

EXECUTIVE SUMMARY OF THE PROJECT

B22GEOKO

The project aimed to forecast tourist flows to Czechia from five key countries, employing diverse methodologies, showcased in the course Predictive Analysis, such as Exponential Smoothing (ETS) models, stepwise regression, Lasso and Ridge regression, and Temporal Hierarchies Forecasting (THieF). The evaluation criteria included the Mean Absolute Percentage Error (MAPE) and visual analysis of the results. Although efforts were made to integrate IMF economic indicators, the predictive models demonstrated suboptimal performance, emphasizing the intricate relationship between economic variables and tourism dynamics. Notably, the project identified that each time series could be better predicted by different models, highlighting the necessity for a tailored approach for each time series.

The investigation emphasized the importance of considering seasonality and trends while modeling time series data, using tools such as dummy variables and differencing techniques. The analysis revealed that models like THieF, ETS, and Ridge regression exhibited stable performance across various time series, while the stepwise regression method showcased its effectiveness in building regression models. However, it was noted that the predictions made before 2019 might not be entirely applicable in the post-pandemic era due to the unprecedented disruptions caused by the Covid-19 pandemic. To enhance forecasting accuracy, the incorporation of real-time data and qualitative insights from industry experts was recommended.

Contents

Abstract	2
I. Introduction	2
II. Mean Absolute Percentage Error (MAPE)	2
III. Dataset & Strategy	3
IV. Exponential Smoothing	3
V. Lags for Regression	4
VI. Stepwise Regression.....	4
VII. Modeling with Trend and Seasonality.....	5
VIII. Lasso & Ridge Regression	5
IX. Temporal Hierarchies Forecasting (THieF)	5
X. Discussion.....	6
XI. Conclusion.....	6
References	6
Appendix	7

Predicting Tourist Flow in Czechia

Georgios Kokolakis
University of Skövde
b22geoko

ABSTRACT

The tourism industry holds significant economic and cultural importance globally, prompting a need for accurate forecasting models. This project focuses on predicting tourist flows to Czechia from five key countries: Germany, Italy, the United Kingdom, Russia, and the United States. Various forecasting techniques were employed, including Exponential Smoothing (ETS) models, stepwise regression, Lasso and Ridge regression, and Temporal Hierarchies Forecasting (THieF). Evaluations were based on the Mean Absolute Percentage Error (MAPE) and visualization of results. Despite efforts to incorporate IMF economic indicators, predictions remained discouraging, emphasizing the complexity of tourism dynamics and the need for further investigation.

I. INTRODUCTION

On a global scale, the tourist industry represents a cornerstone of economic growth and cultural interaction. In an era of rising mobility and worldwide connection, nations around the world understand tourism's significant importance to their economies and cultures [1]. Czechia has evolved as an attractive destination among these nations, recognized for its rich historical history, architectural magnificence, and lively cultural tapestry. Czechia has seen a steady increase in visitor visits throughout the years, reflecting its rising international appeal, according to the World Bank [2].

The precise forecasting of visitor arrivals is critical in this setting. Predictive models are invaluable tools for policymakers, tourism authorities, and industry players looking to forecast and adapt to changing tourist influx patterns. Effective prediction not only helps with resource allocation, but it also aids in strategic planning and marketing

initiatives, ultimately improving the overall tourist experience. This report will look into the complexities of forecasting tourist flows to Czechia for the year 2019. Our investigation broadens its scope to include five main origin countries: Germany, Italy, the United Kingdom, Russia, and the United States of America. These countries contribute significantly to Czechia's international tourism business, making their arrivals of particular interest.

Our fundamental goal, as demonstrated throughout the course, is to construct solid forecasting models suited to each origin country, led by a commitment to precision and reliability. We use a variety of forecasting methodologies throughout our investigation, with the Mean Absolute Percentage Error (MAPE) acting as our preferred statistic for evaluating model performance.

II. MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

The Mean Absolute Percentage Error (MAPE) is a popular metric for assessing a model's forecasting accuracy. It is simple to analyze and explain because it represents the average percentage difference between the projected and actual values (Equation 1). MAPE is valuable because it can be used to compare the accuracy of different models and to evaluate a model's performance over time [3]. MAPE is scale-independent, which means it is unaffected by data scale. It can be used to compare the performance of several models or datasets, regardless of their magnitude, which is appropriate for comparing our models where error metrics amount to the thousands.

Time_Series_Plots

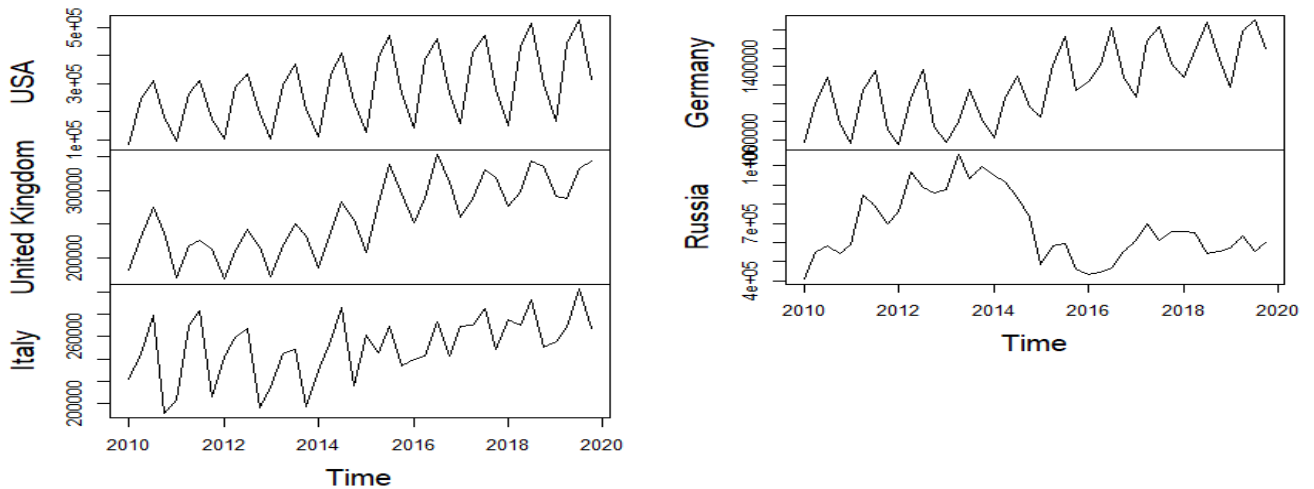


Figure 1. Time Series for each origin Country

$$MAPE = \frac{1}{n} \sum \left| \frac{\text{actual value} - \text{predicted value}}{\text{actual value}} \right| 100$$

III. DATASET & STRATEGY

Our research draws upon historical data pertaining to tourist arrivals in Czechia. These datasets offer a quarterly breakdown of tourist arrivals from five key origin countries: Germany, Italy, the United Kingdom, Russia, and the United States. Each column of the dataset provides a time series of arrivals, allowing us to explore and predict the temporal patterns in tourist flows. We initiated our analysis by setting our dataset from the year 2010 (the first available data entries) to 2019 (target year). Then we plot the time series to gain a better understanding of our data as presented in Figure 1.

From the aforementioned plot one can make these general observations:

- In most plots we have the presence of seasonality.
- One can spot trends in some time series such as USA or UK
- Russia is the rather irregular, specially when compared to the rest. We have an upwards trend and possibly seasonality until 2014. However, in 2014 to 2015 we have a big damp which destabilizes the time series.

Given the above observations, we have to deal with each time series individually to account for their specific characteristics. Before we move into the modeling part let us discuss the train and test intervals. As mentioned previously, the data available to us are quarters from years 2010-2019, which translates to 36 observations per origin country. The number of those observations might be an issue for the accuracy of our predictions. Perhaps not for Exponential Smoothing (Error Trend Seasonality or ETS) or Autoregressive Integrated Moving Average (ARIMA) models which adapt into short term patterns [4] [5], but in

these cases, one should consider time series like Russia, which destabilizes around 2014, and anticipate the performance of some models to be less effective. We set our testing set to be the four quarters of 2019 and use the rest as a training set. We will also utilize a validation set to help us with the model selection [6]. Our validation set will be years 2017 and 2018, roughly the last 8 observations of the training set that is. We can now begin our forecasting by exploring different ETS models tailored to the origin country.

IV. EXPONENTIAL SMOOTHING

Exponential smoothing is a time series forecasting method that uses an exponentially weighted average of past observations to predict future values [7] [8]. This method assigns more weight to recent observations and less to older observations, allowing the forecast to adapt to changing trends in the data [9]. Exponential smoothing is a rule of thumb technique for smoothing time series data using the exponential window function. Whereas in the simple moving average the past observations are weighted equally, exponential functions are used to assign exponentially decreasing weights over time [9].

The implementation of exponential smoothing in R requires the specification of three components Error, Trend and Seasonality. To do so, we consult the time series plots in Figure 1, then narrow down the possible combinations to the ones that apply in each time series. Having established the available combinations for each time series, we use the rolling origin a technique used to evaluate the performance of time series models [10]. We iterate through the different models and obtain several forecast errors for time series, which gives a better understanding of how each model performs. At the end of the process, we create a boxplot to compare the errors of the models as shown in Figure 3.

In the example presented in Figure 3 concerning the USA time series (see Figure 1), given that our time series is seasonal and has an upward trend, we have the following possible combinations of ETS:

- Additive error, trend, and seasonality

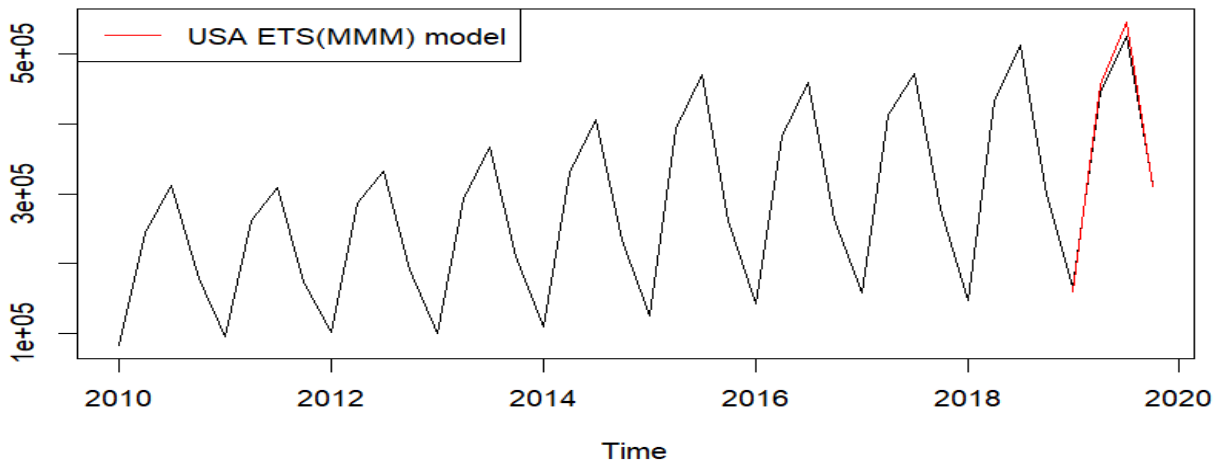


Figure 2. ETS(MMM) prediction for tourist flow from the USA to Czechia for 2019.

- Multiplicative error, additive trend, and seasonality
- Multiplicative error and seasonality, additive trend
- Multiplicative error, trend, and seasonality
- The Naïve
- The mean combination of the above
- The median combination of the above

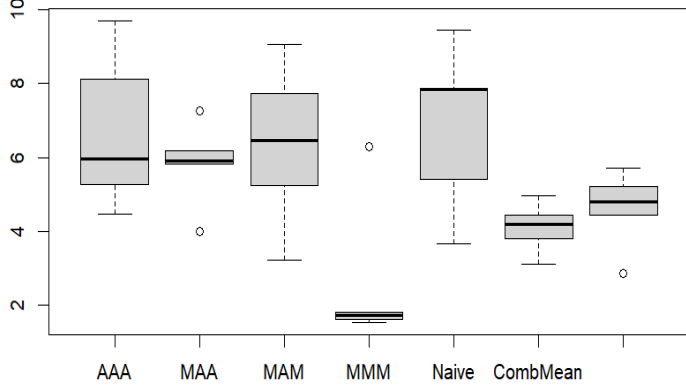


Figure 3. ETS combinations for the USA time series.

From these options ETS(MMM) model performance appears to be the most promising, among other ETS models, for predicting the tourist flow of Americans in Czechia. However, error metrics can always be deceiving thus we visualize the result (Figure 2) for a more concrete evaluation of the model. The visual results match the metric results, and we repeat the process with the appropriate model combinations for the remaining countries. The visual results are featured in the Appendix alongside the error results in Table 1, Figure 7, Figure 9, Figure 8, Figure 5. Something interesting to point out is the fact that for Russia there is a significant gap between the error percentage of the validation and testing set. In the validation set the rolling origin technique points towards the Naïve, however in the testing set MMM is clearly the better approach. We compare Figure 6, Figure 7 to reach the conclusion that the naïve is not actually performing well at all, which solidifies the statement that the metrics can be misleading.

V. LAGS FOR REGRESSION

Time series data presents unique challenges when building regression models, primarily due to the temporal nature of the data and the presence of autocorrelation. Autocorrelation represents the correlation between a variable and its past values, a common feature in time series datasets. To address this, integrating lagged variables into regression models is a powerful strategy. By including lagged variables in regression models, we can capture the effect of past observations on the current observation and improve the accuracy of our predictions [11]. Including them in a regression model offers several benefits.

First and foremost, they introduce memory into the model, allowing it to account for dependencies between observations

at different time points [12]. This is vital when dealing with data exhibiting serial correlation, where current values are influenced by their historical values. Furthermore, lagged variables help preserve temporal patterns and seasonality within the data [12]. This is especially valuable when working with datasets that display recurring patterns at specific time intervals. By incorporating these temporal elements, the model can better capture and understand the underlying structure of the data. Another advantage is enhanced predictive accuracy [12]. The additional lagged variables provide more predictors for the model, improving its ability to forecast future values accurately. This feature is particularly useful when the behavior of the dependent variable is closely related to its own historical values.

