Predictive Analytics

Executive Summary: Predictive Analytics in Exchange Rate Forecasting by Jørgen Leiros

This project focused on time series forecasting of the NOK/EUR exchange rate using regression models such as ARIMA and ETS. The workflow included time series decomposition, structural break analysis, and tests for heteroskedasticity. Based on these findings, pre-financial crisis data was excluded, and first-order differencing was applied to stabilize the series. The project provided valuable experience in predictive modeling and highlighted the limitations of traditional regression approaches compared to multivariate machine learning techniques.

Key Focus Areas:

1. **Objective and Data Preparation** – Jørgen analyzed historical **NOK/EUR exchange rate data**, applying **data preprocessing techniques** such as **structural break detection, seasonality analysis, and missing value imputation** to ensure data reliability.
2. **Comparison of Forecasting Models:**
    a. **ARIMA (AutoRegressive Integrated Moving Average)** – A widely used statistical model that captures linear relationships and autocorrelations.
    b. **TBATS (Trigonometric Box-Cox ARMA Trend and Seasonal Model)** – Handles complex seasonality and trend variations in time-series data.
    c. **NNAR (Neural Network AutoRegressive Model)** – A machine learning approach that captures **nonlinear dependencies and hidden patterns** in financial time series.
3. **Model Evaluation and Performance Metrics:**
    a. Performance was assessed using **Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE)**.
    b. **NNAR significantly outperformed ARIMA and TBATS**, demonstrating **superior accuracy in short-term exchange rate forecasting** due to its ability to capture **nonlinear trends**.
    c. **ARIMA and TBATS models behaved similarly to a naïve forecast**, struggling with the volatility of the exchange rate.
4. **Key Findings and Implications:**
    a. **Neural network-based models (NNAR) are more effective for short-term currency forecasting**, as they can capture subtle patterns in exchange rate fluctuations.

b. **Traditional statistical models (ARIMA, TBATS) perform well in stable financial conditions** but struggle with sudden structural shifts and high volatility.
c. The study highlights the **importance of combining statistical and machine learning approaches** for more **robust financial forecasting strategies**.

Strategic Insights and Future Work:

- **Enhance time-series forecasting** by incorporating external macroeconomic indicators (e.g., interest rates, trade balances).
- **Explore deep learning architectures (e.g., LSTM, GRU)** to improve long-term exchange rate predictions.
- **Develop hybrid models** that integrate **ARIMA's interpretability** with **NNAR's predictive power** to optimize financial decision-making.

Through this research, Jørgen has demonstrated **expertise in predictive analytics, timeseries modeling, and financial forecasting**, contributing valuable insights for businesses, policymakers, and financial analysts seeking to **enhance exchange rate prediction strategies**.

Comparing Time Series Models

1. Introduction

Exchange rates express the relative value of a currency in terms of another currency, and it can reflect and affect the economic situation of a country. The exchange rate between two currencies is determined by the foreign exchange market in operation and can be highly volatile, changing multiple times a day (Lowry, 2024). Being able to predict the exchange rate between currencies can be useful for many reasons, especially if you make frequent cross-border payments. With robust predictions of future exchange rates, one can save and even earn money through better informed decisions.

Time series modelling can leverage historical data to forecast future values by considering trends, cycles and seasonal variations (Hyndman & Athanasopoulos, 2021). Comparing different time series models can reveal which techniques are most effective for predicting exchange rates, providing insights into the strengths and weaknesses of each approach. This comparison can highlight key factors and patterns that are crucial for accurate forecasts, such as the importance of capturing trends and handling nonlinear relationships. This leads me to the research question:

In this paper, I aim to compare predictions based on historical exchange rate data, using AutoregRessive Integrated Moving Average (ARIMA), Exponential Smoothing (ES), and Neural Network AutoRegressive (NNAR). Ultimately, understanding which models perform best can provide robust exchange rate prediction and guide further improvements and model development.

2. Related Work

There are many possible approaches to forecasting exchange rates, and a lot of data is available to help these predictions. Methods like Purchasing Power Parity (PPP), Interest Rate Parity (IRP) and Behavioural Equilibrium Exchange Rate (BEER) consider factors such as the price of common goods, interest rates, productivity, government debt and terms of trade to produce exchange rate predictions (European Central Bank, 2017). While these models can perform well under specific conditions, it is hard to produce models that consistently outperforms a simple random walk model, which suggests that the current exchange rate represents the best prediction of future exchange rates. Many researchers have aimed to identify the optimal approach to exchange rate prediction, and time series models represent sophisticated approaches to utilizing previous exchange rates to produce predictions.

