Forecasting Gold Prices

Report Objective

The objective of this forecasting project is to develop a robust, interpretable, and theoretically grounded model for forecasting gold prices and to forecast gold prices for the next 12 months, from March 2025 - February 2026. This model will support informed decision-making for investors, jewellers, traders, and policymakers by integrating economic theory, rigorous statistical modelling, and real-world insights. A forecast horizon of 12 months is suitable for these various stakeholders as it offers a reasonable balance between shorter-term responsiveness and long-term planning. Investors and businessmen often utilize shorter-term horizons, such as 12 months, for businesses seeking to navigate immediate operational decisions, who may not have the privilege to invest long-term (Brixx, 2023). While some stakeholders may value longer-term forecasts, 12 month forecasts are well-suited for most stakeholders and businesses seeking short-term forecasts and, thus, will be implemented in this report.

For investors, the model will help assess market timing and portfolio diversification strategies by identifying potential trends and volatility in gold prices. Jewellers can use the forecasts to better manage inventory costs and pricing strategies, particularly in anticipation of peak demand periods or expected price shifts. Traders will benefit from short-term predictive signals to optimize entry and exit points in gold markets, while policymakers can leverage the forecasts to monitor inflation expectations, currency stability, and economic sentiment, all of which are factors in which gold prices often serve as a key barometer.

Data Collection & Preprocessing

The first step is to collect monthly data about the prices of gold from a reputable and unbiased source. The *World Gold Council,* a nonprofit company, is one of the most popular and attractive sources of information regarding the gold commodity. They are a globally-recognized organization consisting of the world's largest gold mining companies and industry experts (BullionByPost, 2025). The *World Gold Council* collects data that contains detailed information about average gold and gold reference prices in many different units of measurement, frequencies, and lengths of time.

The monthly price of gold was collected from January 1978 to February 2025 from the *World Gold Council* website. The downloaded dataset was an Excel workbook in .xlsx format and contained several sheets, each with gold prices recorded in different frequencies at different currencies. The .xlsx file containing the data was loaded into a new R Studio project and script. I first applied the here() function to my dataset before passing it to read_excel(). The here() function creates a file path based on the main folder of your R project, which helps keep your code organized and easy to run on any computer. When reading the data using read_excel(), I set “sheet = 4” to specify that I want to select the 4th sheet in the excel workbook. The 4th sheet contains monthly data for gold prices. The 4th sheet was then viewed using the View() function. There were multiple columns, each for the price of gold using different currencies. For my analysis, I will only use the price of gold in US dollars. Because of this, I will only select the 2nd and 3rd columns from this excel sheet which contains the date and the price of gold in US dollars respectively. I then renamed my two columns with their appropriate title as the original column names were nonsensical.

I visually inspected the dataset first using the ‘View()’ function. Gold prices >= 1,000 contained commas, which would interfere with the analysis. I then checked the structure of the dataset using the ‘str()’ function. The variable types for my two columns, date and price, were incorrect (both were ‘character’ data types). The dataset was also inspected for missing values using the ‘colSums(is.na())’ function. The dataset was free of missing values, so no imputation had to be done. I then converted the ‘date’ column to a ‘Date’ data type and the average price column to a ‘Numeric’ data type. I also removed the commas from the average price column to ensure all observations are purely numeric. Finally, I converted the average price column to a time series object using the ‘ts()’ function. I set the start of the time series as the minimum (earliest) date and specified the frequency as 12 for monthly data. Using the ‘autoplot()’ function on my cleaned dataset, I can visualize the overall pattern of the entire time series. This is shown in *Figure 1.* I will be using data from 2015 onwards. This is due to several reasons. Firstly, the time plot exhibits significantly different behavior pre-2000 and post-2000. In addition, there is a very non-linear pattern between 2010 and 2015.