Nevertheless, there are important considerations to keep in mind. Selecting the appropriate lag length can be a complex task, often requiring data analysis and domain knowledge. Tools like autocorrelation and partial autocorrelation function (PACF) plots are useful for determining the optimal lag length. Five PACF plots are produced, one for each origin country to determine the appropriate lag length.

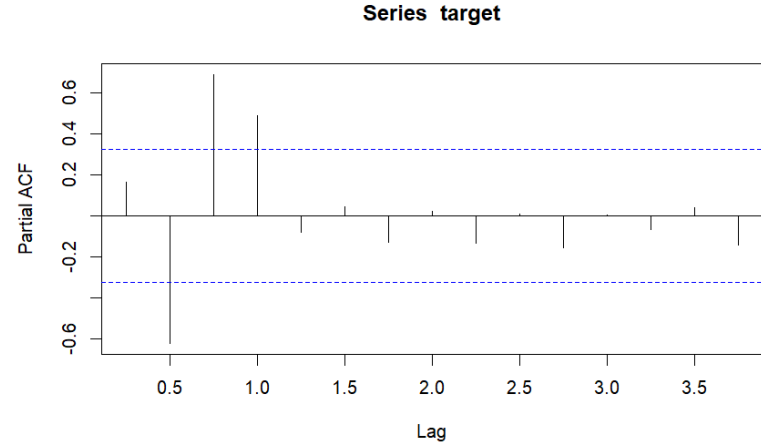


Figure 4. USA PACF

As demonstrated in Figure 4, in the PACF plot of the USA showcases 3 significant lags (lag 2, 3 & 4). Thus, the lag length for the USA time series will be set to 4. We repeat the above process to determine the rest of the lag lengths in the remaining time series.

VI. STEPWISE REGRESSION

Stepwise regression is a procedure used to build a regression model from a set of predictor variables by entering and removing predictors in a stepwise manner into the model until there is no statistically valid reason to enter or remove any more [13]. The goal of stepwise regression is to build a regression model that includes all of the predictor variables that are statistically significantly related to the response variable [13]. We will be using the Stepwise AIC method which uses the Akaike information criterion (AIC) for variable selection in linear regression modeling [13].

The above model guarantees the selection of the most significant lags for our time series and produces an adequate

outcome in most of our cases, as featured in Figure 10, Figure 11, Figure 12, Figure 13, Figure 14 and in Table 1, in the Appendix.

VII. MODELING WITH TREND AND SEASONALITY

As presented in Figure 1, each time series has different properties in terms of seasonality and trend. The stepwise regression model, although reliable, does not account for those properties. When we model the lags of a time series, we need to account for seasonality and trends because they can have a significant impact on the dependent variable. Seasonality refers to the pattern that repeats itself at regular intervals, such as daily, weekly, or yearly. Trends refer to the long-term changes in the data that may be increasing or decreasing over time. The inability to gain useful insights from time series data might be hampered by a lack of proper modeling for seasonality and patterns. Failure to account for these critical components might result in inefficient forecasting and decision-making, limiting the model's value in giving accurate predictions and useful insights for informed decision-making [14].

Dummy variables are a frequent method for dealing with seasonality. These variables are binary indicators that indicate the dataset's various seasons or time periods. By including these dummy variables in the model, the influence of each season on the target variable may be estimated, allowing for seasonal fluctuations in the data [15]. Furthermore, modeled differences can be used to identify trends in time series data. Identifying and incorporating the underlying long-term trends or movements within the data is what trend modeling entails [16]. The term "modeled differences" refers to the use of differencing techniques to turn data into a stable series. This procedure helps to remove any non-stationary underlying trends or patterns, making it easier to model the data using traditional time series models. It is feasible to remove trends from time series data by differencing it and then modeling the resulting stationary series. This method is frequently used to account for any linear or nonlinear trends that may exist in the data. Differences in modeling can aid in understanding the pace of change or growth within a time series, allowing the detection of any underlying patterns that may not be immediately evident [16].

Our initial exploration of the data showed the presence of seasonality and trend in many of our time series. Thus, we create a new regression model for each time series which accounts for trend and seasonality. The results presented in Table 1 indicate that in time series such as Germany and USA, where trend and seasonality are more evident, the "Difference model" performs better than the stepwise regression. On the other hand, as expected when it comes to times series such as Russia, where there is no trend or seasonality the model indicates poor performance. Visual results are shown in the Appendix, namely in Figure 15, Figure 16, Figure 17, Figure 18, Figure 19.

VIII. LASSO & RIDGE REGRESSION

Having established the lags and created well performing models we explore Lasso and Ridge regression for comparison.

Lasso regression, also known as Least Absolute Shrinkage and Selection Operator, is a regression analysis approach that is used for variable selection and regularization. It is primarily used when dealing with datasets with a big number of variables, especially when some of these variables are useless or redundant [17]. By reducing the coefficient estimates of less significant predictors to zero, Lasso regression helps to reduce overfitting and enhance model prediction accuracy [17] [18].

Ridge regression is a regression analysis method for dealing with multicollinearity in a dataset with strongly linked independent variables [19]. Multicollinearity can result in inconsistent estimations of regression coefficients, making model interpretation complicated and untrustworthy. Ridge regression solves this problem by including a penalty element in the ordinary least squares (OLS) regression equation, which helps to stabilize the regression coefficient estimations [20]. Ridge regression, like lasso regression, includes a regularization element in the OLS equation that is added to the sum of squared residuals. Unlike lasso, however, ridge regression employs the total of the squared coefficient values multiplied by a tuning parameter (commonly represented as λ) as the penalty term. This ensures that the coefficients are reduced to zero but not pushed to be exactly zero [21]. As a result, ridge regression allows all variables to remain in the model while reducing their impact, eliminating overfitting and increasing the model's overall predictive performance [21].

The results are presented in Figure 24, Figure 25, Figure 26, Figure 27, Figure 31, Figure 28, Figure 29, Figure 30, Figure 32, Figure 33 in the Appendix. Alongside Table 1, the outcome exhibits the strength of the models without however giving us a clear indication of which of the models has an overall better performance.

IX. TEMPORAL HIERARCHIES FORECASTING (THiEF)

Last but not least, we investigate the use of THiEF in our time series. According to [22] the Temporal Hierarchical Forecasting (THiEF) method uses a hierarchical time series methodology to provide forecasts at multiple temporal frequencies. Taking a seasonal time series, the approach computes all potential temporal aggregations that result in an integer number of observations per year. Predictions made at all aggregation levels are merged to produce forecasts that are temporally reconciled, accurate, and resilient. The implied combination reduces modeling uncertainty, while the reconciled nature of the forecasts results in a coherent prediction that facilitates synchronized judgments across many planning horizons, ranging from short-term operational planning to long-term strategic planning. THiEF can integrate high-level managerial forecasts with complicated and unstructured data with lower-level statistical projections. The process is independent of forecasting models and has the

potential to improve accuracy over traditional forecasting, particularly under increased modeling uncertainty.

THieF outperforms the majority of the models in terms of metrics as shown in Table 1. The performance of the model is stable across all time series as one can see in Figure 20, Figure 21, Figure 22, Figure 23, Figure 34.

X. DISCUSSION

To achieve accurate predictions, we employed various forecasting models, including Exponential Smoothing (ETS) models, stepwise regression, Lasso regression, Ridge regression, and Temporal Hierarchies Forecasting (THieF). The models were evaluated using the Mean Absolute Percentage Error (MAPE) and visualization of the results.

Utilizing the International Monetary Fund (IMF) dataset comprising essential economic indicators such as inflation indices and Gross Domestic Product (GDP) for each country, we attempted to leverage this information to predict tourist flows. Despite the disparity between the quarterly tourist flow dataset and the annual IMF data, we devised a strategy to align the datasets by dividing the yearly figures by four. This approach allowed us to merge the datasets and explore potential relationships between the economic indicators and the tourist flow. Regrettably, our predictive models did not yield the desired results, with the metrics indicating suboptimal performance (Table 1). This outcome prompts further examination of the underlying dynamics and factors influencing tourist flows, emphasizing the complex interplay between economic variables and tourism trends that warrant deeper investigation for more accurate predictions and insights.

Table 1 showcases the most prominent, in terms of MAPE, models for every time series. All models have strengths and weaknesses, which is supported by the fact that each time series is better approximated by different models. Namely, Germany is better predicted by the regression model with the dummies accounting for seasonality and modeled differences. Italy is better approached by ETS, USA by THieF, the UK by Ridge regression and Russia by the stepwise regression. The summation of all errors gives us an interesting perspective on models with overall stable performance. The stepwise regression has the lowest sum, which indicates that the AIC is a powerful tool for building regression models. What needs to be mentioned is that the model that accounts for the trend differences accumulates the highest error values in the Russian time series. Furthermore, comparing the results between stepwise model and “difference model” one can deduct that it is better choice for time series with even a slight presence of trend and seasonality. Last but not least, the THieF approach scores remarkably well across all time series. The results of ThieF appear to be stable even in time series like Russia which are destabilized.

A concern that needs to be addressed in this project is the validity of our model predictions in the years following 2019. The arrival of the Covid-19 pandemic in 2020 caused major disruption in a variety of industries, including tourism, and

presented unprecedented hurdles to forecasting tactics. These difficulties derive principally from the pandemic's sudden and unexpected changes in consumer behavior, travel restrictions, and worldwide economic collapse. Several challenges may occur in the context of visitor flow estimates based on historical data and economic indicators. Because of the abrupt shift in travel behavior and limits, historical patterns and linkages become obsolete. Previous trends and indicators no longer accurately reflect the situation of the tourism industry today or in the future. Furthermore, the pandemic caused enormous economic instability, dramatically affecting GDP, inflation rates, and other economic indicators. As a result, forecasting models based on pre-pandemic data no longer adequately describe the underlying economic conditions.

To address these challenges, the forecasting technique must be modified. Up-to-date data representing current market conditions, travel trends, and consumer attitude should be included to assist capture the continually dynamic scene more precisely. Furthermore, including qualitative insights from industry professionals and stakeholders can provide a deeper understanding of the dynamic tourism sector, adding context to the quantitative forecasting models.

XI. CONCLUSION

The project's findings shed light on the intricacies of predicting tourist flows to Czechia, revealing various stable models. A key takeaway from the project was the recognition that each time series can be more accurately forecasted by different models, owing to its unique properties. This emphasizes the need for a tailored approach for each dataset. The research underscored the significance of integrating real-time data and qualitative insights into forecasting methodologies, highlighting the importance of adaptable and agile strategies in navigating the dynamic tourism landscape. Further research and exploration are recommended to refine and improve the precision and reliability of future tourism flow predictions.

REFERENCES

- [1] C. L. Bunghez, “The importance of tourism to a destination’s economy,” *J. East. Eur. Res. Bus. Econ.*, vol. 2016, pp. 1–9, 2016.
- [2] “World Bank Open Data,” World Bank Open Data. Accessed: Oct. 10, 2023. [Online]. Available: <https://data.worldbank.org>
- [3] A. De Myttenaere, B. Golden, B. Le Grand, and F. Rossi, “Mean absolute percentage error for regression models,” *Neurocomputing*, vol. 192, pp. 38–48, 2016.
- [4] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.
- [5] G. Perone, “Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy,” *Eur. J. Health Econ.*, pp. 1–24, 2021.
- [6] C. Bergmeir and J. M. Benítez, “On the use of cross-validation for time series predictor evaluation,” *Data Min. Softw. Trust.*, vol. 191, pp. 192–213, May 2012, doi: 10.1016/j.ins.2011.12.028.

- [7] R. G. Brown, *Smoothing, forecasting and prediction of discrete time series*. Courier Corporation, 2004.
- [8] C. C. Holt, “Forecasting seasonals and trends by exponentially weighted moving averages,” *Int. J. Forecast.*, vol. 20, no. 1, pp. 5–10, 2004.
- [9] E. S. Gardner, “Exponential smoothing: The state of the art—Part II,” *Int. J. Forecast.*, vol. 22, no. 4, pp. 637–666, Oct. 2006, doi: 10.1016/j.ijforecast.2006.03.005.
- [10] L. J. Tashman, “Out-of-sample tests of forecasting accuracy: an analysis and review,” *Int. J. Forecast.*, vol. 16, no. 4, pp. 437–450, 2000.
- [11] V. Cerqueira, L. Torgo, and I. Mozetič, “Evaluating time series forecasting models: An empirical study on performance estimation methods,” *Mach. Learn.*, vol. 109, pp. 1997–2028, 2020.
- [12] A. S. Wilkins, “To lag or not to lag?: Re-evaluating the use of lagged dependent variables in regression analysis,” *Polit. Sci. Res. Methods*, vol. 6, no. 2, pp. 393–411, 2018.
- [13] T. Yamashita, K. Yamashita, and R. Kamimura, “A stepwise AIC method for variable selection in linear regression,” *Commun. Stat. Methods*, vol. 36, no. 13, pp. 2395–2403, 2007.
- [14] A. Dokumentov and R. J. Hyndman, “STR: Seasonal-trend decomposition using regression,” *ArXiv Prepr. ArXiv200905894*, 2020.
- [15] E. Ostertagová and O. Ostertag, “Regression analysis and seasonal adjustment of time series,” *J. Autom. Control*, vol. 3, no. 3, pp. 118–121, 2015.
- [16] Z. Wu, N. E. Huang, S. R. Long, and C.-K. Peng, “On the trend, detrending, and variability of nonlinear and nonstationary time series,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 38, pp. 14889–14894, 2007.
- [17] J. Ranstam and J. Cook, “LASSO regression,” *J. Br. Surg.*, vol. 105, no. 10, pp. 1348–1348, 2018.
- [18] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.
- [19] A. E. Hoerl and R. W. Kennard, “Ridge regression: applications to nonorthogonal problems,” *Technometrics*, vol. 12, no. 1, pp. 69–82, 1970.
- [20] G. C. McDonald, “Ridge regression,” *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 1, no. 1, pp. 93–100, 2009.
- [21] D. Gabauer, R. Gupta, H. A. Marfatia, and S. M. Miller, “Estimating us housing price network connectedness: Evidence from dynamic elastic net, lasso, and ridge vector autoregressive models,” *Int. Rev. Econ. Finance*, 2023.
- [22] G. Athanasopoulos, R. J. Hyndman, N. Kourentzes, and F. Petropoulos, “Forecasting with temporal hierarchies,” *Eur. J. Oper. Res.*, vol. 262, no. 1, pp. 60–74, 2017.

