

CHAPTER IV

RESULTS

4.1 Overview of Research Results

This chapter presents the results of the development and performance evaluation of the models used for gold price forecasting. The models considered in this study include MIDAS, MIDAS-AR, MIDAS-AR-Threshold, MIDAS-AR-HMM, and MIDAS-AR-HMM-SOM, each evaluated using three lag lengths: $K = 4$, $K = 8$, and $K = 12$. In addition, the ARIMA model is employed as a baseline model for performance comparison.

The performance of the models is evaluated using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Ljung–Box test, and forecasting error measures, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Furthermore, the robustness of the models is examined under different Train:Test ratios, namely 70:30, 75:25, and 80:20.

4.2 Results of ARIMA Model Analysis

The ARIMA model was used as the baseline model for comparison with the other models in this study. Monthly gold price data were used to construct the model, and the data were divided into training and testing sets under different Train ratios, namely 70:30, 75:25, and 80:20. Prior to model construction, the stationarity of the data was examined using the Augmented Dickey–Fuller test (ADF test). The results indicated that the original series was non-stationary, as the p-value was greater than 0.05. Therefore, first-order differencing was applied before the data were used in the model-building process. Subsequently, model order selection in this study was performed using the Auto ARIMA function in Python, which selects the most appropriate model based on the minimum AIC criterion.

4.2.1 Stationarity Testing

Table 4.1 Results of the ADF stationarity test before and after first-order differencing.

Train:Test	Data	ADF Statistic	p-value	Conclusion
70:30	Original Series	-1.474	0.547	Non-Stationary
70:30	First Difference	-3.646	0.005	Stationary
75:25	Original Series	-1.268	0.644	Non-Stationary
75:25	First Difference	-3.708	0.004	Stationary
80:20	Original Series	-0.835	0.809	Non-Stationary
80:20	First Difference	-3.677	0.004	Stationary

Table 4.1 presents the results of the ADF test, which was used to examine the stationarity of the data under different Train:Test ratios. The results show that the original gold price series has p-values greater than 0.05 in all cases, indicating that the data are non-stationary. After applying first-order differencing, the p-values are less than 0.05 in all cases. Therefore, the first-order differenced data can be used for constructing the ARIMA model.

4.2.2 ARIMA Model Selection Results

Table 4.2 presents the results of ARIMA model selection under different Train ratios.

Table 4.2 ARIMA Model Selection Results.

Train:Test	Selected Model	AIC	BIC
70:30	ARIMA(0,1,1)	2267.155	2277.182
75:25	ARIMA(0,1,1)	2425.280	2435.515
80:20	ARIMA(0,1,1)	2601.736	2612.165

Based on the Auto ARIMA procedure implemented in Python, the results indicate that the ARIMA(0,1,1) model is the most appropriate model under all Train ratios.

4.2.3 Residual Diagnostic Checking

The residuals of the model were examined using the Ljung-Box test, which was employed to assess the white noise property of the residuals.

Table 4.3 Ljung–Box Test Results for the ARIMA Model.

Train:Test	Model	Ljung–Box p-value	Conclusion
70:30	ARIMA(0,1,1)	0.295	Pass
75:25	ARIMA(0,1,1)	0.307	Pass
80:20	ARIMA(0,1,1)	0.402	Pass

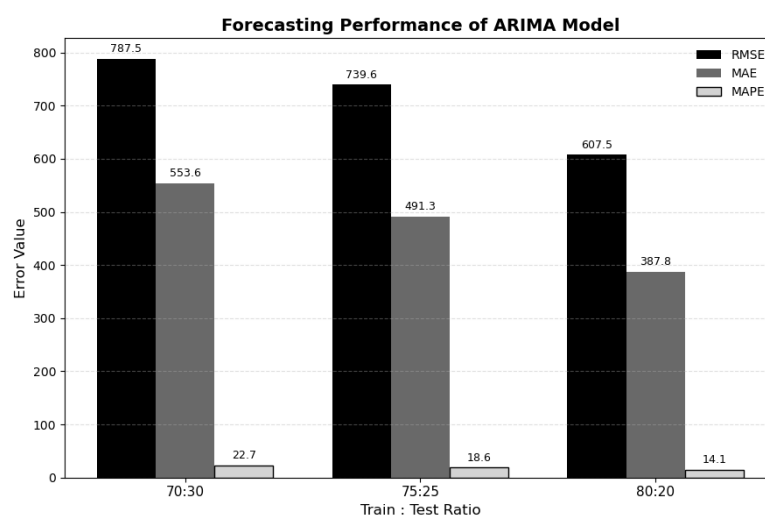
The results indicate that the residuals of the ARIMA model have p-values greater than 0.05 in all cases, suggesting that the residuals do not exhibit statistically significant autocorrelation.