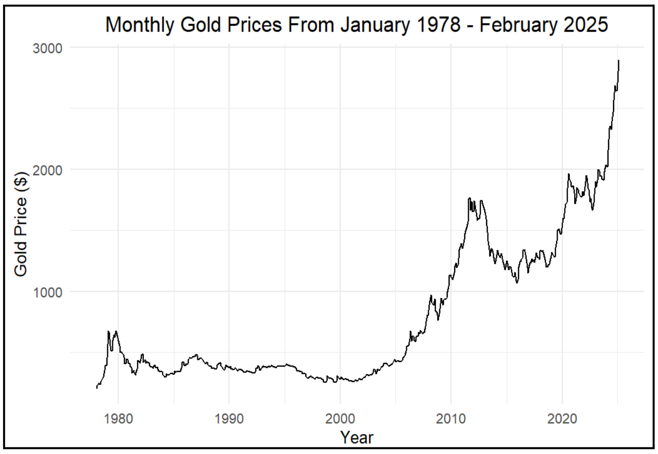
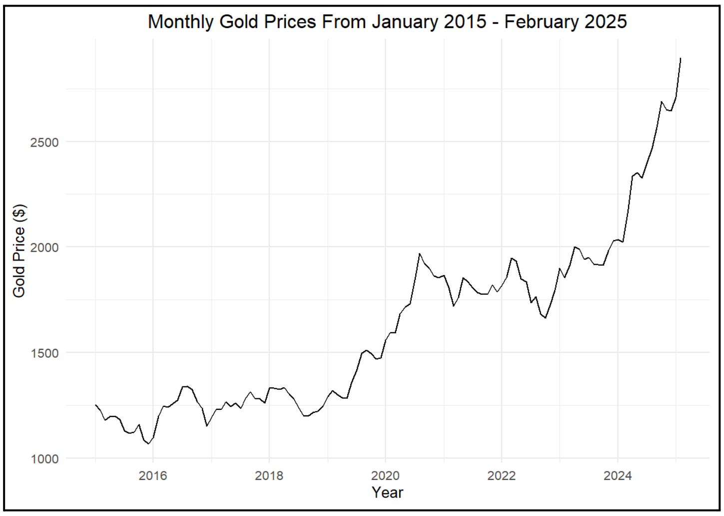


Figure 1

In addition, the older data behave significantly differently than the newer data. This is likely due to the different economic landscape around that time. Prior to 2015, gold prices were significantly influenced by extraordinary economic events and policy decisions that introduced volatility and anomalies, potentially skewing long-term trend analyses. The global financial crisis that lasted from 2007 - 2009 led investors to seek out safer assets that are immune to economic and financial shocks. In the United States, drastic measures taken by the government, such as quantitative easing introduced liquidity into the market and increased the money supply, thus lowering the value of the dollar. As a result, investors sought after financial instruments that maintained the value of their assets. Commodities like gold provided a safe haven for investors during this period of economic turmoil due to their stable intrinsic value (Hergt, 2013). This led to a sharp increase in the price of gold which peaked in 2011. Such economic crises are quite rare, and basing future forecasts on past realizations of the price of gold influenced by such rare crises might lead to biased forecasts. Moreover, to achieve accurate forecasts using simple and exponential smoothing models, non-linearity must be eliminated as much as possible from the time plot. This is because these models assume a linear relationship in the time series and forecasts are also linear (Katy, 2025). Significant nonlinearity can be seen from 2008 - 2015, where the graph exhibits a significant upward trend until around 2011, followed by an abrupt drop.

Eliminating such nonlinearities would significantly improve the accuracy of the forecasts. After discarding data before 2015, I am left with 122 observations, which is generally considered sufficient for time series forecasting. However, there is no hard rule. This final data is then plotted and presented in *Figure 2*.

Exploratory Data Analysis



The time series exhibits an overall linear upwards trend, with a stagnation period around the end of 2020 until 2023, likely due to the COVID-19 pandemic. Around the end of 2020, a COVID-19 vaccine was discovered. This reduced the risk of the COVID-19 threat with economists and investors thinking the world economy will recover. With an optimistic economic outlook, investors saw less need to hold their assets in safer financial instruments like gold and, thus, its

Figure 2

demand dropped, which resulted in the price of gold stagnating towards the end of 2020 after sharply increasing for several months (The Malaysian Reserve, 2020). There doesn't seem to be any seasonality in the time plot, as there are no consistent and predictable fluctuations. Instead, there are irregular fluctuations of inconsistent frequency and magnitude. The ACF (autocorrelation function) plot and subseries plot are examined in *Figure 3* below to confirm these patterns.