APPENDIX

Table 1. MAPE Results

	Germany	Italy	USA	UK	Russia	Sum of MAPE Errors
ETS	3.23 % (MAA)	3.86 % (MAA)	2.72 % (MMM)	7.11 % (MAA)	5.80 % (NAÏVE)	4.5 %
Step Regression	2.81 %	4.29 %	3.02 %	4.28 %	2.23 %	3.3 %
Modeled Differences	2.70 %	4.43 %	2.78 %	5.43 %	14.83 %	6 % (more than 3% of that error is from the Russian ts)
Lasso	5.96 %	5.38 %	1.38 %	3.91 %	6.93 %	4.7 %
Ridge	5.26 %	4.78 %	4.13 %	3.85 %	7.84 %	5.1 %
THieF (Arima)	4.42 %	4.15 %	1.09 %	4.77 %	4.40 %	3.7 %
IMF Regression	> 40 %	> 40 %	> 40 %	> 40 %	> 40 %	-

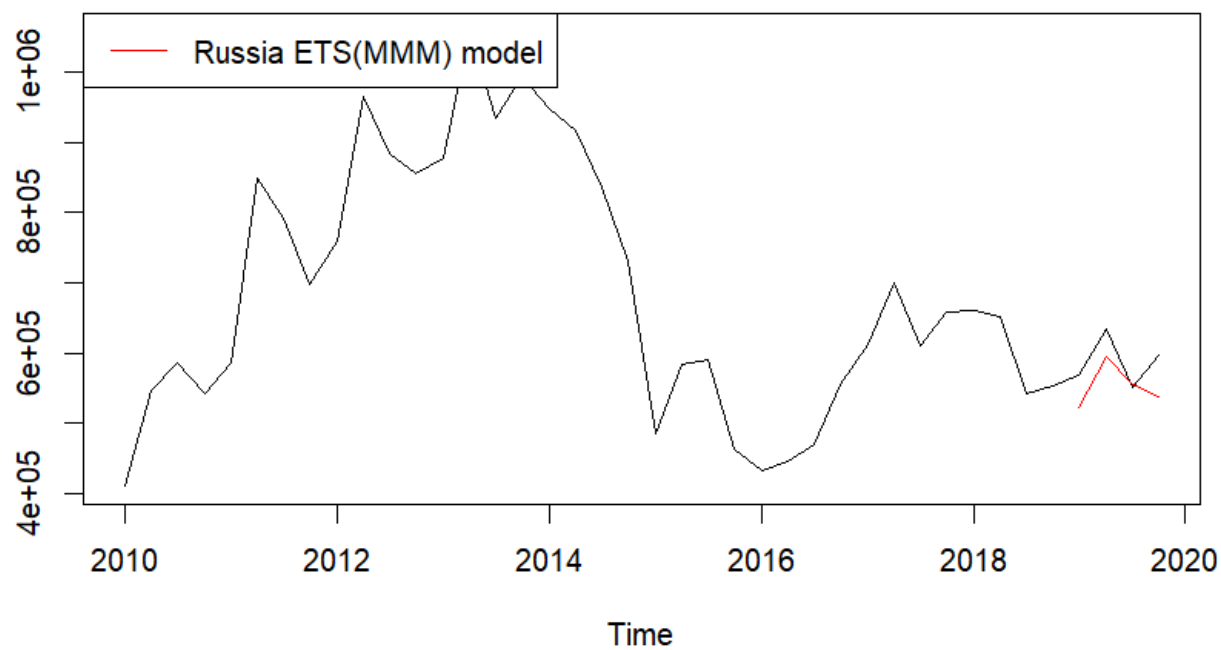


Figure 5. ETS(MMM) Russia

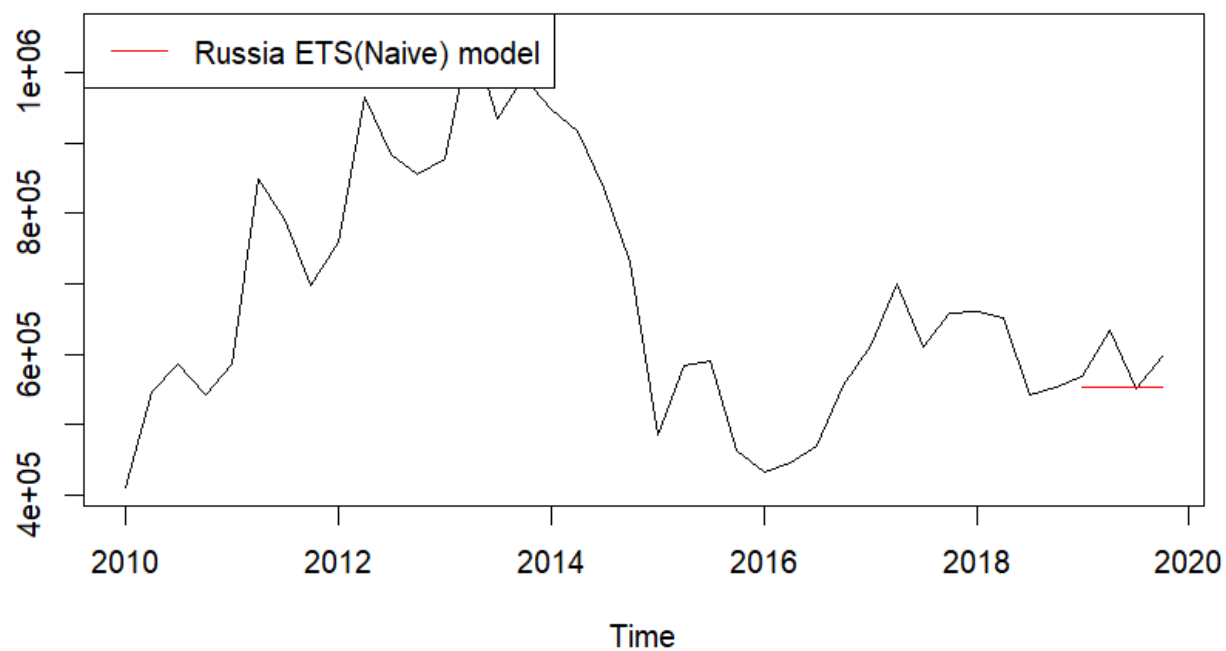


Figure 6. ETS(Naive) Russia

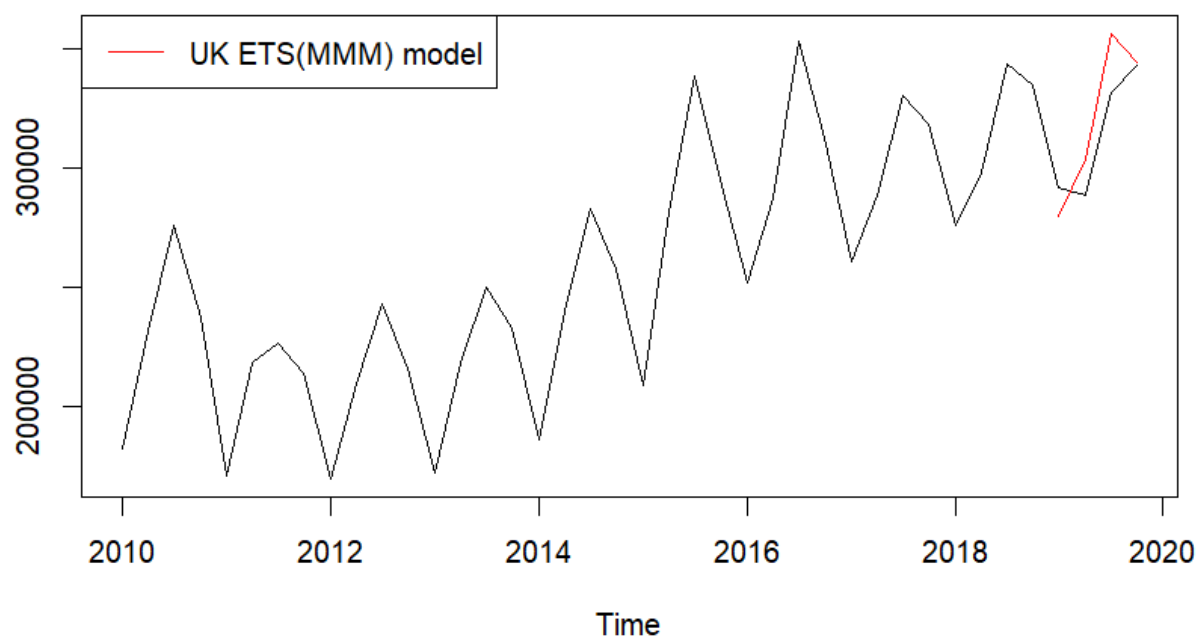


Figure 7. ETS(MMM) UK

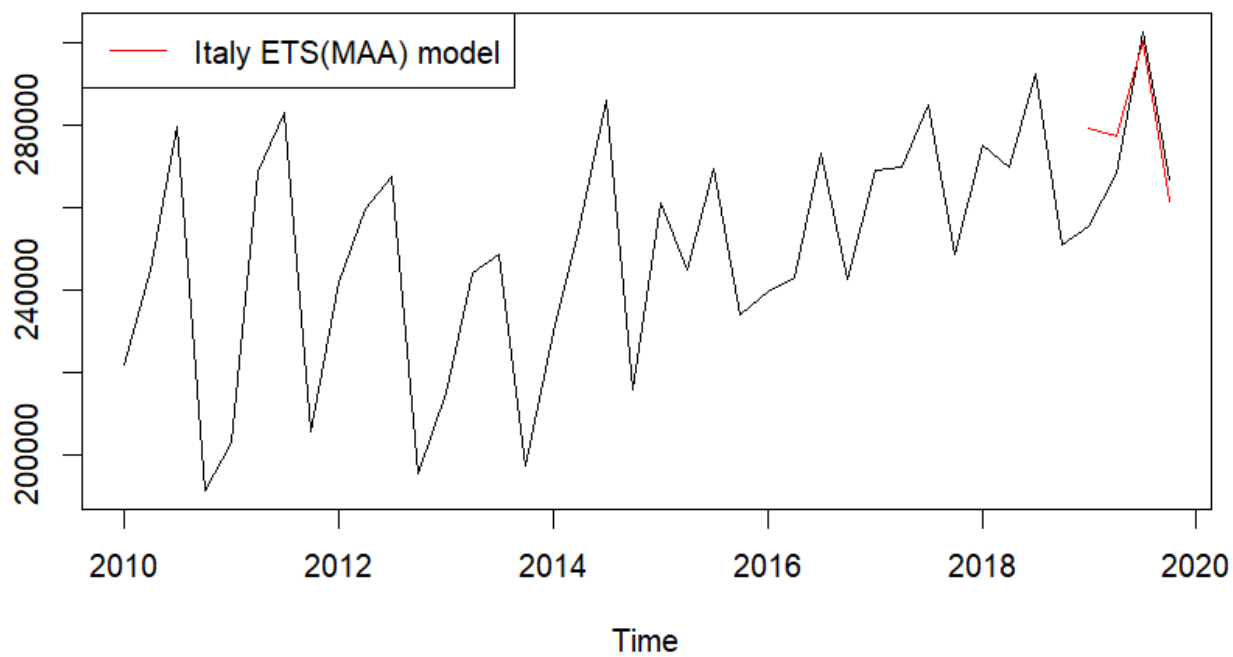


Figure 8. ETS(MAA) Italy

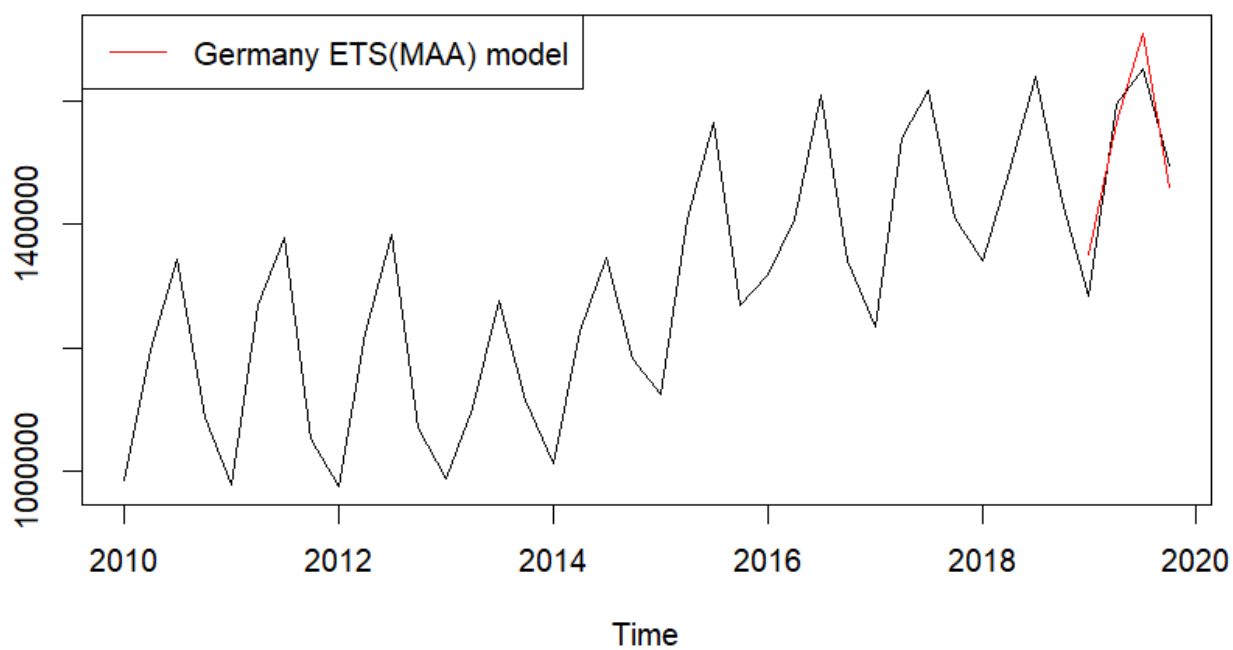


Figure 9. ETS(MAA) Germany

Stepwise Regression USA

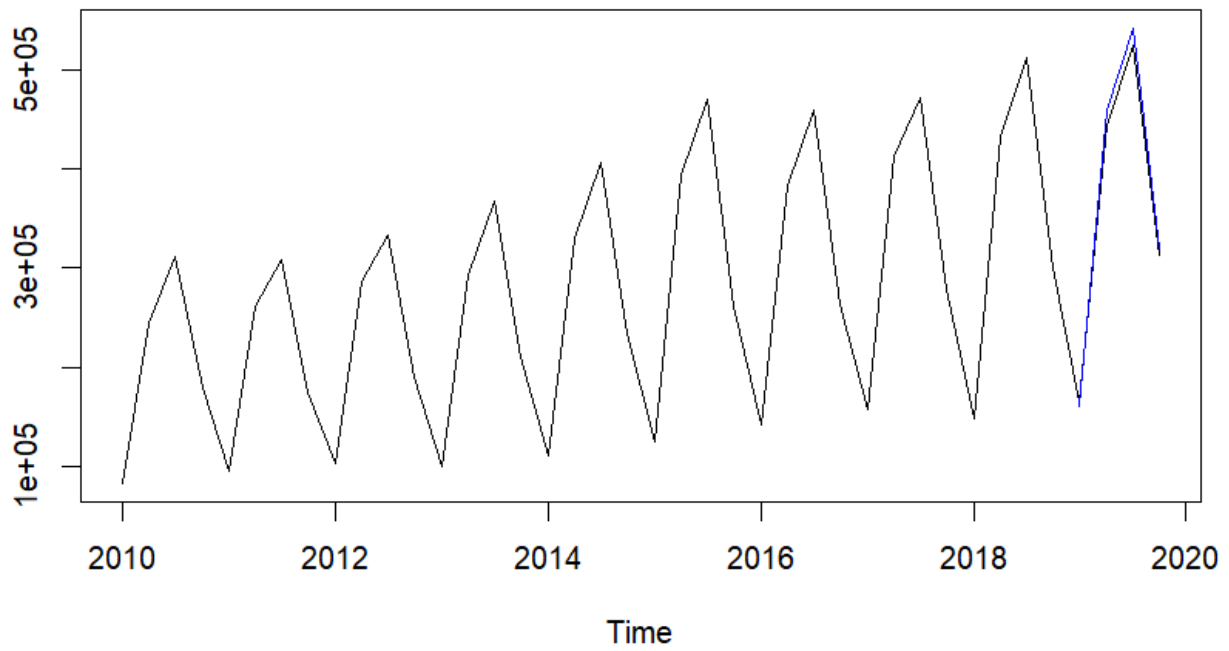