4.2.4 Performance Evaluation Results of the ARIMA Model on the Test Set

To evaluate the performance of the model, forecasting error measures, including RMSE, MAE, and MAPE, were calculated using the test dataset.

Table 4.4 Forecasting Performance of the ARIMA Model.

Train:Test	Model	RMSE	MAE	MAPE
70:30	ARIMA(0,1,1)	787.544	553.646	22.733
75:25	ARIMA(0,1,1)	739.552	491.300	18.592
80:20	ARIMA(0,1,1)	607.538	387.759	14.082

**Figure 4.1** Forecasting Performance of ARIMA Model under Different Train:Test Ratios.

The results indicate that as the proportion of training data decreases, the forecasting errors of the ARIMA model tend to increase. Among the ARIMA models, the Train ratio of 80:20 provides the best forecasting performance.

Table 4.5 Parameter Estimates of the ARIMA(0,1,1) Model for the Train:Test Ratio of 80:20.

Parameter	Coefficient	Std. Error	p-value
Intercept	6.755	3.052	0.027
MA(1)	-0.152	0.042	0.000
Variance (σ^2)	3050.172	197.161	0.000

Table 4.5 presents the parameter estimates of the ARIMA(0,1,1) model for the Train ratio of 80:20. The results show that the MA(1) parameter is statistically significant at the 0.05 significance level, with a p-value less than 0.05. In addition, the Ljung–Box test yields a p-value of 0.402, indicating that the residuals of the model do not exhibit significant autocorrelation. The estimated ARIMA(0,1,1) model can be expressed as follows:

$$\Delta Y_t = 6.755 - 0.152\varepsilon_{t-1} + \varepsilon_t.$$

4.3 Results of the MIDAS Model

This section presents the forecasting results of the MIDAS model, which uses gold prices as the dependent variable and the DXY index together with oil prices as independent variables. The main advantage of the MIDAS model is its ability to incorporate variables with different frequencies into the estimation and forecasting process. In this study, the dependent variable, gold price, is observed at a monthly frequency, while the independent variables, the DXY index and oil prices, are observed at a weekly frequency. Three lag structures are considered: $K = 4$, $K = 8$, and $K = 12$.

4.3.1 Analysis of AIC and BIC Values for the MIDAS Model

Table 4.6 AIC and BIC Values of the MIDAS Model.

Train:Test	$K = 4$		$K = 8$		$K = 12$	
	AIC	BIC	AIC	BIC	AIC	BIC
70:30	1294.700	1312.200	1294.200	1311.700	1293.110	1310.610
75:25	1087.799	1104.022	1098.268	1114.491	1086.048	1102.270
80:20	876.970	891.630	884.475	899.135	880.730	895.397

The results indicate that the MIDAS model with the lowest AIC value differs across the Train:Test ratios. For the Train:Test ratios of 70:30 and 75:25, the model with $K = 12$

yields the lowest AIC value, whereas for the Train:Test ratio of 80:20, the model with $K = 4$ yields the lowest AIC value.

When considering the BIC values, the model with $K = 12$ also produces the lowest BIC value for the Train:Test ratios of 70:30 and 75:25. In contrast, for the Train:Test ratio of 80:20, the model with $K = 4$ yields the lowest BIC value.

4.3.2 Residual Diagnostic Checking

The residuals of the MIDAS model were examined using the Ljung–Box test to assess the white noise property of the residuals.

Table 4.7 Ljung–Box Test Results for the MIDAS Model.

Train:Test	$K = 4$	$K = 8$	$K = 12$
	Ljung–Box p-value	Ljung–Box p-value	Ljung–Box p-value
70:30	4.42×10^{-72}	3.98×10^{-72}	3.86×10^{-71}
75:25	2.34×10^{-58}	5.41×10^{-61}	2.55×10^{-55}
80:20	2.74×10^{-51}	2.41×10^{-52}	3.30×10^{-51}

The Ljung–Box test results indicate that the residuals of the MIDAS model exhibit autocorrelation in all cases.

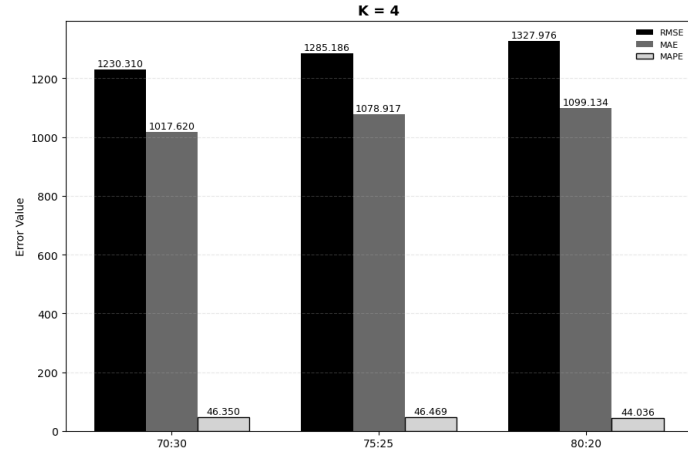
4.3.3 Performance Evaluation of the Model on the Test Dataset

Table 4.8 Performance Evaluation Results of the MIDAS Model under Different Train:Test Ratios.