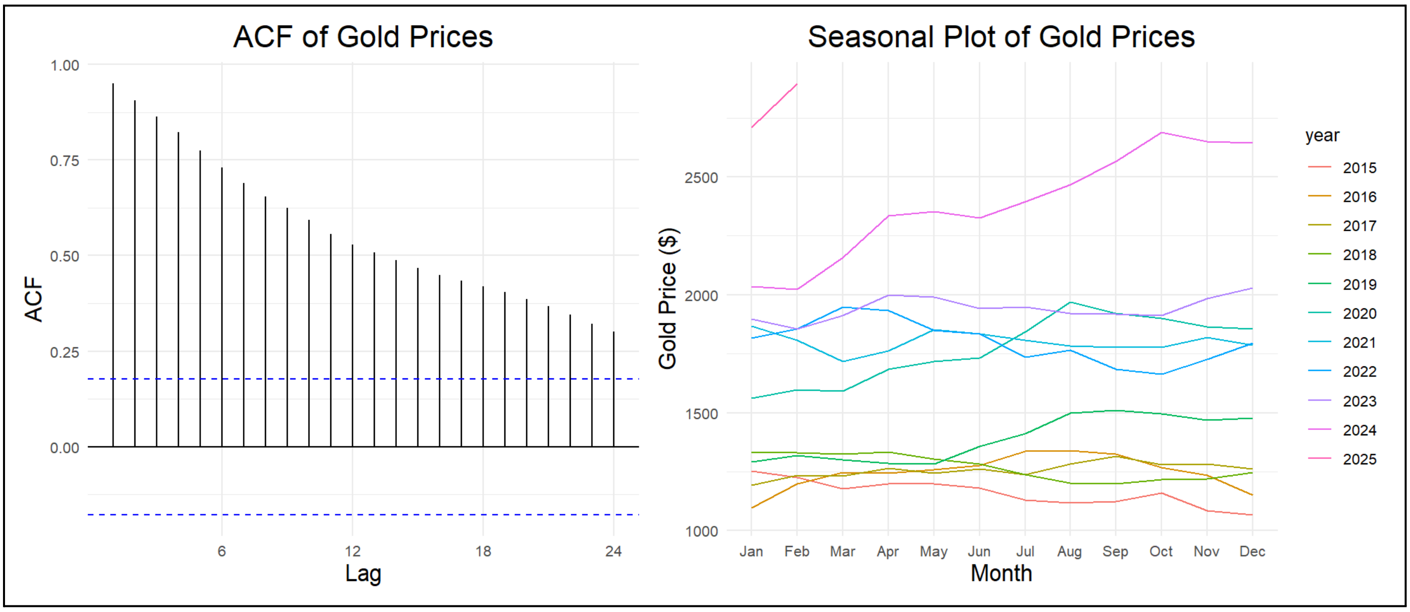


Figure 3

The ACF plot shows slowly-decaying autocorrelation functions at different lags. This is indicative of a trend. In addition, there are no clear patterns or fluctuations at specific lags of the autocorrelations as they are decreasing through the lags, indicative of no seasonality. The season plot shows no consistent pattern at any specific month for all the years, on average. This is also indicative of no seasonality. While there's some natural variability in the time plot, it is not strong enough to clearly suggest heteroscedasticity (non-constant variance). As a result, applying a Box-Cox transformation to this time series would not improve model accuracy, as there is no changing variance throughout the series that needs to be stabilized. It may even worsen the accuracy of the forecast by unnecessarily introducing complexities (Sakia, 1995).

Given the presence of an upward trend and no seasonality, I will choose a random walk with drift as my simple forecasting and benchmark model and ETS (A, A, N) model as my exponential smoothing model. Among the simple forecasting methods, the random walk with drift is the most theoretically sound simple forecasting model for a time series with a trend, given the drift term that accounts for the average increase in the level of the series. Given the established pattern of an upward trend and no seasonality, a random walk with drift is chosen as my simple forecasting model. The relatively stable variance suggests that an additive error would be more suitable than a multiplicative error for my ETS model. In general, a multiplicative trend is not suitable due to its poor forecasting performance (Hyndman & Athanasopoulos, 2018). A damped trend would also not be suitable as the stagnant gold prices around mid 2020-2023 would decrease the overall forecasted values and potentially underestimate the price of gold as ETS takes an exponentially-weighted average of all past values, including the values from 2021-2023.

This period, like the period around the 2008 global financial crisis, was also influenced by a rare event (COVID-19 Pandemic). Including a damped trend would therefore decrease the forecasted values even further, potentially leading to even worse forecasts with bigger errors. Therefore, it is logical to choose an additive trend over an additive damped trend. No seasonality was chosen for the ETS model based on the previous assessment of no seasonality in the original time plot.

Forecasting Model Development

I will implement 3 models in my analysis. The first model will use a simple forecasting method that will act as a benchmark for the performance of the remaining 2 models. The second model will be an ETS model with manually-set parameters while the third model will be an ETS model with parameters automatically set by R.

Before implementing any model, I need to split my data into a training set and a test set. The models will be fitted using the training data and their forecast performance will be assessed by forecasting the models on the test data. This way, a model's forecast accuracy can be assessed without waiting for new data to emerge. The training set consisted of the first 80% of observations while the test set consisted of the final 20%. This is visualized in *Figure 4*.