Figure 10. Stepwise USA

Stepwise Regression UK

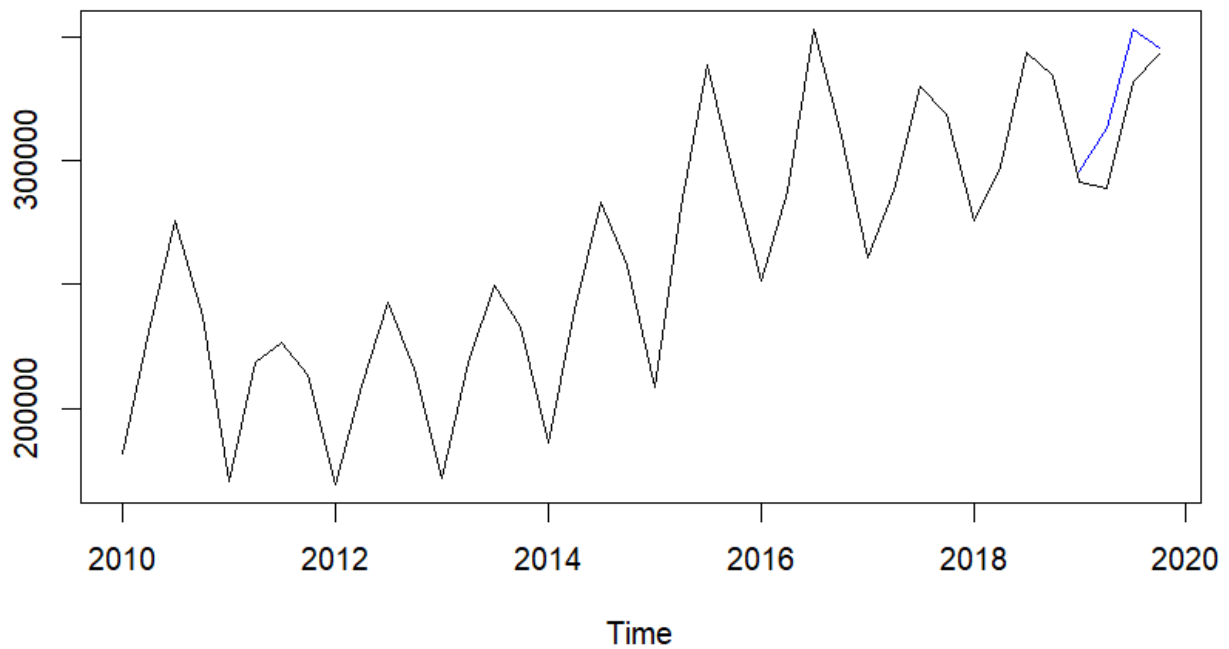


Figure 11. Stepwise UK

Stepwise Regression Russia

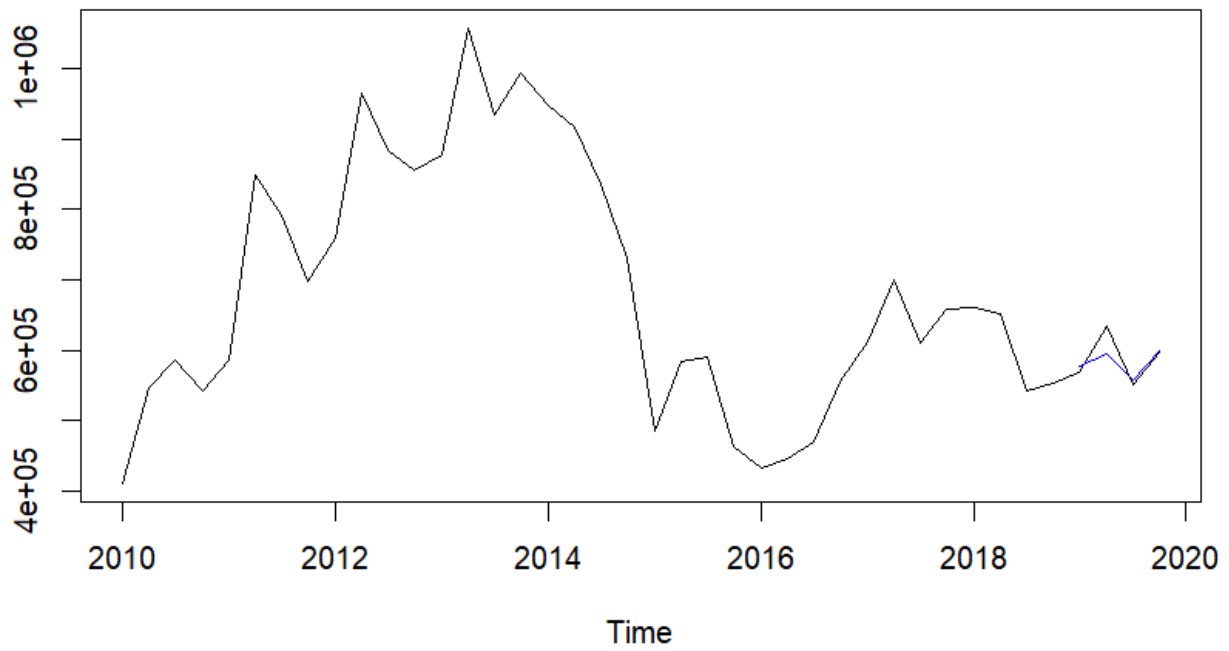


Figure 12. Stepwise Russia

Stepwise Regression Italy

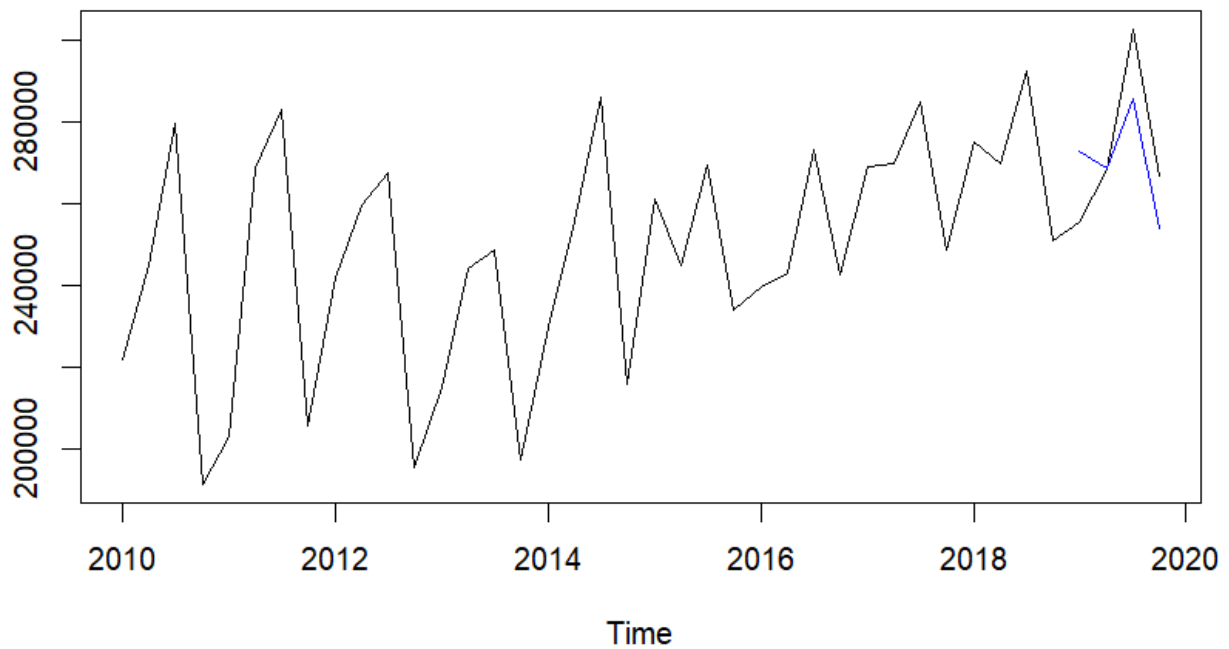


Figure 13 Stepwise Italy

Stepwise Regression Germany

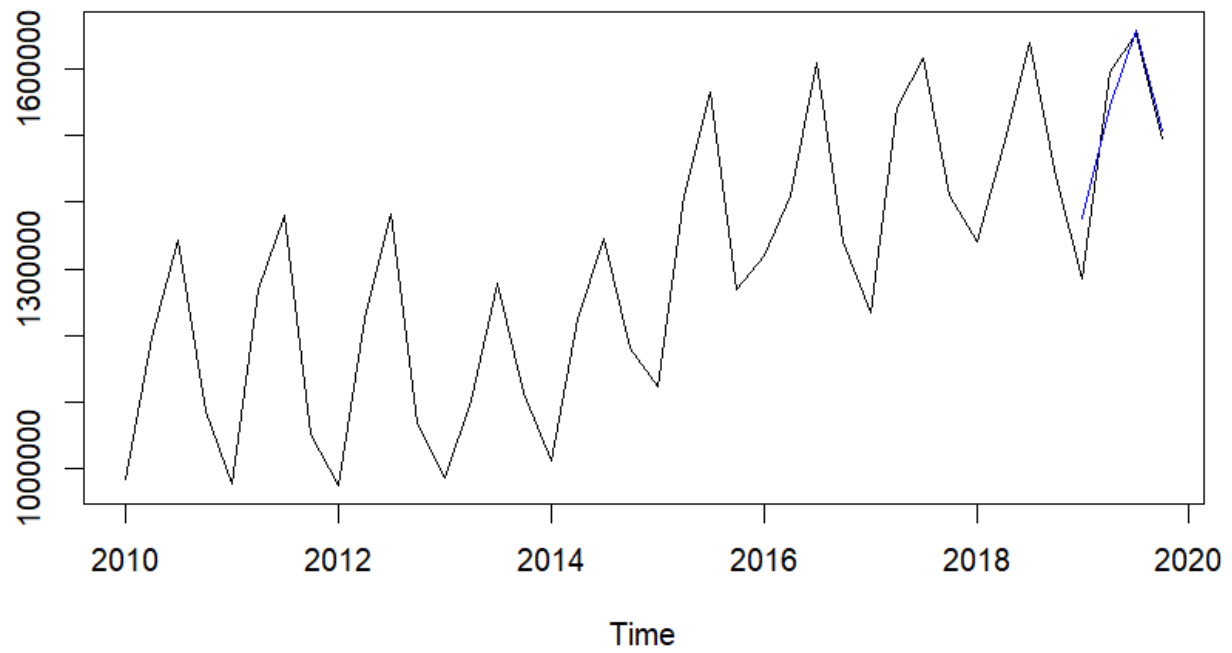


Figure 14. Stepwise Germany

Italy

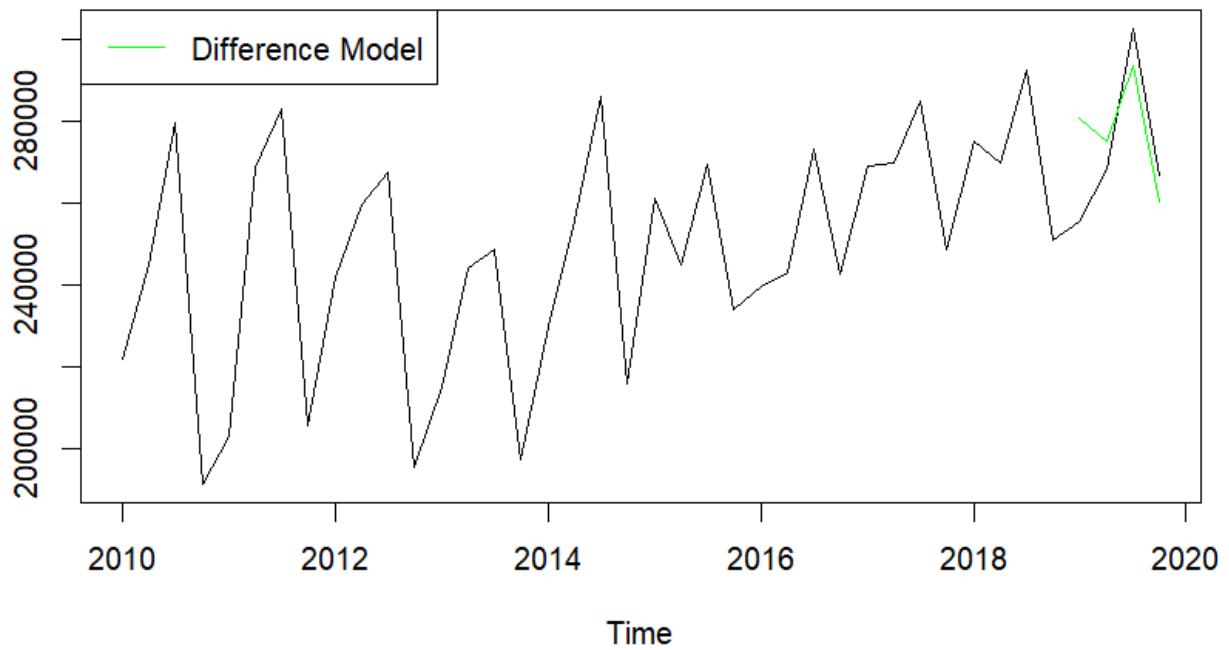


