Among the combination models that formed by averaging predictions from both TSLM models, exhibited a reduced RMSE of 0.447, suggesting improved performance compared to the individual models. This indicates that combining the insights captured by both TSLM yields a more accurate prediction.

Similarly, another combination, created by blending predictions from The TSLM model, which incorporates TVadverts as a predictor, and ARIMA models, ARIMA(0,1,1) errors, integrating TVadverts as a predictor, achieved an RMSE of 0.564, showcasing a competitive performance.

Contrastingly, some combination models involving NNET, displayed increased RMSE values, indicating that certain model combinations may not necessarily lead to improved forecasting accuracy. It is essential to carefully evaluate the impact of combining specific models and consider the dynamics of the dataset.

The forecast accuracy of the first four models with the lowest RMSE:

Model	RMSE
Combination of TSLM2 and TSLM4	0.4468999
TSLM2	0.4651792
TSLM4	0.4751055
Combination of TSLM2 and ARIMA4	0.5640734

Table 2- RMSE Forecasting

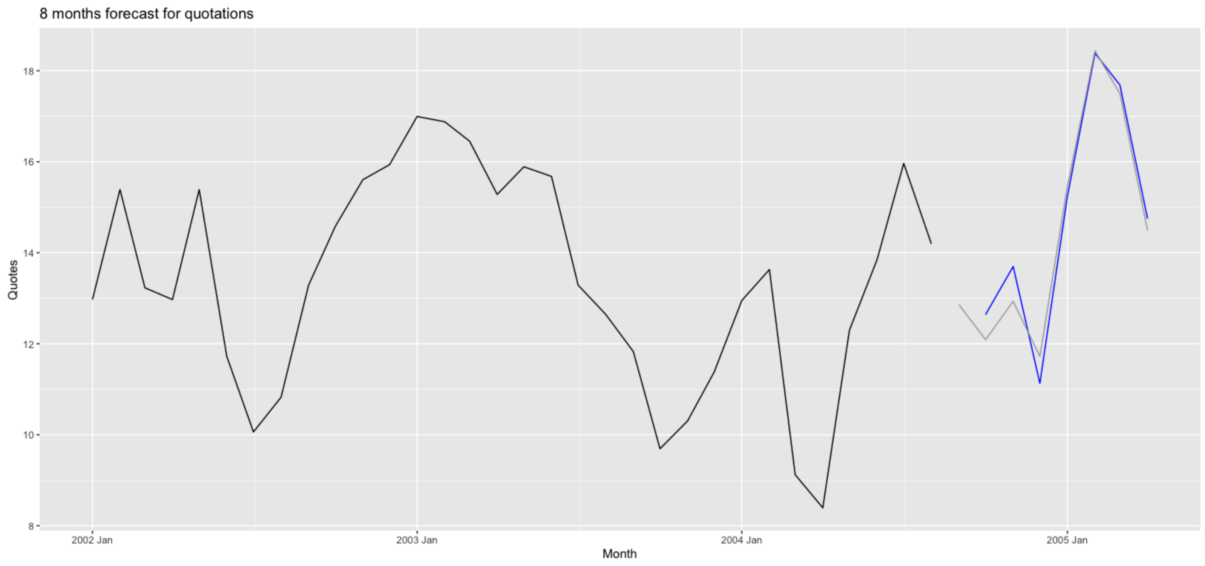

Figure 5 - Forecasting Graph from Combination of TSLM2 and TSLM4 models.

Cross Validation

By utilizing cross-validation for the combination of TSLM2 and TSLM4, we enhance the reliability of our model evaluation. The average RMSE of 0.5465743 obtained through cross-validation reflects the performance of the combined model across diverse subsets of the data. The combination model, which averages predictions from TSLM2 and TSLM4, demonstrates stability in predictive performance across different partitions of the dataset.

Comparatively, the RMSE without cross-validation, specifically on the Test data, was 0.4468999. This single RMSE value serves as a baseline reference for the model's accuracy when trained and tested on the entire dataset without considering cross-validation. The marginally higher average RMSE with cross-validation (0.5465743) suggests that the combined model's predictive accuracy is consistent and provides a more realistic estimate of its performance on new, unseen data.

Conclusion

In this project, spanning from January 2002 to April 2005, we delved into the relationship between television advertising and monthly quotations for an insurance company. Utilizing diverse forecasting models, the analysis highlighted the effectiveness of the NNAR model in capturing changes in market conditions through volatility forecasting. Among the models, combination of TSLM (Quotes ~ TVadverts) and TSLM(Quotes ~ TVadverts + lag(TVadverts)) emerged as the top performer, emphasizing the crucial role of TVadverts and its lag in predicting monthly quotations accurately.

Our findings underscore the practical implications for decision-makers in the insurance industry, providing valuable insights into adjusting marketing strategies based on TV advertising dynamics. Looking forward, further research could explore the impact of external factors on forecasting accuracy. Overall, this project bridges theoretical methodologies with practical applications, offering actionable intelligence for navigating the complexities of marketing and strategic decision-making in the insurance landscape.

References

1. Hyndman, RJ, Athanasopoulos G. Forecasting: Principles and practice (3rd ed) [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://otexts.com/fpp3/