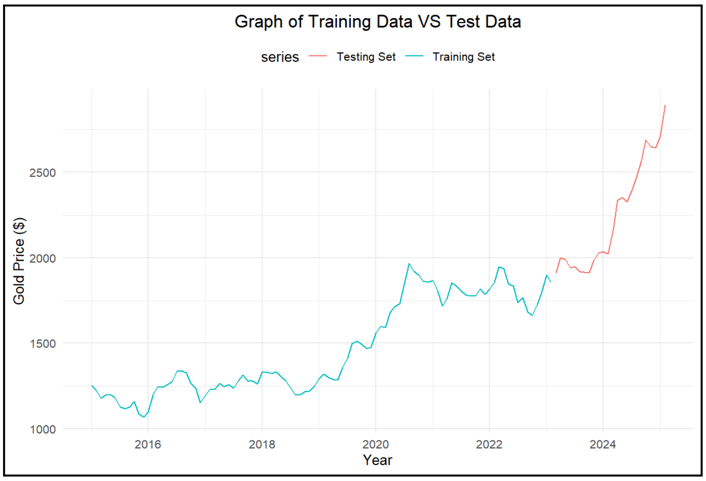


Figure 4

This split ensures that a sufficient number of data points are included in the training set. However, this split is not perfect as the late 2023 structural break has not been fully captured by the training data.

The h-step ahead forecast using a random walk with drift model can be written as follows: $\hat{y}_{T+h|T} = y_T + \frac{h}{T-1} \times \sum_{t=2}^{T} (y_t - y_{t-1})$ where h is the forecasted period. It can be seen that the h-step ahead forecast is the sum of the previous value of the series and the average change. This is equivalent to drawing a line from y_1 through y_T. The drift, or average change, of this model is 6.21. This means that the model will forecast an average increase of $6.21/month for the price of gold. Substituting the drift into our formula, we get the following equation: $\hat{y}_{T+h|T} = y_T + h\,(10.83)$. The h-step ahead forecast is the sum of the series/forecast of the previous month plus h multiplied by the average change of $6.21.

Based on the original plot of the time series, I concluded that the time plot possessed an overall upward trend with no seasonality. In addition, I noted that the variance, or random fluctuations in the series, is relatively constant over time. As a result, I implemented an ETS(A, A, N) model.

The choice of additive errors (first A) is inferred based on the constant variance in the time series. The additive trend (2nd A) is due to the overall upwards trend (note: multiplicative trends result in unreliable and unstable models). The final parameter, N, represents no seasonal component as discussed. This model uses Holt's linear exponential smoothing method and provides a prediction interval for its forecast.

The state-space equations for an ETS(A, A, N) model can be represented as follows:

$y_t = l_{t-1} + b_{t-1} + \epsilon_t$ (observation equation)

$l_t = l_{t-1} + b_{t-1} + \alpha\epsilon_t$ (level/mean-updating equation)

$b_t = b_{t-1} + \beta\epsilon_t$ (trend-updating equation)

The additive error can be seen as a result of adding ϵ_t or its weighted average in all equations. The additive trend can be seen when adding b_{t-1} in both the observation and level-updating equations. The lack of seasonality results in no seasonal γ parameter and no initial seasonal s_0 state.

In order to implement these equations, initial states for the level and trend must be realized by recursive substitution until the initial level and trend states l_0 and b_0 respectively as well as the smoothing parameters α and β. α represents the level smoothing parameter, controlling how much weight the model gives to new information when updating the level.
β represents the trend smoothing parameter, determining how responsive the trend component is to recent changes.

These are presented in *Table 1* below. The parameters can be interpreted as follows:

ETS(A, A, N) Model Parameters	
Parameter	Value
Alpha	0.99
Beta	0.01
Initial Level (l)	1,245.79
Initial Trend (b)	−0.32
Sigma	48.25

Table 1

Alpha (α): The alpha value of 0.99 means that almost all the weight is placed on the most recent observation when updating the level/mean of the series. The level estimate reacts very quickly to changes in the observed data. Past observations have negligible effect on level estimates.

Beta (β): A beta value of 0.01 suggests that the model places very little weight on recent changes in the trend, instead favoring the historical trend component.

Initial Level (l_0): The model's best estimate of the baseline value of the price of gold is $1,245.79

Initial Trend (b_0): According to the model's best estimate, the time series initially started with an average decrease of $0.32/month for the price of gold

Sigma (σ): On average, the forecasts differ from the actual values by about $48.25. This defines the forecast's prediction intervals.

These parameters assume linearity. For example, the trend assumes the same increase every month. The final ETS model is one where the parameters are chosen automatically by R. R decided that an ETS(M, N, N) model would be optimal to forecast this time series.
This chosen state-space model has multiplicative errors, no trend, and no seasonality. The first 2 arguments are different compared to the previous model with manually-chosen parameters. The effect of having no trend is a perfectly horizontal forecast and the multiplicative error results in a wider prediction interval. These forecasts are clearly significantly less accurate as compared to both previous models as we have established that the time series has an upward trend. The state-space equations for the estimated ETS(M, N, N) state-space model can be formulated as follows:

$y_t = l_{t-1}(1 + \epsilon_t)$ (observation equation)

$l_t = l_{t-1}(1 + \alpha\epsilon_t)$ (level/mean-updating equation)

ETS (M, N, N) Parameters (Auto Chosen)	
Parameter	Value
Alpha	0.99
Initial Level (l)	1,250.54
Sigma	0.03

Table 2

In these models, the error is a product rather than an addition to an equation. This is due to the multiplicative nature of the error in this model whereas the error was additive in the previous ETS model. The initial level is once again calculated. The estimated parameters of this model are shown in *Table 2* below.