Figure 15. Difference model Italy

Russia

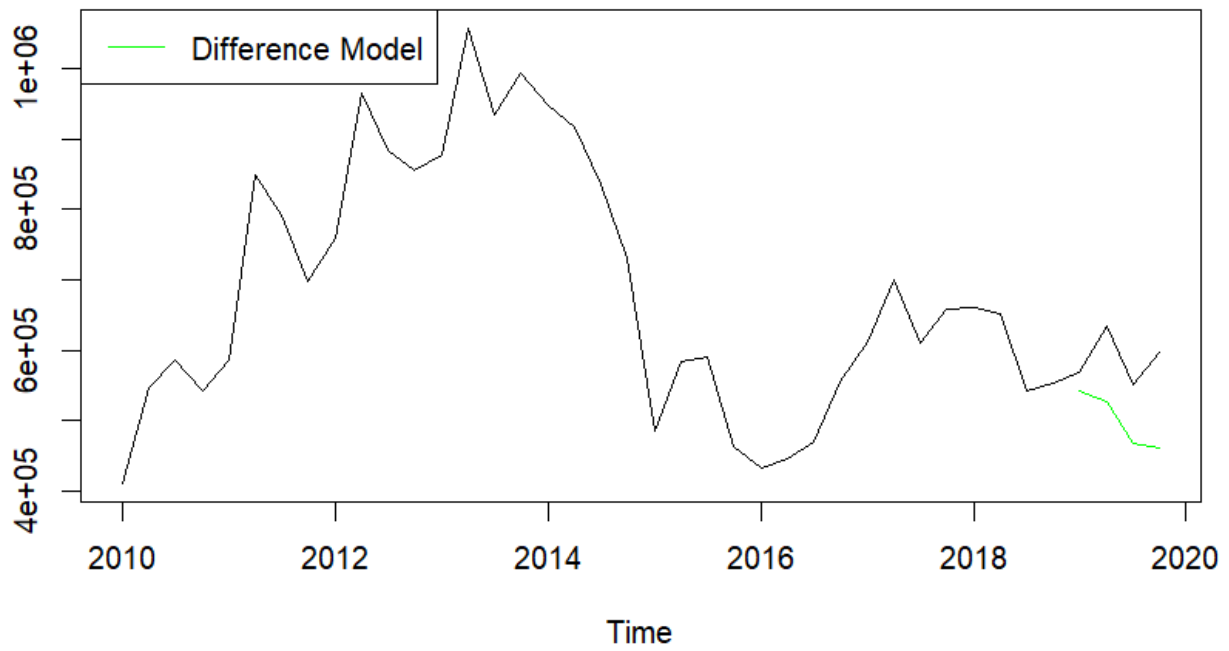


Figure 16. Difference model Russia

UK

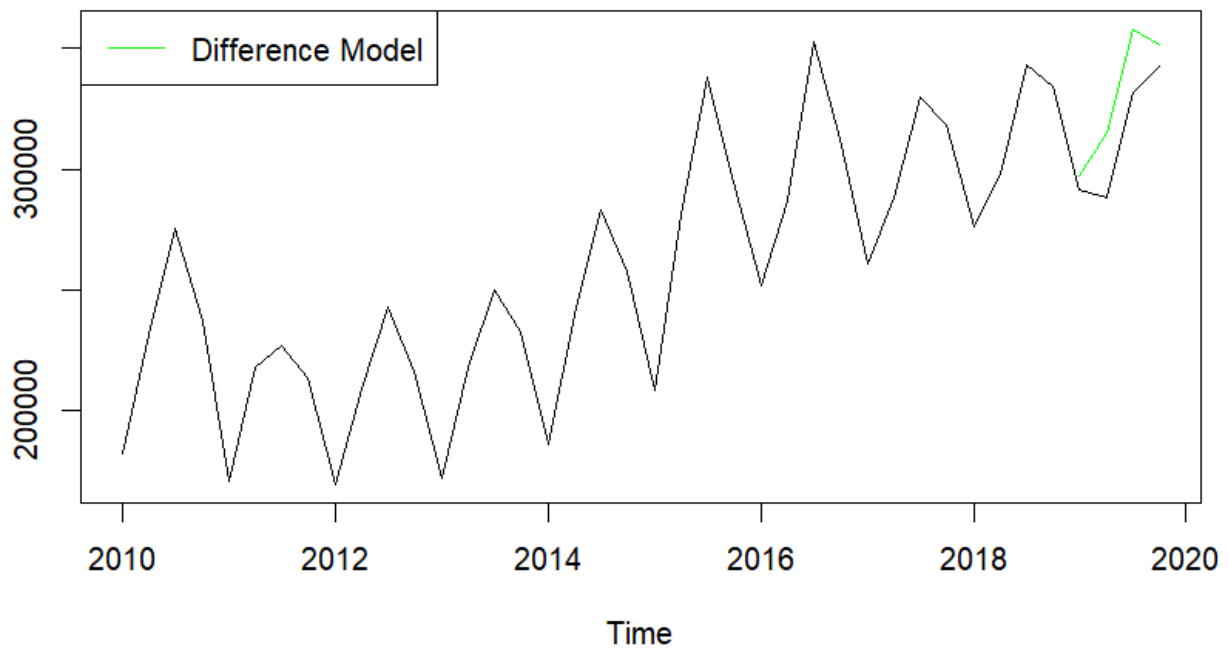


Figure 17. Difference model UK

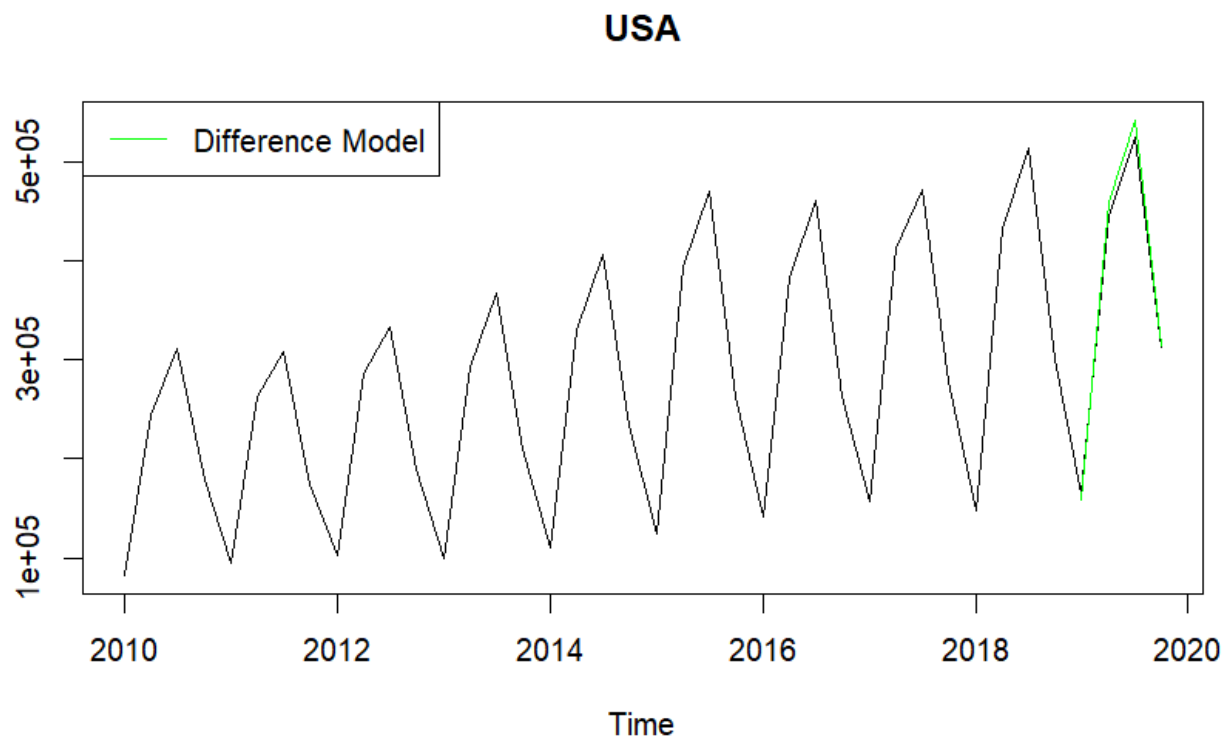


Figure 18. Difference model USA

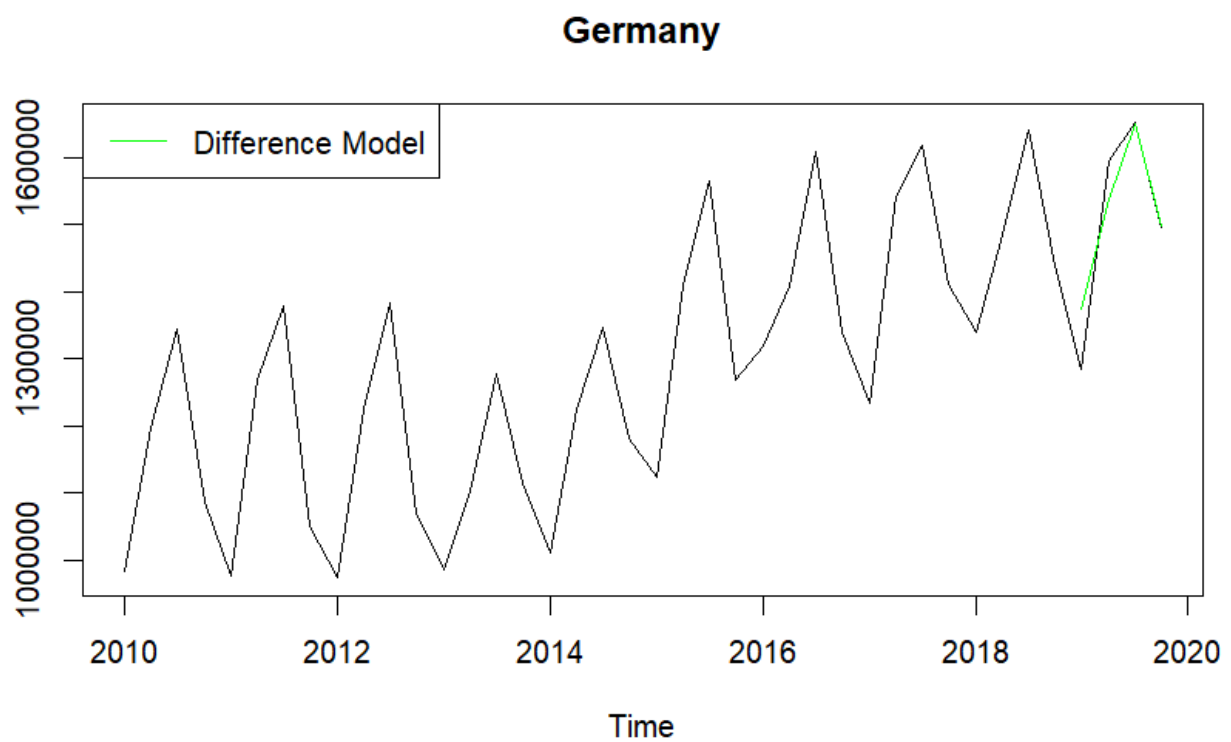


Figure 19. Difference model Germany

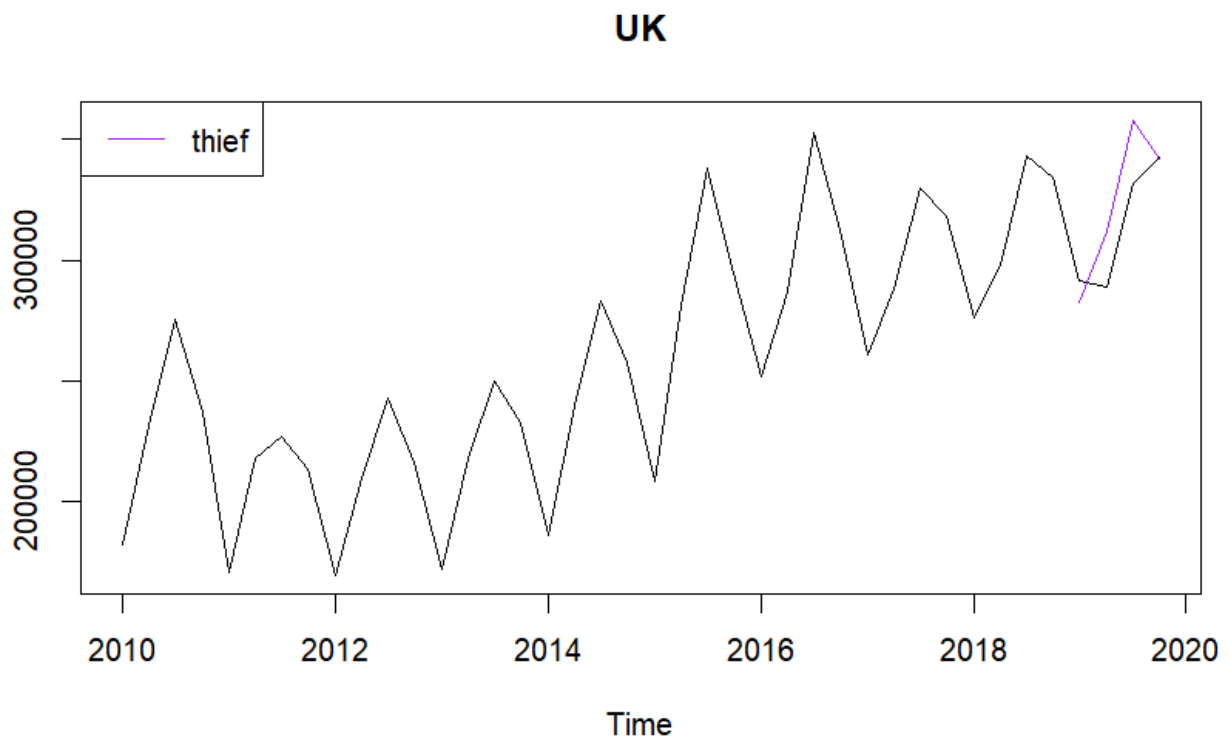


Figure 20. THieF UK

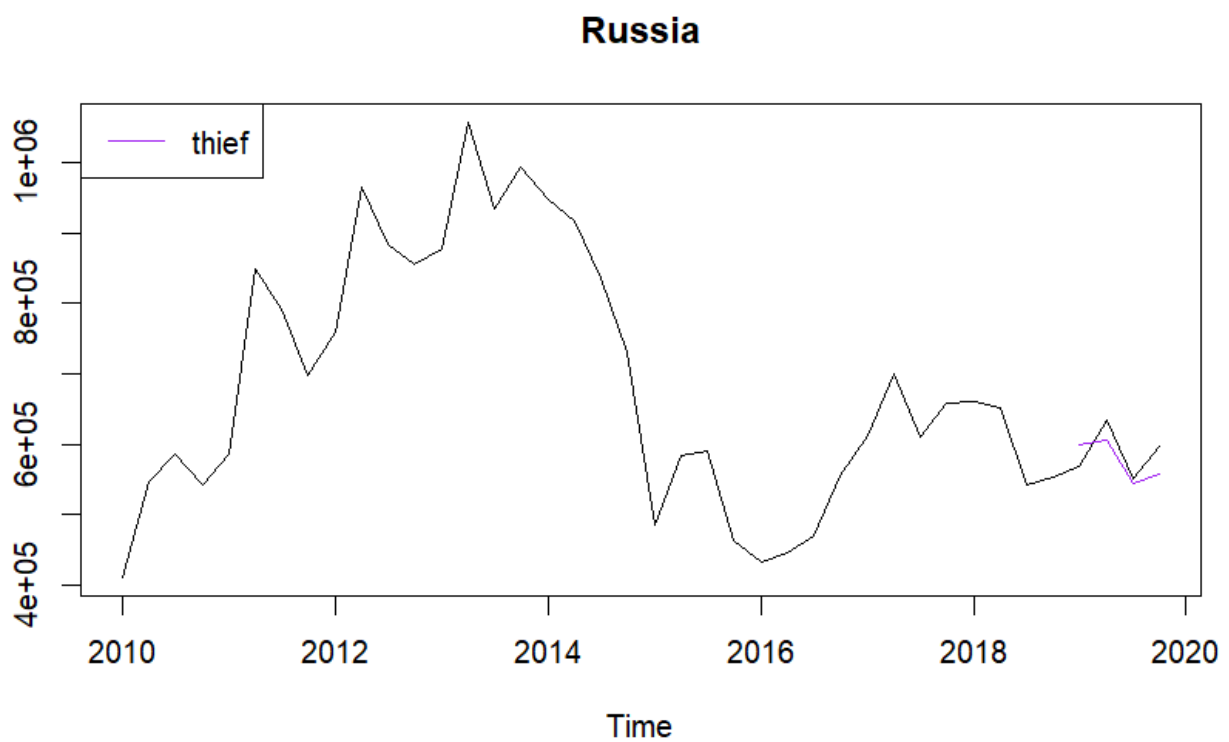