Newaz (2008) examines exchange rate prediction through Exponential Smoothing (ETS) and ARIMA models. By applying these models to historical exchange rate data between various currencies, the study assesses their forecasting accuracy. The results show that both ETS and ARIMA models can offer reliable short-term forecasts, though their accuracy decreases over longer periods. Notably, the ARIMA model slightly outperforms ETS in terms of Mean Absolute Percentage Error (MAPE). Despite the complexity of ARIMA models, the study concludes that they offer marginally better predictive capabilities, highlighting their usefulness in financial forecasting.

Islam and Hossain (2021) present a markedly different approach within time series models, utilizing a hybrid neural network that combines Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) for exchange rate prediction. Focusing on forecasting the closing prices of four major currency pairs with high-frequency data, their model includes 20 hidden neurons in the GRU layer and 256 in the LSTM layer. This hybrid model outperforms standalone GRU and LSTM models as well as a simple moving average model, achieving lower Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), along with the highest R-squared scores. This indicates a better fit and reduced prediction risk, underscoring the benefits of combining different neural network architectures for enhanced predictive accuracy and reliability in volatile markets.

Vipul Mehra (2017) investigates the forecasting of the USD to INR exchange rate using various time series analysis techniques, including Holt-Winters Simple Exponential Smoothing, ARIMA, and Neural Networks. Analyzing ten years of daily exchange rate data, the study compares the performance of these models. The findings reveal that the HoltWinters Simple Exponential Smoothing model outperforms both ARIMA and Neural

Network models in terms of Mean Absolute Percentage Error (MAPE) on training and test datasets. Specifically, the Holt-Winters model shows superior accuracy with the lowest MAPE values, while the Neural Network model is outperformed by the other two. This study demonstrates the effectiveness of traditional time series models in exchange rate forecasting compared to more complex neural network approaches.

3. Methodology

3.1 Data Exploration

Initial data exploration was carried out to identify potential berrors in the data, missing values and checking for anomalies. This was done by printing all necessary information about the dataset and visualizing the time series. The exploration process revealed some missing values which were imputed using na.interp, an R implementation for univariate time series imputation which has been shown to provide good results (Moritz et al., 2015). Additionally, there are no exchange rate observations for Saturdays and Sundays, likely because the foreign exchange market is closed on weekends, causing no changes in exchange rates during this period. Therefore, the values for the weekends were forward imputed using the exchange rate from Friday for both days.

3.2 STL Decomposition

As part of the data exploration process, I decided to decompose the data to get a better understanding of its underlying structure. Information gathered by isolating and analysing trend, seasonal, and residual components of the data separately, will be useful to determine further processing steps and choice of models. There are multiple decomposition techniques which can be used for this, but most of them are designed to handle monthly or quarterly data (Hyndman & Athanasopoulos, 2021). Seasonal and Trend decomposition using Loess (STL) is a robust decomposition method which handles any type of seasonality, which is why I have chosen to use it for decomposing my daily data.

3.2.1 Trend

The decomposition reveals some fluctuations in the trend values although a general positive trend can be observed, particularly since around 2012. This point represents a potential structural break as the trend goes from slightly declining to a clear positive trend.

3.2.2 Seasonality

Mathematical transformations can be useful in time series modelling to normalize seasonal variation across time series levels (Hyndman & Athanasopoulos, 2021). Looking at the trend and seasonal variation, there is little evidence of heteroskedasticity but rather an increase in variability from around 2008 which can be partly explained by the financial crisis which impacted the global markets during that time (Leung et al., 2017). A mathematical transformation could still be appropriate to stabilize variance. To assess the effect of a Box-Cox transformation using Guerrero to find a suitable value for λ, I plotted a decomposition plot for the transformed data which can be seen in appendix A. Looking at the plot, the transformation did not seem to have a stabilizing effect on either the residuals or seasonal variation.

3.2.3 Remainder

The remainder shows residuals with significant variability although close to a mean of zero for the most part. Notable spikes can be observed around 2008 and from 2020. As mentioned earlier, increased variation can be related to the financial crisis which affected the volatility of the foreign exchange market and the spikes from 2020 can be explained at least partly by the spread of COVID-19 (Feng et al., 2021).