The interpretation is as follows:

Alpha: The alpha smoothing parameter is the same as the chosen ETS model. almost all the weight is placed on the most recent observation when updating the level/mean of the series.

Initial Level (l_0): The model's best estimate of the baseline value of the price of gold is $1,096.33. As the model has estimated no trend, there is no trend-updating equation and no initial trend level.

Model Selection & Evaluation

Now that I have fit all my models using the training data, I can measure their accuracy by comparing their forecasts with the testing data. Firstly, we must examine the residuals of each model, which is the difference between the observed values and the model's forecasted values. This is crucial for several reasons. Having correlated residuals means there is important information left in the residuals that has not been captured by the forecasting model. This is indicative that the forecasting model used is not very robust and could be improved. The residuals should also have an average value of 0. If the mean is different from zero then the forecasts are biased; $E(\hat{Y}_{T+h|T}) \neq Y_{T+h}$. Residuals that are uncorrelated with a mean of 0 and constant variance are white noise processes. For prediction intervals of the forecasted values, there are another 2 useful residual properties. The first is that they are homoscedastic; they have constant variance. The second is that they are normally distributed. Homoscedasticity results in prediction intervals that have equal width regardless of the time horizon of the forecasts. The normality assumption allows for precise prediction intervals because it assumes the forecast errors follow a normal distribution. This means we can use the standard normal distribution to calculate the critical values (e.g., 1.96 for 95% confidence level) and construct prediction intervals that reflect the uncertainty of the forecasts. To test for autocorrelations in the residuals, we can employ the Ljung-Box test. The null and alternative hypotheses for this test are as follows:

H_0: $\rho_j = 0 \ \ \forall j$ $\qquad$ H_1: $At\ least\ one\ \rho_j \neq 0$

Where ρ_j represents the autocorrelation at lag j.

In addition, we can visually examine the residuals as well as the autocorrelation plot and distribution of residuals. The results of the Ljung-Box test as well as relevant visuals for the random walk with drift model are seen in *Table 3* and *Figure 5* respectively.

Ljung-Box Test Results for Random Walk with Drift Residuals	
Statistic	Value
Q*	22.52
Degrees of Freedom	20
P-value	0.31
Total Lags Used	20

Table 3

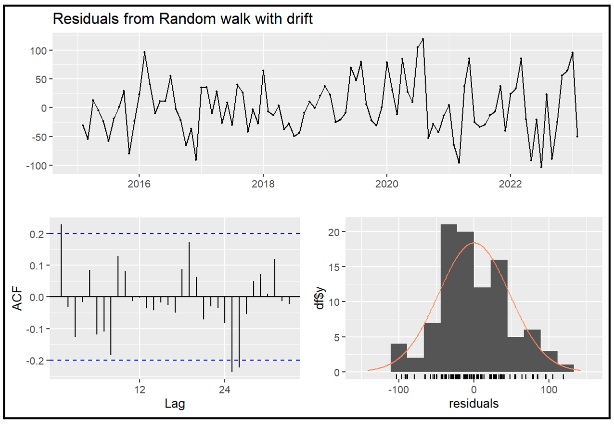


Figure 5

The p-value of the Ljung-Box test is greater than 0.05. As a result, we do not reject the null hypothesis and conclude that there is no autocorrelation in the residuals for up to 20 lags. Examining the plots in *Figure 5*, The ACF plot affirms no correlation in the lags. The residuals appear to have a constant variance as shown by the main (top) plot. In addition, they appear to have a mean-reverting behavior around a value of slightly less than 0 which is also reflected in the histogram of the residuals. The residuals also appear to be normally distributed. These properties ensure that our forecasts are relatively unbiased and the prediction intervals are consistently wide and accurate.

The results of the Ljung-Box test as well as relevant visuals for the ETS(A, A, N) mode with manually-chosen parameters are seen in *Table 4* and *Figure 6* respectively.