Figure 21. THieF Russia

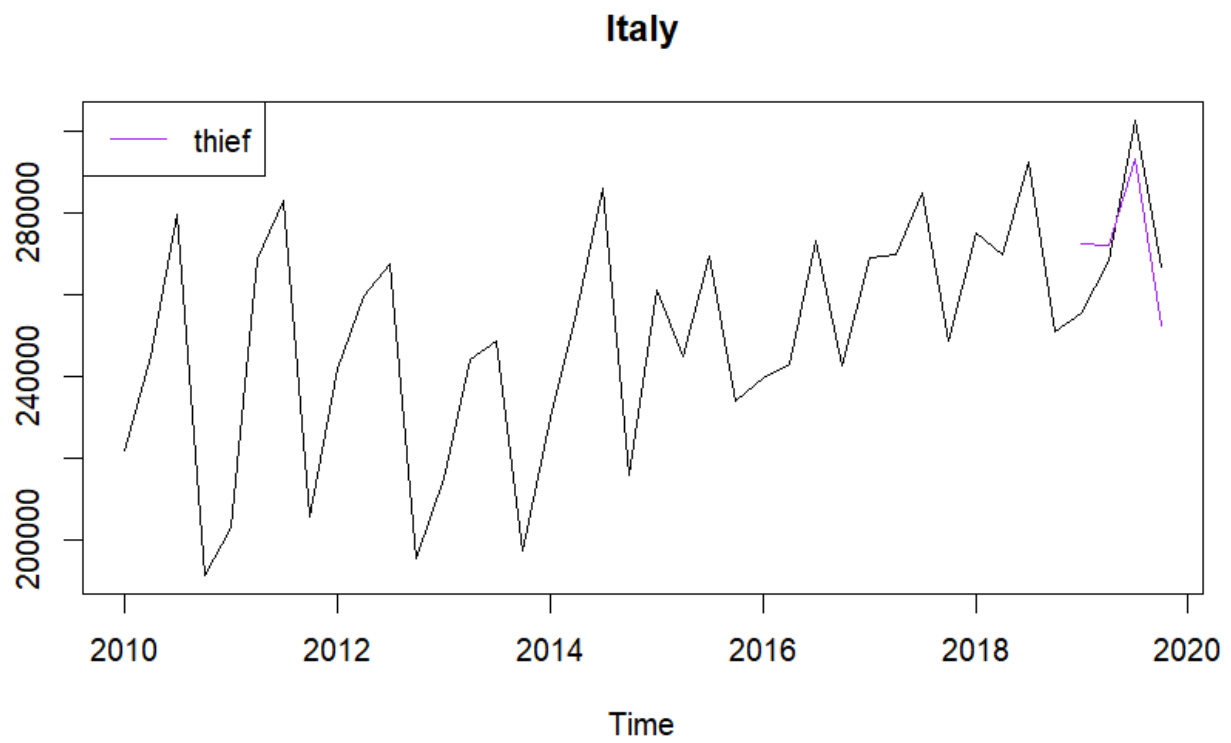


Figure 22. THieF Italy

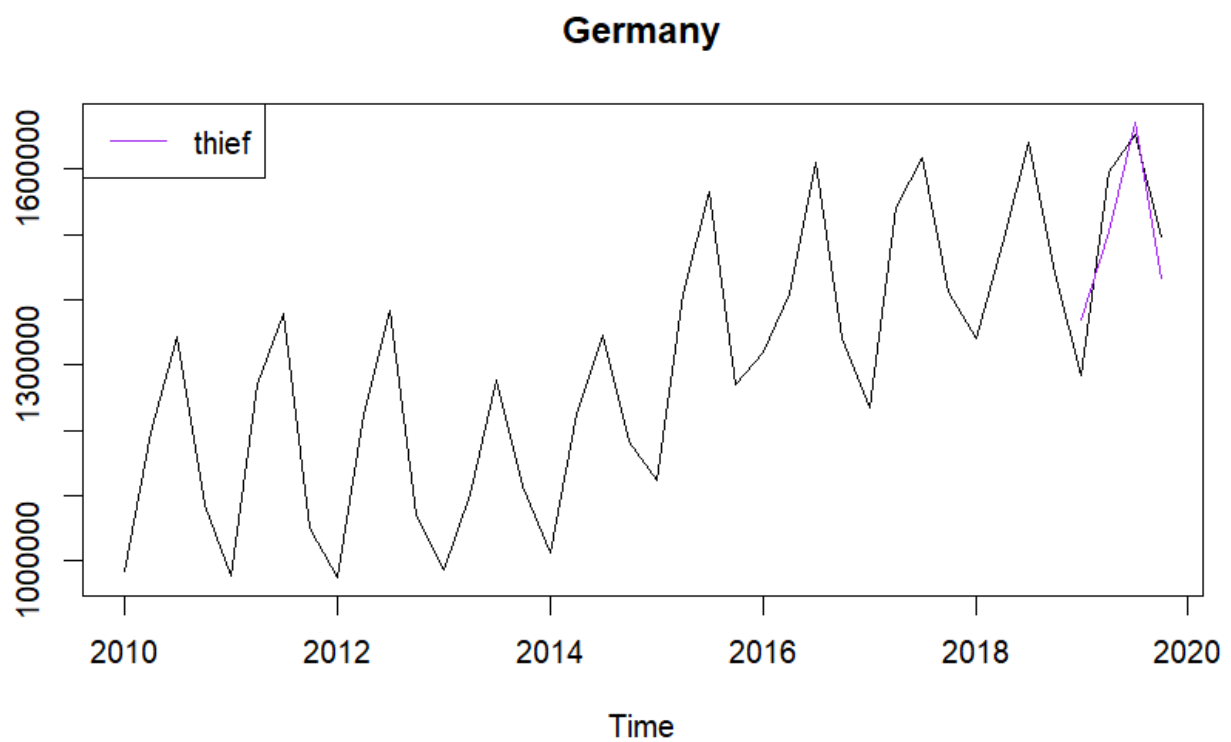


Figure 23. THieF Germany

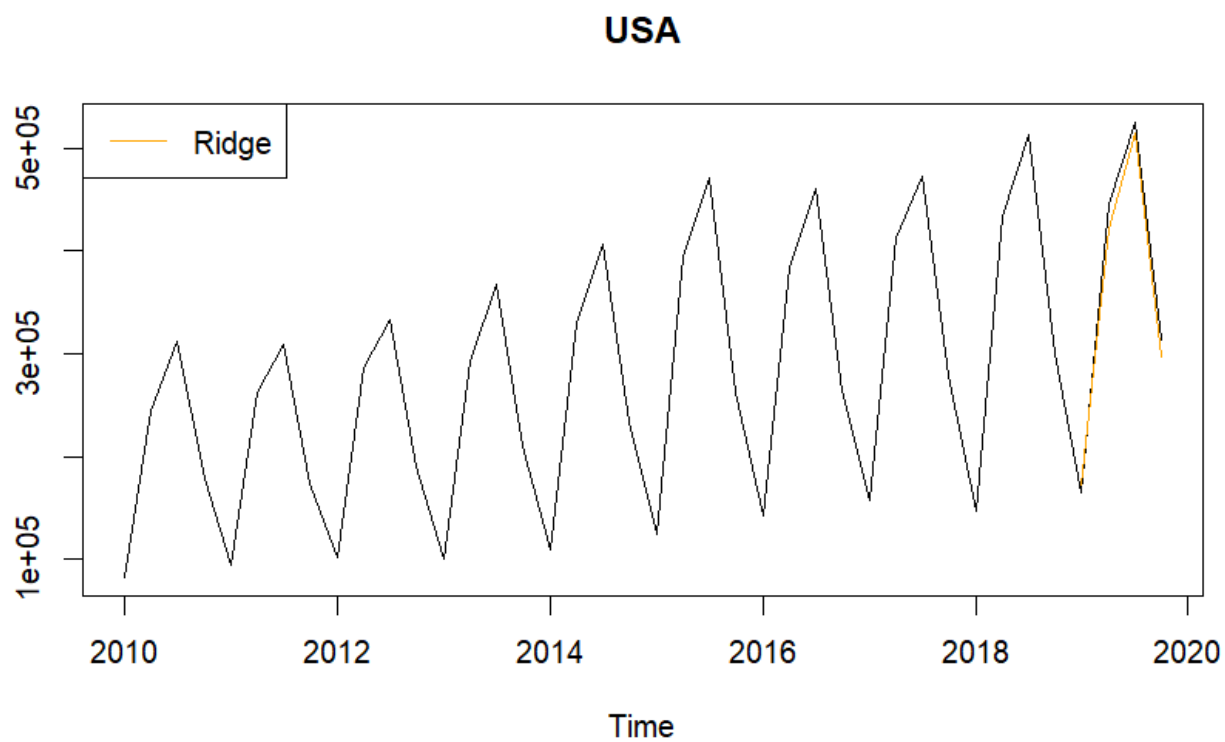


Figure 24. Ridge USA

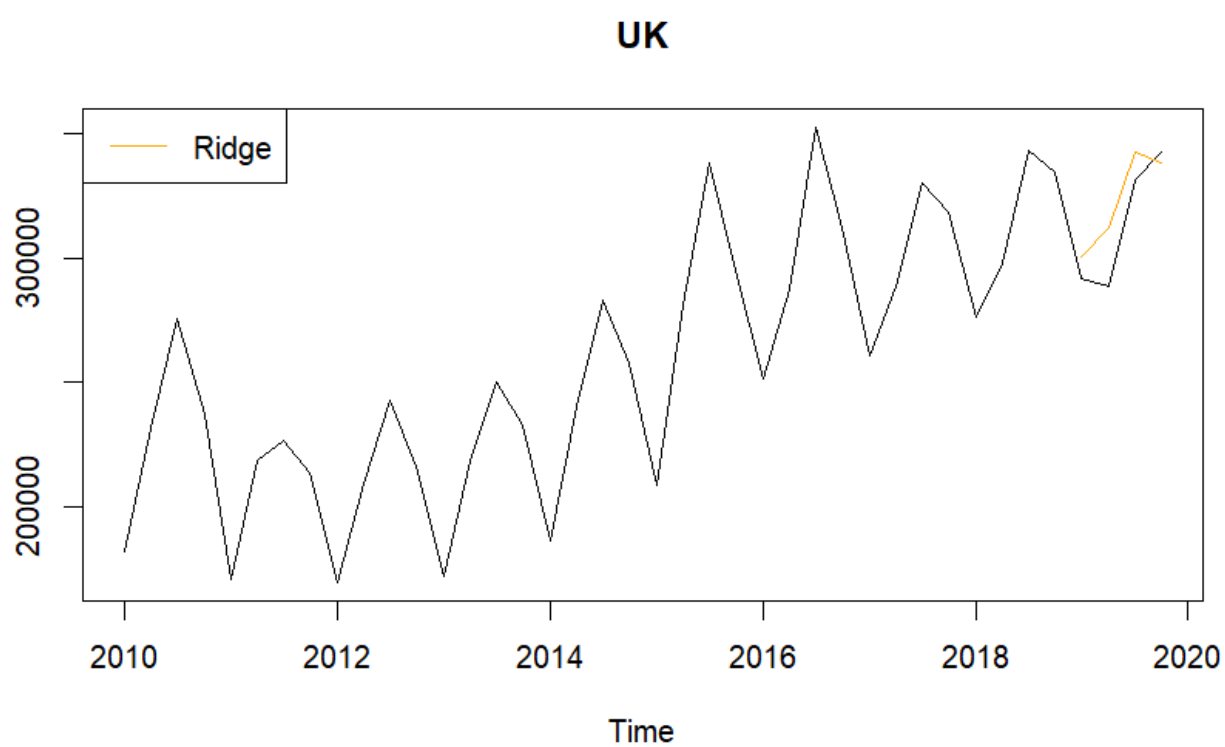


Figure 25. Ridge UK

Russia

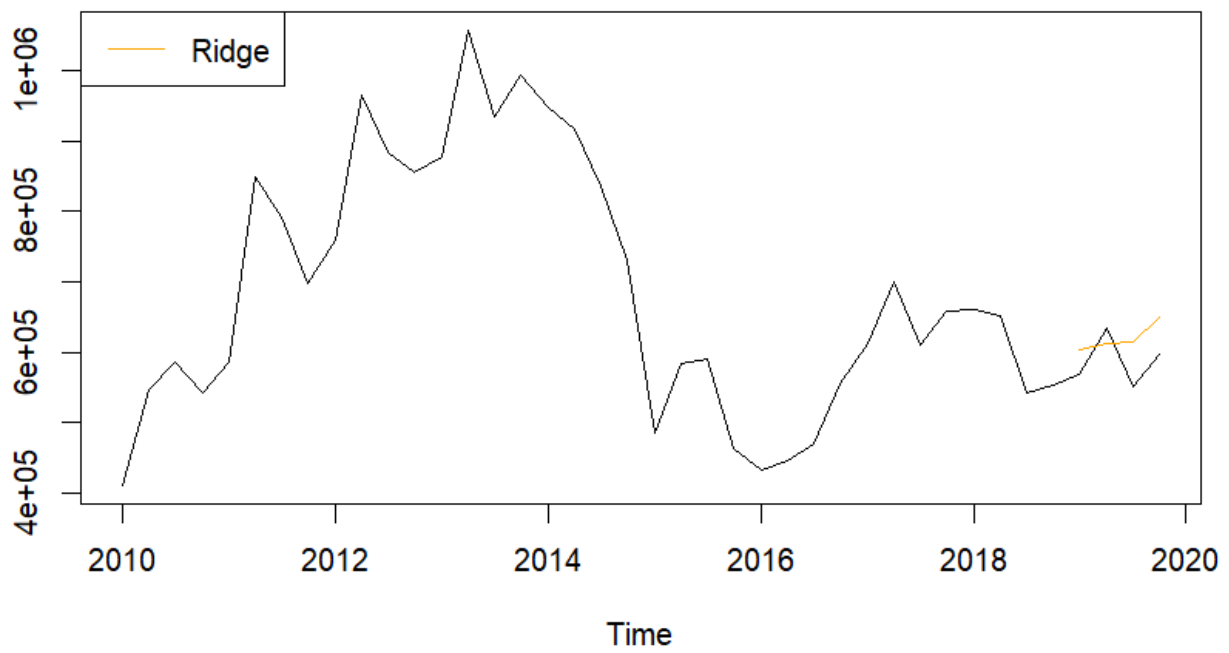


Figure 26. Ridge Russia

Italy

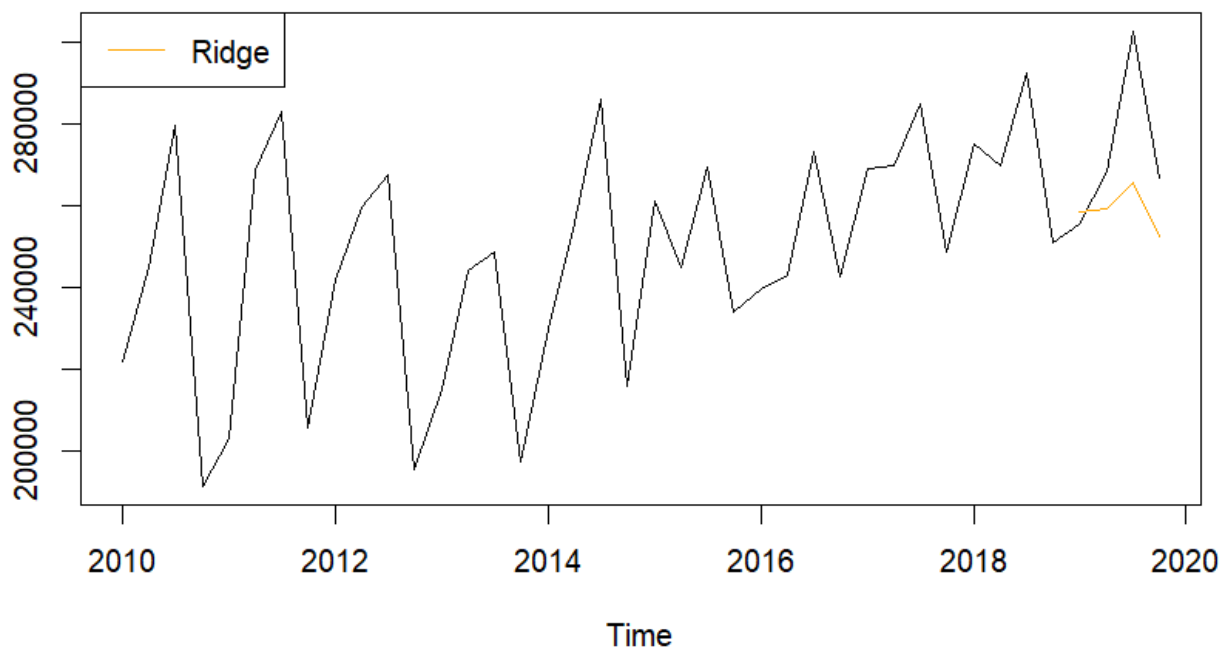


Figure 27. Ridge Italy