3.3 Structural breaks

As put forward in the analysis of the decomposed time series, there is some evidence of structural breaks in the time series. To confirm these suspicions and pinpoint the period where the structural break occurs, a Quandt Likelihood Ratio (QLR) test and an Information Matrix (IS) test was conducted. The QLR revealed a significant structural break around 2011, indicated by a peak in the F-statistics. This is also supported by the IS test which produces a p-value of 0.005, which suggests there is strong evidence against the null hypothesis of no structural breaks. A possible explanation for this point could be a combination of reduced turmoil after the financial crisis and close to the point where the value of the Euro starts to steadily increase compared to the Norwegian krone.

There are multiple ways to deal with structural breaks, but as it appears relatively early in the time series, leaving a significant number of observations after it, I decided to drop the data before this structural break. By dropping data from before the structural break, the overall time series becomes more stable, contributing to a simplified modelling process. This approach will also lead to leaving out trends that are not likely to be relevant for current and future predictions. The QLR test identified the 176th day of 2011 as the breaking point, but looking at the time series plot, it seems that there might be some transition effects. For this reason, I decided to drop all data before 2012.

Carrying out the QLR test on the remaining data reveals significant spikes in F statistics starting in 2020 which as explained earlier can is likely to be related to COVID-19. These findings support the sentiment by Hyndman (2014) that structural break tests are often mis- and overused, stating that "... most things change slowly over time, and only occasionally with sudden discontinuous change... But that is not very often, and we know about it when it happens". The most significant spike in F statistics is in late 2021, and it is reasonable to assume that COVID-19 would cause a structural break in currency exchange rates. However, high F statistics after this time is likely due to increased volatility which can be explained as long term economic effects of this event. Considering the statement by Hyndman, I decided to remove data up until the structural break indicated by the second QLR test, but not removing any additional data. After doing this, the remaining time series contained 895 data points providing the most accurate representation of the current currency exchange rate market.

4. Model Selection

4.1 ARIMA

ARIMA models perform forecasting based on the autocorrelation of time series data (Hyndman & Athanasopoulos, 2023). This model represents one of the two most commonly used approaches to time series forecasting and it is highly represented in the exchange rate forecasting literature. ARIMA models consist of autoregression (AR) which makes a prediction based on the linear combination of lagged values, differencing (I) to make the data stationary and moving average (MA) which performs regression on weighted previous forecast errors. An ARIMA model combines both the AR and MA predictions on the stationary data to produce a final prediction.

ARIMA is chosen because it can address the identified trends and seasonal patterns of the time series data through differencing. Furthermore, it captures the strong autocorrelation of exchange rate data withs its AR and MA components, providing flexibility to model the complex temporal patterns observed.

4.2 TBATS

Exponential smoothing is the second of the two most common approaches to time-series forecasting (Hyndman & Athanasopoulos, 2021). ETS models are common implementations of the exponential smoothing approach which provide flexibility to capture different characteristics such as level changes, trends and seasonal components. Trigonometric Box-Cox transformed ARMA Trend and Seasonal Components (TBATS) model extends simple exponential smoothing models by adding a trigonometric function to account for complex seasonality.

While ETS methods can provide reliable predictions efficiently, they struggle to deal with the complex seasonality of daily data (Hyndman & Athanasopoulos, 2018). The TBATS model proposed by De Livera et al. (2011) accounts for multiple seasonal periods, highfrequency seasonality, non-integer seasonality and dual-calendar effects. It should be noted that a TBATS model is less flexible than the ETS model and can be slow to estimate. Another alternative would have been to perform STL decomposition and model forecast based on the data, neglecting the seasonal component. TBATS has been chosen for this project as the seasonal component has potential to contribute valuable input for more accurate predictions.

4.3 NNAR

Neural networks consist of processing units and associated weights which in combination allows the network to model complex nonlinear relationships in data (Hyndman & Athanasopoulos, 2021). Neural Network Autoregression (NNAR) models are similar to ARIMA models, but the architecture as well as the input to be considered can be specified leading to a nonlinear model capable of identifying complex relationships which may not be captured by the ARIMA model. The NNAR model also does not require stationary data. As the data fed to the NNAR model does not need to be differenced, the model may consider valuable information in the trend and seasonality of the data.