Ljung-Box Test Results for ETS(A,A,N) model	
Statistic	Value
Q*	22.60
Degrees of Freedom	20
p-value	0.31
Total Lags Used	20

Table 4

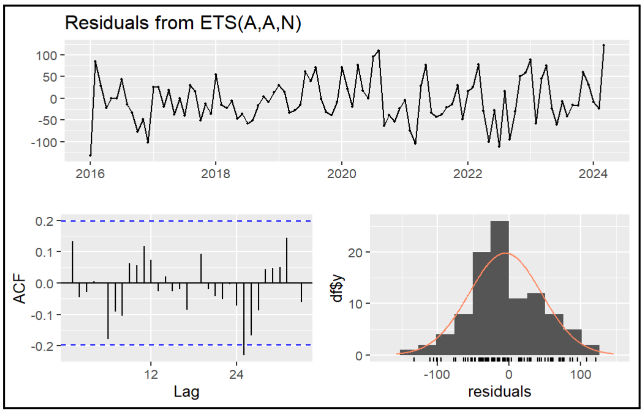


Figure 6

The p-value of the Ljung-Box test is also greater than 0.05 for the manually-chosen ETS(A,A,N) model. As a result, we do not reject the null hypothesis and conclude that there is no autocorrelation in the residuals up to the 20th lag. Also, most lags in the ACF plot are insignificant, which further supports this.The plots in *Figure 6* behave similarly to the previous model indicating that the forecasts produced by this ETS model are relatively unbiased and the prediction intervals are consistently-wide and accurate. The results of the Ljung-Box test as well as relevant visuals for the random walk with drift model are seen in *Table 5* and *Figure 7* respectively.

Ljung-Box Test Results for Auto ETS(M,N,N) model	
Statistic	Value
Q*	19.15
Degrees of Freedom	20
p-value	0.51
Total Lags Used	20

Table 5

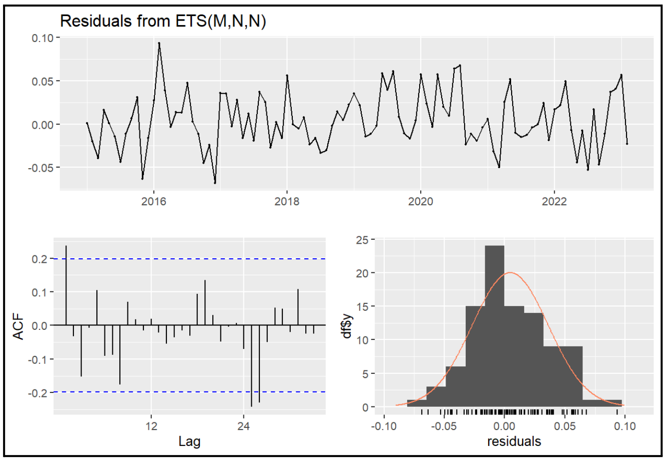


Figure 7

The p-value for the automatic ETS() model is also greater than 0.05. This means that we, once again, fail to reject the null hypothesis and conclude that there is no autocorrelation in the residuals up to the 20th lag. The ACF plot of the residuals in *Figure 7* shows most autocorrelations up to 30+ lags to be insignificant. In addition, the residuals have a relatively constant mean and insignificant lags. The residuals appear relatively normally distributed, and the mean appears to be close to 0. As a result, the forecasts are relatively unbiased and prediction intervals are relatively accurate. The residual diagnostics indicate uncorrelated residuals for all 3 models. The auto ETS(M, N, N) model has a slightly less desirable distribution of its residuals, potentially indicating that it may not be the best fit. Empirically, we have already examined an overall upward trend in the time plot, so it is highly unlikely for this model to have the best fit. However, each model's accuracy must be assessed by comparing forecasted values with actual data observed in the test set. The 3 performance metrics that will be used are the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The formulas for these 3 metrics are given below:

$$RMSE: \sqrt{\frac{1}{H}\sum_{h=1}^{H} e^{2}_{T+h}} \quad MAE: \frac{1}{H}\sum_{h=1}^{H}\left|e_{T+H}\right| \quad MAPE: \frac{1}{H}\sum_{h=1}^{H}\left|\frac{e_{T+h}}{y_{T+h}}\right| \times 100$$

RMSE and MAE are scaled according to the scale of the original data while MAPE is measured in percentage. Given that our data contains no 0 values, MAPE is able to be computed for every value. The RMSE, MAE, and MAPE of each of the 3 models against the test set is shown in *Table 6 below*. The ETS model with manually-chosen parameters ETS(A,A,N) has the best forecasting accuracy according to the testing set compared to all 3 measures as they are all the lowest compared to the benchmark random walk with drift model and the ETS model with parameters automatically chosen by R.

Forecast Accuracy Comparison (Test Set)

Model	RMSE	MAE	MAPE (%)
Random Walk w/ Drift	83.15	73.16	3.68
ETS(A,A,N)	97.45	88.56	4.46
ETS(M,N,N) (Auto)	122.36	113.54	5.72

Table 6

However, these forecast accuracies are not very robust because there is only 1 training and test set. There is a more robust, modern method to assess forecast accuracy, time series cross-validation. Instead of having only 1 training set and 1 test set, cross-validation splits the data into training and test sets multiple times. Each time, there is only 1 data point for the test set. For each training-test split, the training set “rolls” forward in time and captures newer and newer data points. The overall forecast accuracy is then computed by averaging the forecast errors from each training-test split. Time series cross-validation can be applied in R using the tsCV() function and specifying the forecasting model to be used. All 3 models are tested again using time series cross-validation and the results for 1-12 month ahead forecasts are presented in *Table 7*.