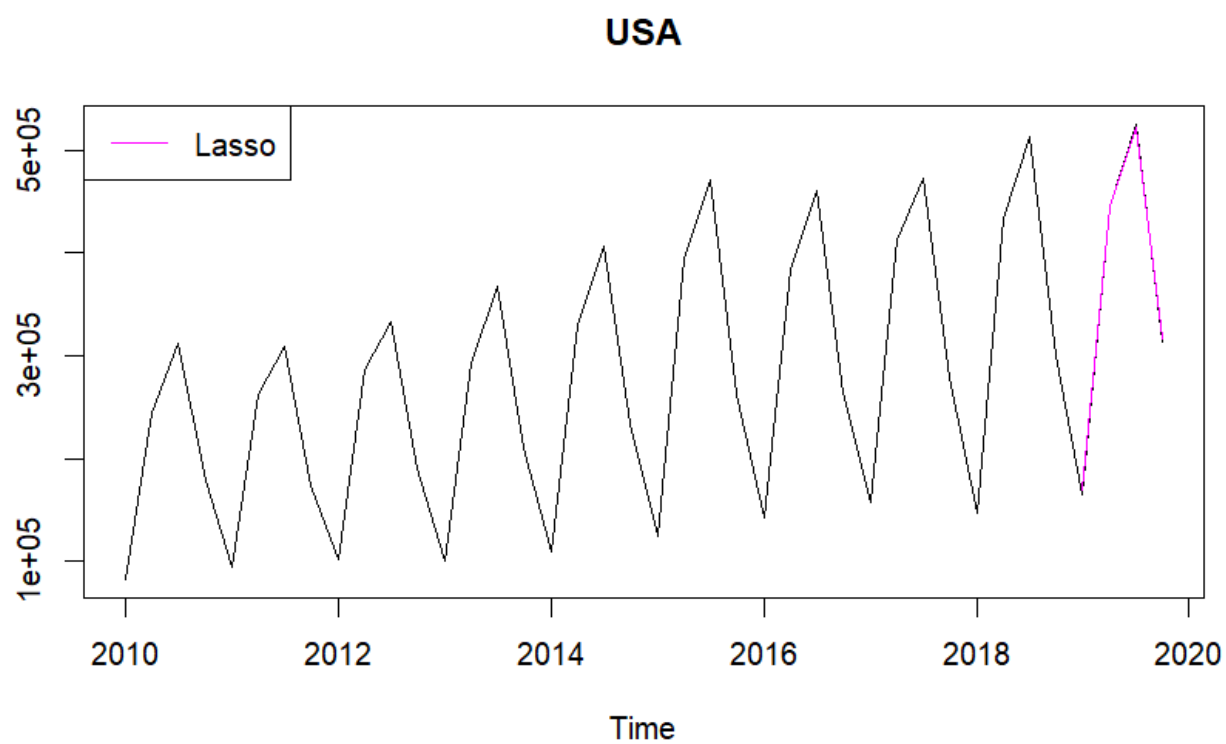


Figure 28. Lasso USA

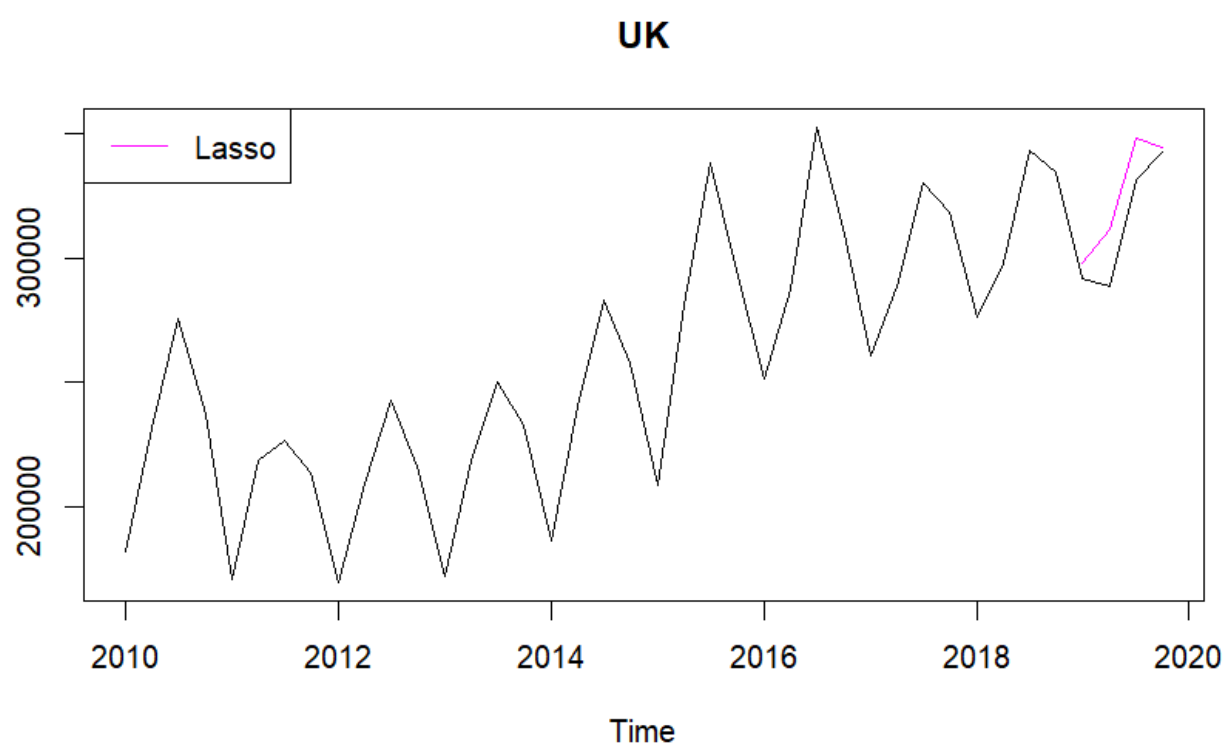


Figure 29. Lasso UK

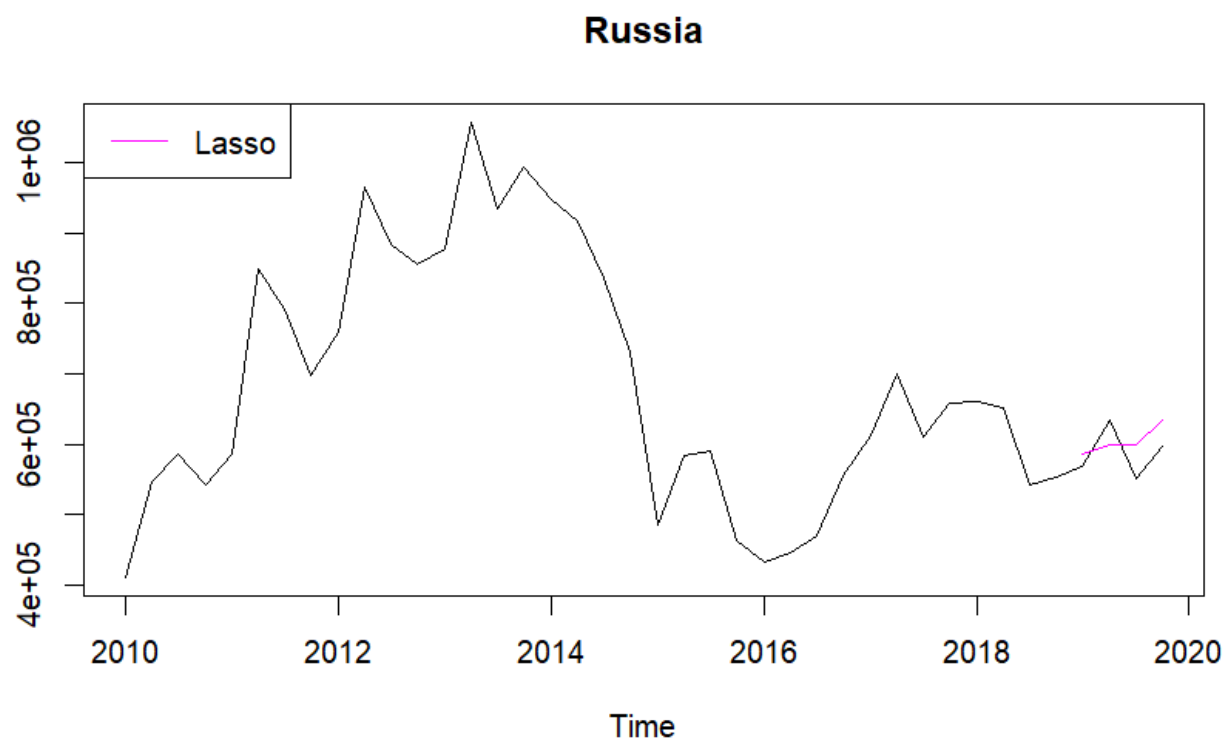


Figure 30. Lasso Russia

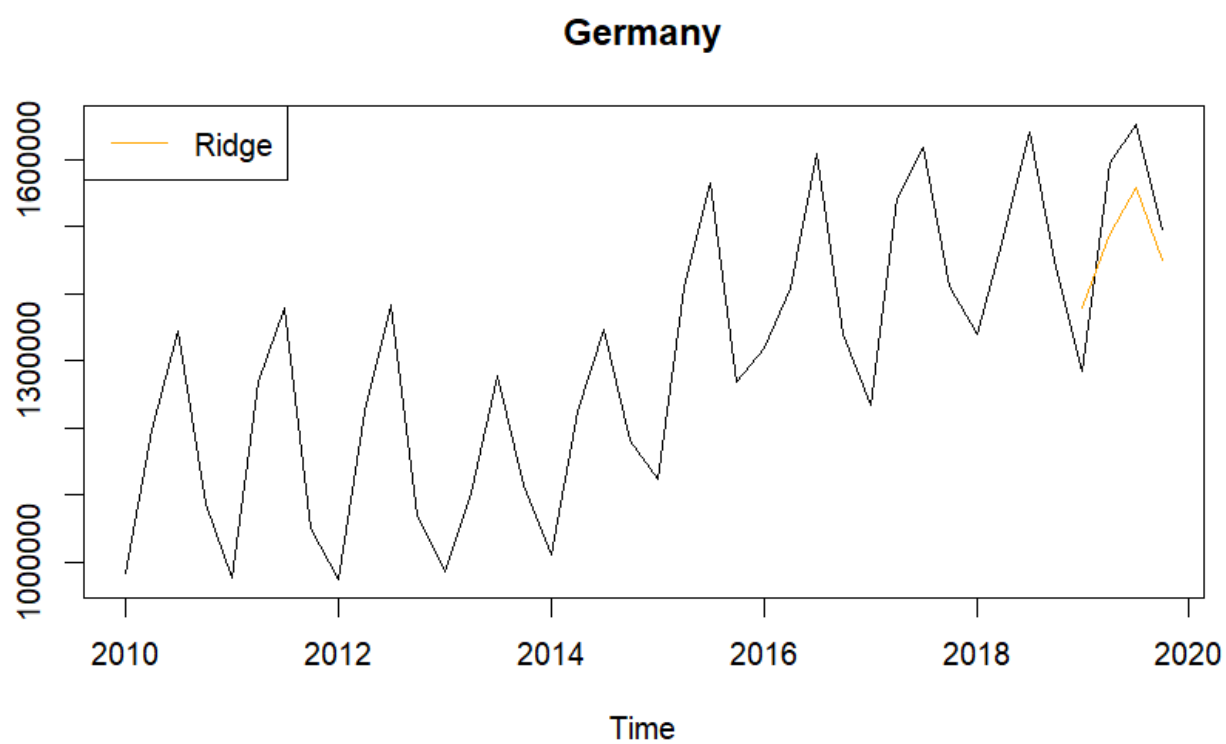


Figure 31. Ridge Germany

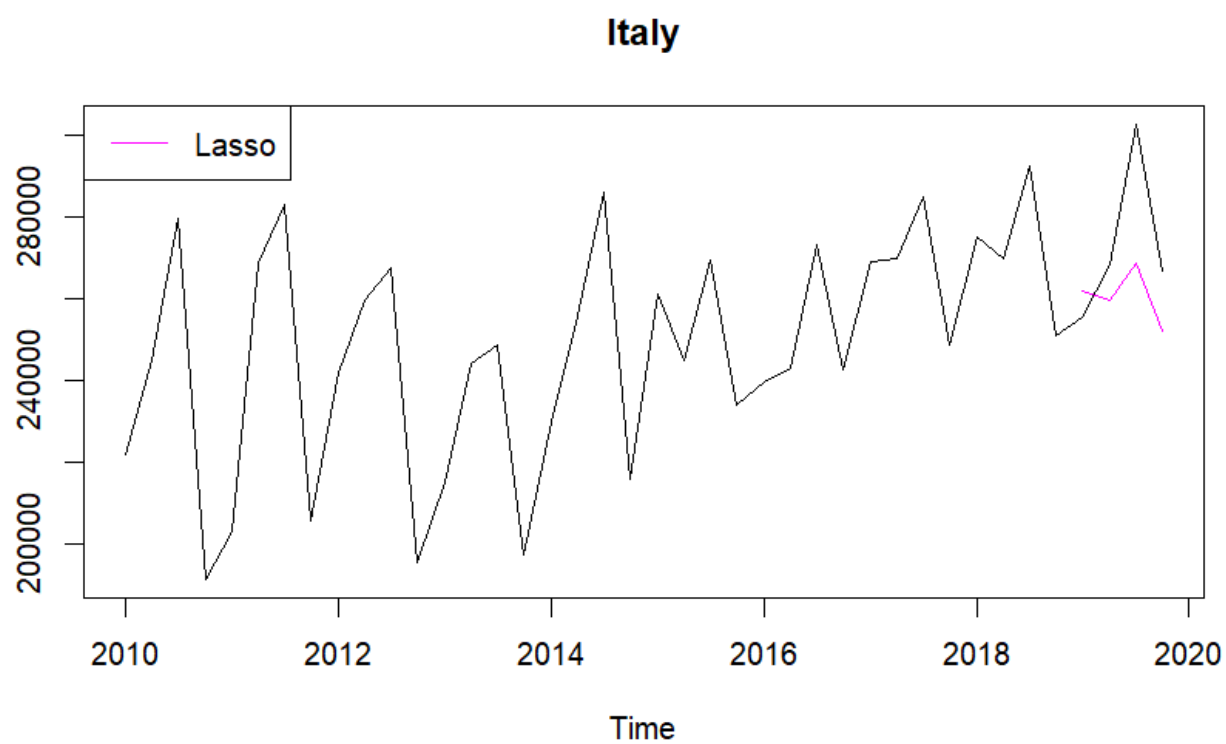


Figure 32. Lasso Italy

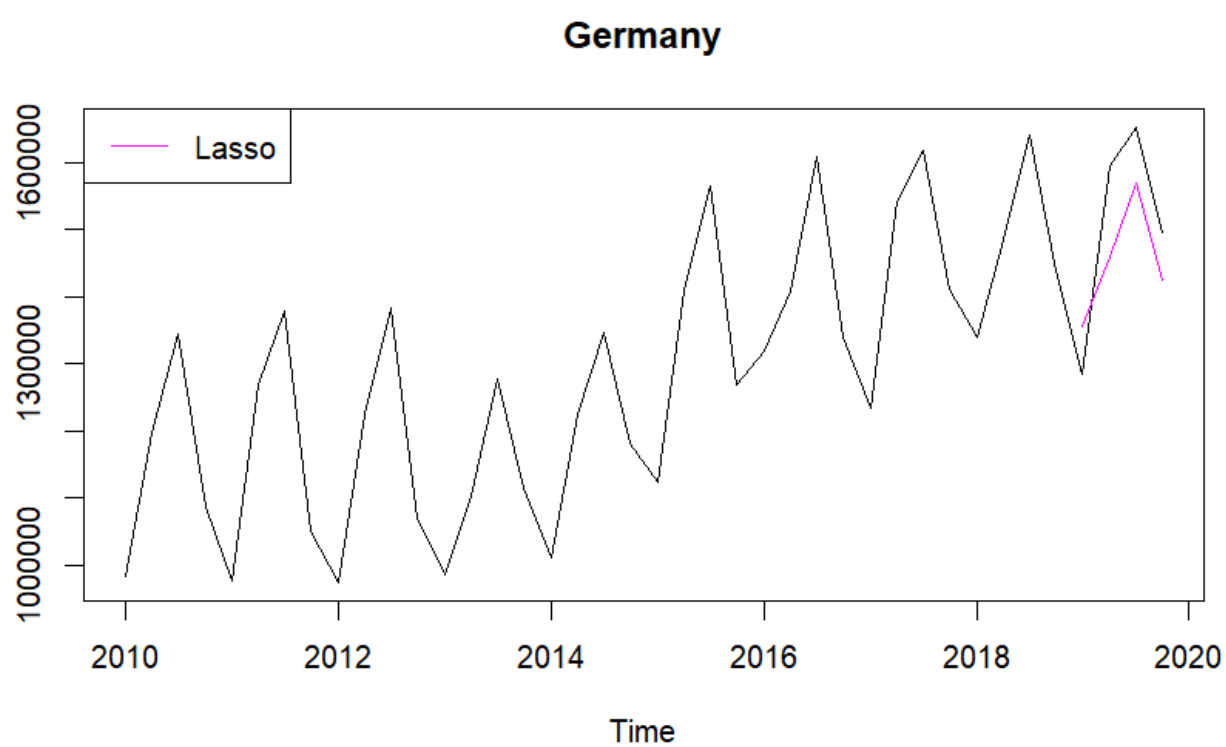


Figure 33. Lasso Germany

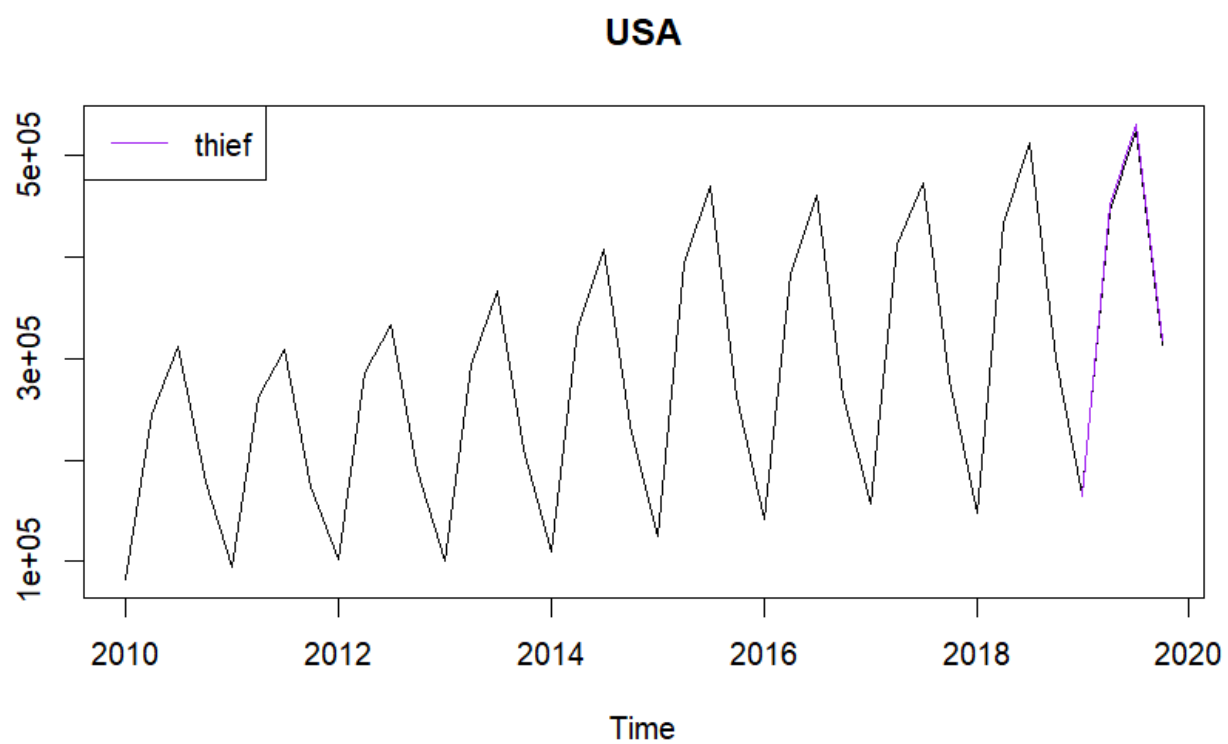


Figure 34. THieF USA