5. Model Evaluation

To evaluate and compare the three models chosen, various methods will be employed to determine the optimal hyperparameters for each of them. When evaluating models, we are not concerned with the scale of the data as each model using the same algorithm also uses the same scaled data. However, when comparing algorithms for our use case, the ARIMA model will produce predictions on differenced data, resulting in errors on a different scale than TBATS and NNAR, which forecast on the raw data. To address this, we will use scaled error measures as proposed by Hyndman & Koehler (2006). Specifically, we will use the Mean Absolute Scaled Error (MASE) and Root Mean Squared Scaled Error (RMSSE) to evaluate and compare the models. These measures provide a normalized approach to error evaluation that accounts for the scale of the data, allowing for fair comparison between models applied to different data transformations.

5.1 ARIMA

ARIMA models require time series data to be stationary. To evaluate the stationarity of the data and potential need for differencing, unit root tests have been carried out. Various unit root tests are based on different assumptions which can lead to conflicting answers (Hyndman & Athanasopoulos, 2021). By combining the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the confidence in the result increases. The stationarity of the data was initially tested using a simple implementation of both ADF and KPSS using default parameters, and both tests suggested non-stationarity in the data. The detailed tests after first-order differencing confirm that the data is stationary. All ADF tests reject the null hypothesis of a unit root, and KPSS tests fail to reject stationarity, confirming that the data is stationary in all respects. Detailed metrics are in Appendix B.

To estimate the optimal parameters for the ARIMA model it can be useful to look at the ACF and PACF of the differenced time series. The ACF plot reveals significant spikes at the first seven lags with rapid decline to values near zero. This indicates that the data is now stationary and that there is an AR component. The PACF plot shows a significant spike at the first lag, with additional significant spikes at every 7 lags, indicating a weekly seasonal MA component.

Given these observations, the data has been successfully transformed to a stationary state through first-order differencing and there is no indication of integration components. The non-seasonal AR component should be one but the I component should be zero. The significant spikes at the first seven lags in the ACF plot is indicative of an MA component of 7. The seasonal AR parameter should be zero but the seasonal MA component should be one, with a seasonal period of 7. Based on these plots, the suggested model is a seasonal ARIMA(1,0,7)(0,0,1)[7].

In addition to estimating the parameters based on the ACF and PACF plots, I used the auto.arima function to automatically identify and fit the best seasonal ARIMA model to the time series data. When fitting the ARIMA model using with exhaustive search without approximation the algorithm identified an ARIMA (0,1,0) model as the optimal fit with a

Corrected Akaike's Information Criterion (AICc) score of -29757.5. With default parameters the optimal identified model was an ARIMA (1,1,2) with an AIC of 29755.01. The model identified through the exhaustive search is very simple and is essentially a random walk with first degree differencing to make the data stationary.

The residuals of each of the models are centered around the mean without obvious patterns. Both the automatically selected models exhibit no significant autocorrelation based on their ACF plots. The seasonal ARIMA model which was selected based on the initial ACF and PACF plots exhibit significant spikes at lag 1, 4, 7 and 14, indicating that the model does not handle autocorrelation as well as the other models.

Looking at the Ljung-Box test results for all of the ARIMA models we have tested show high p-values, suggesting no evidence to reject the null hypothesis of remaining autocorrelation. Examining the residuals visually supports this finding as there is no clear skewness and residuals are centred around the mean, displaying white noise traits.

To decide on the ARIMA model which will be used for forecasting and comparison with the other algorithms in this project, I examine the information criteria for each of these. Among the three models, ARIMA(0,1,0) has the lowest Akaike's Information Criterion (AIC), AICc and Schwarz's Bayesian Information Criterion (BIC) values, making it the most parsimonious and effective model for the time series. This suggest that the additional parameters in the more complex models do not sufficiently improve the model fit to justify their complexity.

5.2 TBATS

TBATS is a completely automated model, which does not require nor allow for hyperparameters. This model handles complex seasonality but as it is fully automated, it is not very flexible which can lead to poor forecasting.

Examining the residuals, their distribution and the autocorrelation of the TBATS model suggest remaining autocorrelation in the first lag, but other than this equally distributed residuals and little to no skewness. The Ljung-Box test also returns a p-value of 0.7275 which is much higher than the confidence level of 0.05, suggesting no remaining autocorrelation.