RMSE for Different Models and Forecast Using Time Series Cross-Validation

Forecast Horizon (h)	Model	Root Mean Squared Error	Forecast Horizon (h)	Model	Root Mean Squared Error	Forecast Horizon (h)	Model	Root Mean Squared Error
1	RWF	54.22	5	RWF	146.77	9	RWF	214.69
1	ETS (A,A,N)	56.64	5	ETS (A,A,N)	155.90	9	ETS (A,A,N)	238.35
1	ETS (M,N,N)	55.76	5	ETS (M,N,N)	152.79	9	ETS (M,N,N)	220.95
2	RWF	86.29	6	RWF	168.60	10	RWF	228.62
2	ETS (A,A,N)	90.21	6	ETS (A,A,N)	181.68	10	ETS (A,A,N)	254.64
2	ETS (M,N,N)	88.76	6	ETS (M,N,N)	175.04	10	ETS (M,N,N)	235.50
3	RWF	108.10	7	RWF	187.89	11	RWF	245.11
3	ETS (A,A,N)	113.11	7	ETS (A,A,N)	204.87	11	ETS (A,A,N)	275.15
3	ETS (M,N,N)	111.12	7	ETS (M,N,N)	193.84	11	ETS (M,N,N)	251.22
4	RWF	126.07	8	RWF	203.39	12	RWF	259.17
4	ETS (A,A,N)	132.44	8	ETS (A,A,N)	223.84	12	ETS (A,A,N)	292.51
4	ETS (M,N,N)	130.59	8	ETS (M,N,N)	208.89	12	ETS (M,N,N)	263.62

Table 7

Table 7 shows that the random walk with drift model performs better over all the different forecasting horizons. The manually-chosen ETS(A,A,N) model continuously performs

The worst. These results show that a more complicated model isn't necessarily always going to outperform more complicated models. In addition, the ETS(M,N,N) model which was automatically chosen performed better than the manually chosen ETS(A,A,N) model, even though it was seen that the original data had an overall upward trend. The result from the time series cross-validation affirms the random walk with drift to be the best model, showing the consistently lowest RMSE among all forecast horizons. The manually-chosen ETS(A,A,N) performs the worst among all forecast horizons, showing the highest RMSE among all models consistently. The consistent results from both traditional accuracy measures using a single training-test split and the more robust cross-validation technique suggest that the random walk with drift model can be confidently chosen as the champion model.

Forecasting

Now that the random walk with drift has been chosen as the champion model, we can use it to generate forecasts for the next 12 months. This means that the forecasted values will be for months between February 2025 and February 2026. *Figure 8* shows these forecasts.

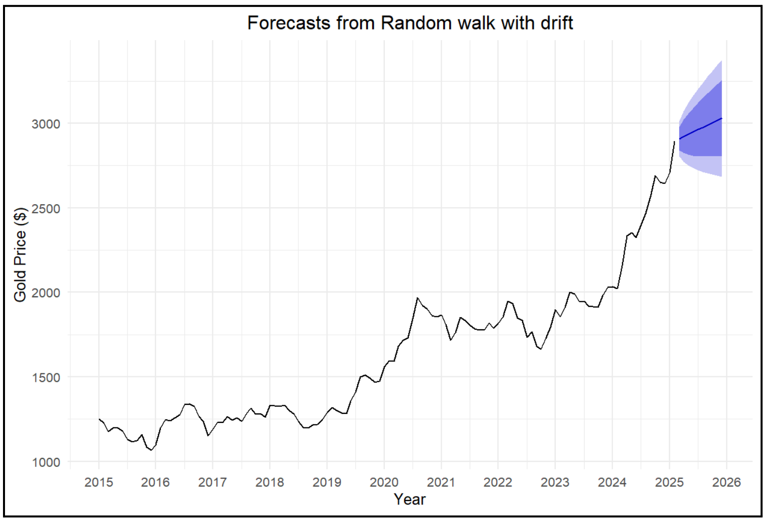


Figure 8

The point forecast (dark purple line) in *Figure 8* shows that the price of gold will continue its upward trend 12 months into the future. This is reasonable given the overall upward trend observed in the time plot for more recent years. The 80% and 95% forecast intervals are given by the dark purple and light purple shaded regions respectively. It is more important to focus on the forecast prediction interval rather than the point forecast.