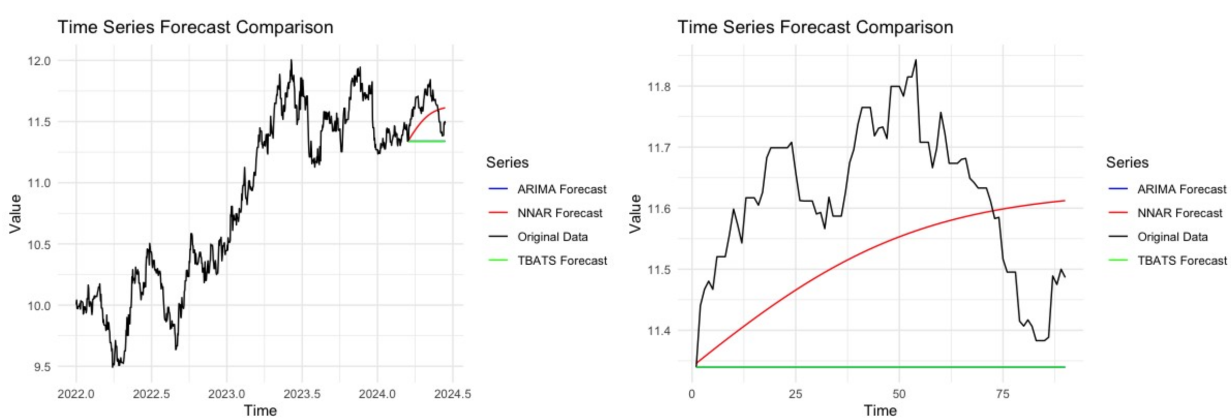
5.3 NNAR

Like the ARIMA model, the NNAR model is too customizable so parameters can be chosen based on established knowledge of the time series. The NNAR (p, P, K) m model takes four hyperparameters which select default values if not specified. The default values of the model for the training set were NNAR (1, 1, 1)1 which means the default model uses one lagged value, one seasonal lagged value and has a single hidden node. An m value of 1 also indicates that the model uses a frequency of 1 rather than 7, 30 or 365 which can be useful to capture any complex seasonal patterns. This indicates that there is no significant seasonal component which will improve forecasting. This means that the seasonal lagged value for this model will be the same as the lagged value. The default model gets a relatively low sigma2 score of 0.001072 indicating that it fits the training data well. Still, there is a lot of information not being utilized when only considering one lagged value and no seasonal lagged values, and the NNAR's ability to model complex relationships is being limited by only utilizing one hidden node. For this reason, I decided to also consider a model that accounts for a higher degree autoregressive term with a p value of 4 and a continue with a seasonal term of 1 but changing the frequency to 7 to include the value of that same day the week before. For the number of neurons, we use the default algorithm for selection which is normally used when k is not specified: (p + P + 1) / 2, resulting in a value of k equal to 3. The resulting custom model is an NNAR (4, 1, 3)7.

The residuals of the NNAR (1,1) model are slightly skewed with some spikes outside the significance interval of the ACF plot. The p-value of the Ljung-Box test is 0.3404 suggesting that most autocorrelation is accounted for and that the residuals are mostly due to white noise. The residuals of the NNAR (4,1,3) model on the other hand, show less significant spikes in the ACF plot and residuals slightly more centred around the mean. The p-value is also very high for this model with a value of 0.9125, providing no evidence of remaining autocorrelation.

When comparing the two NNAR models based on the training set results, the customized NNAR(2,1,3) model outperforms the default NNAR(1,1,1) model in several key metrics. The customized model has a lower RMSE (0.03249 vs. 0.03275) and MAE (0.01953 vs. 0.01996), indicating it provides a better fit to the training data. Additionally, the customized model also shows a slightly lower MAPE (0.2323% vs. 0.2375%), suggesting it is more accurate in percentage terms. The near-zero ME and lower ACF1 value further confirm that the customized NNAR model captures the patterns in the training data more effectively, making it the preferred model based on these metrics. It should be noted that the model is more complex and as such it may not generalize as well to unseen data. This effect has been mitigated by choosing a relatively small size for the network as well as a high number of repeats to average the performance over.

6. Forecasts



Picture 2: Forecasted values plotted over the original time series. The ARIMA forecast is hard to see as it has close to an identical forecast to the TBATS model.

The forecasts of the ARIMA(1,1,0) and TBATS models primarily reflect a naive prediction strategy rather than capturing the nuanced trends and patterns present in the data. The STL decomposition plot highlights the complexity of the time series, characterized by some trend and seasonal components although irregular, and a significant portion classified as remainder. This complexity likely contributes to the challenges faced by the models and is to be expected for volatile exchange rate data. These models' inability to capture the slight trend in the data is likely due to the absence of a clear linearity in this trend.

Model	MAE	RMSE	MAPE
ARIMA(1,1,0)	0.2745	0.2988	2.3538
TBATS	0.2743	0.2986	2.3521
NNAR(4,1,3)	0.1531	0.1678	1.3167

Table 1: Accuracy measures of each of the three selected models.

The NNAR (4,1,3) model significantly outperforms the ARIMA (1,1,0) and TBATS model across all accuracy measures considered (MAE, RMSE and MAPE). The neural network model's performance is also evident when comparing the forecasts in the plot above. We can see from the plot that the NNAR model can capture a short-term trend, but it is no better than the other models in capturing seasonal variations. Still, given the high volatility of the data, the performance of the neural network model is promising.

7. Conclusion

This study aimed to assess the effectiveness of statistical models compared to machine learning models in forecasting exchange rate data. The first step of this project was to determine whether any form of mathematical transformation would be beneficial, then testing for structural breaks and finally, stationarity. I found that a mathematical transformation of the data had little to no stabilizing effect on the variance, that first order differencing was sufficient to achieve stationarity and that structural breaks were present in 2011, 2021 and throughout the time series after removing data up until 2022. I compared models suggested by automated algorithms with customized models based on characteristics of the time series data, and model selection was based on information criteria, residual characteristics and accuracy on the test set.

The NNAR was found to be the superior model based on all accuracy measures considered due to its ability to capture the trend in the time series. None of the models were able to capture seasonal variance resulting in a smooth forecast with trend for the NNAR and close to naïve forecasts from both the ARIMA and TBATS models. This suggests that the data is too volatile, displaying little to no significant patterns except a non-linear trend which is only captured by the non-linear model. It is possible that more complex models would be able to capture even more complex relationships in the data. Furthermore, my decision to train the models and forecast on data post COVID-19 clearly makes the task more difficult, and forecasts before 2020 would likely be more accurate.

The neural network model is much more complex than the ARIMA and TBATS models chosen, and it is able to capture non-linear relationships in the data which explains its effectiveness in forecasting exchange rate data over short time-intervals. Neither the TBATS nor the ARIMA models were effective in this case but it is possible that forecasting on more stable data could favour linear models, although it is likely that the ability to customize the models as with ARIMA is a notable advantage compared to TBATS.

References

De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. Journal of the American statistical association, 106(496), 1513-1527.

European Central Bank. (2017). Exchange rate prediction redux: New models, new data, new currencies. No 2018 / February 2017. Publications Office. https://data.europa.eu/doi/10.2866/524144

Feng, G., Yang, H., Gong, Q., & Chang, C. (2021). What is the exchange rate volatility response to COVID-19 and government interventions?. Economic Analysis and Policy, 69, 705-719. https://doi.org/10.1016/J.EAP.2021.01.018. Galeshchuk, S. (2016). Neural networks performance in exchange rate prediction. Neurocomputing, 172, 446-452.

Islam, M. S., & Hossain, E. (2021). Foreign exchange currency rate prediction using a GRU LSTM hybrid network. Soft Computing Letters, 3, 100009.

Leung, H., Schiereck, D., & Schroeder, F. (2017). Volatility spillovers and determinants of contagion: Exchange rate and equity markets during crises ☆. Economic Modelling, 61, 169-180. https://doi.org/10.1016/J.ECONMOD.2016.12.011.

Lowry, C. (2024, April 18). How often are exchange rates updated. Blog | Western Union. https://www.westernunion.com/blog/en/us/how-often-are-exchange-rates-updated/

Mehra, V. (2017). Forecasting USD to INR foreign exchange rate using Time Series Analysis techniques like HoltWinters Simple Exponential Smoothing, ARIMA and Neural Networks. Academia. edu, 4-5.

Moritz, S., Sardá, A., Bartz-Beielstein, T., Zaefferer, M., & Stork, J. (2015). Comparison of different methods for univariate time series imputation in R. arXiv preprint arXiv:1510.03924.

Newaz, M. K. (2008). Comparing the performance of time series models for forecasting exchange rate. http://dspace.bracu.ac.bd:8080/xmlui/handle/10361/438 Rob

J Hyndman. (2014, May 23). Structural breaks. Rob J Hyndman. https://robjhyndman.com/hyndsight/structural-breaks/

Rob J Hyndman & George Athanasopoulos. (2018). Forecasting: Principles and Practice (2nd ed). OTexts. Retrieved July 12, 2024, from https://otexts.com/fpp2/

Rob J Hyndman & George Athanasopoulos. (2021). Forecasting: Principles and Practice (3rd ed). OTexts. Retrieved June 7, 2024, from https://otexts.com/fpp3/

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001