It is more important to focus on the forecast prediction interval rather than the point forecast. This is because the forecast interval reflects the level of uncertainty associated with our point forecast. A point forecast without a forecast interval is useless because there is no way of quantifying how certain or uncertain the forecast is. Forecast intervals provide a range of values within which we are confident the true future value of the time series would be. In the plot in *Figure 8*, the dark purple and light purple shaded regions are the range of values

within which there is 95% and 80% confidence, respectively, that the true future value would fall in. This quantification of uncertainty in our forecasts would allow stakeholders to make informed decisions and investments based on their level of risk. Rationally, however, the 95% prediction interval does not make full sense. Looking at the lower bound of the 95% prediction interval, it can be seen that there is a possibility that the price of gold may decrease. The reason this is illogical is because there has never been a decrease in the price of gold throughout the observed history. Even in dire economic times, such as the 2008 global financial crisis, the price of gold remained stagnant, but it did not decrease. It may also be argued that the 80% confidence interval may also not be fully accurate. This is because the lower bound indicates that the price of gold may be stagnant in the coming 12 months. This may not be logical as the price of gold has been skyrocketing for the past year.
Based on the predicted forecast and logical analysis, my recommendation for stakeholders is to focus on the middle and upper range of values in the 80% and 95% forecast intervals. While all possible values in the shaded areas are theoretically possible, it is logical to assume that forecasts will lie in the middle/upper regions of both shaded areas for the reasons just discussed above. Actual predicted point forecasts and lower and upper prediction intervals can be seen in *Table 8* below. There are several limitations and risks of this analysis, however. The main limitation of our model is that it assumes a linear model. In reality, and as seen clearly in the original time plot, gold price (or any other commodity) does not necessarily display a linear pattern. This is due to the inherent complexity of economic variables and market sentiment which makes it difficult to capture all nuances via a simple model. Another limitation is that the model doesn't (nor can any model) account for potential future economic shocks.

12-Month Forecast of Gold Prices Using Random Walk with Drift

Month	Point Forecast	Lower 80% Bound	Upper 80% Bound	Lower 95% Bound	Upper 95% Bound
Mar 2025	$2,911.22	$2,840.86	$2,981.58	$2,803.61	$3,018.83
Apr 2025	$2,927.71	$2,827.75	$3,027.66	$2,774.84	$3,080.58
May 2025	$2,944.20	$2,821.23	$3,067.17	$2,756.13	$3,132.27
Jun 2025	$2,960.69	$2,818.06	$3,103.32	$2,742.56	$3,178.82
Jul 2025	$2,977.18	$2,817.01	$3,137.34	$2,732.22	$3,222.13
Aug 2025	$2,993.67	$2,817.45	$3,169.89	$2,724.16	$3,263.17
Sep 2025	$3,010.16	$2,818.99	$3,201.32	$2,717.79	$3,302.52
Oct 2025	$3,026.65	$2,821.40	$3,231.89	$2,712.75	$3,340.54
Nov 2025	$3,043.14	$2,824.51	$3,261.76	$2,708.78	$3,377.49
Dec 2025	$3,059.62	$2,828.20	$3,291.05	$2,705.70	$3,413.56
Jan 2026	$3,076.11	$2,832.38	$3,319.85	$2,703.35	$3,448.88
Feb 2026	$3,092.60	$2,836.97	$3,348.24	$2,701.65	$3,483.56

Table 8

This is due to the model being trained on past data which cannot foresee future events. Lastly, the forecasts were conducted by purely technical analysis; chart patterns were the sole factor in forecasting gold prices. Current economic conditions and fundamental analysis were not incorporated into this analysis, which may have incorporated crucial information about the forecast of the price of gold. These risks and limitations warrant caution for stakeholders interested in leveraging gold as a financial instrument for financial gains.

References

Brixx. (2023). *How Long Should Your Business Forecast Be? - Brixx*. Brixx. https://brixx.com/how-long-should-a-financial-forecast-be/

Hergt, B. (2013). *Gold prices during and after the Great Recession : Beyond the Numbers: U.S. Bureau of Labor Statistics*. Www.bls.gov; U.S Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-2/gold-prices-during-and-after-the-great-recession.htm

Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2.

Katy. (2025). *Tackling nonlinear time series forecasting with STAR models*. Medium; Python's Gurus. https://medium.com/pythons-gurus/tackling-nonlinear-time-series-forecasting-with-star-models-dddeb7612deb

Sakia, R. (1992). The Box-Cox Transformation Technique: A Review. The Statistician. 41. doi: 10.2307/2348250.

The Malaysian Reserve. (2020, December). *News on discovery of COVID-19 vaccine causes drastic drop in gold price*. https://themalaysianreserve.com/2020/12/01/news-on-discovery-of-covid-19-vaccine-causes-drastic-drop-in-gold-price/

World Gold Council | BullionByPost. (2025). Bullionbypost.com. https://www.bullionbypost.com/index/gold/world-gold-council/