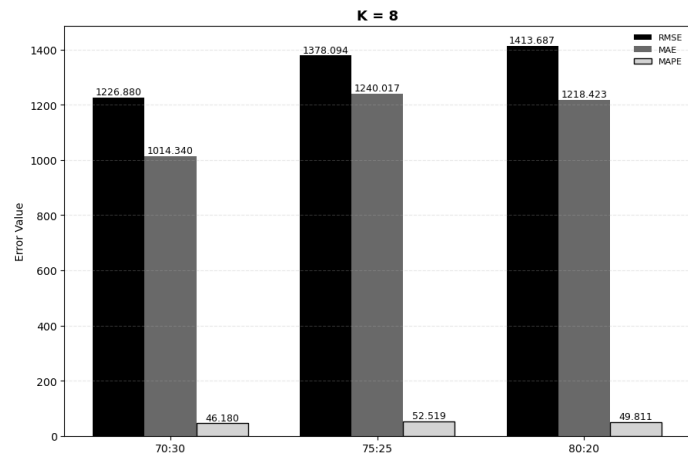
Train:Test	$K = 4$			$K = 8$			$K = 12$		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
70:30	1230.310	1017.620	46.350	1226.880	1014.340	46.180	1219.480	1009.940	45.830
75:25	1285.186	1078.917	46.469	1378.094	1240.017	52.519	1270.270	1070.580	46.205
80:20	1327.976	1099.134	44.036	1413.687	1218.423	49.811	1370.329	1167.458	47.477

The results indicate that the MIDAS model yields relatively high forecasting errors in all cases. When comparing different lag lengths, the model with $K = 4$ provides the best forecasting performance for the Train:Test ratio of 80:20, with an RMSE of 1327.976, an MAE of 1099.134, and a MAPE of 44.036%.

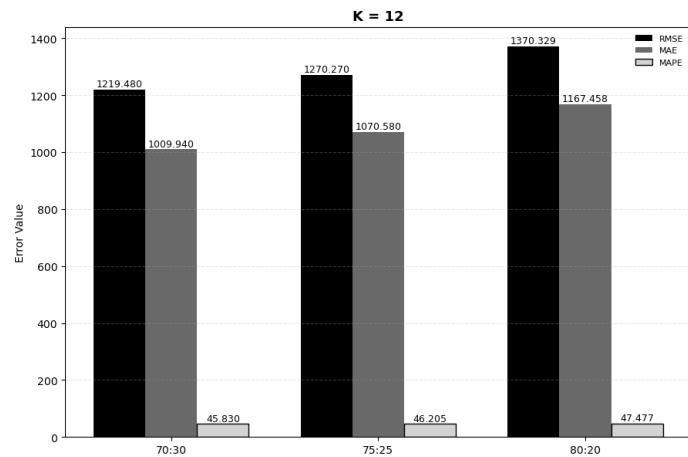
For the Train:Test ratios of 70:30 and 75:25, the model with $K = 12$ yields the lowest forecasting errors when compared with the cases of $K = 4$ and $K = 8$.



(a) $K = 4$



(b) $K = 8$



(c) $K = 12$

Figure 4.2 Forecasting Performance of the MIDAS Model under Different K Values and Train:Test Ratios.

4.3.4 The Selected MIDAS Model

Based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and forecasting error measures, including RMSE, MAE, and MAPE, the MIDAS model with $K = 4$ under the Train:Test ratio of 80:20 provides the best forecasting performance on the test dataset. In particular, this model yields the lowest MAPE value compared with the other cases. In addition, this case utilizes a larger proportion of training data. Therefore, the MIDAS model with $K = 4$ under the Train:Test ratio of 80:20 is selected as the representative model for presenting the structure of the MIDAS model.

The general form of the MIDAS model can be expressed as follows:

$$Y_t = \beta_0 + \beta_1 \sum_{k=1}^4 w_{1,k}(\nu_1) X_{t-k/m}^{(DXY)} + \beta_2 \sum_{k=1}^4 w_{2,k}(\nu_2) X_{t-k/m}^{(OIL)} + \varepsilon_t,$$

where Y_t denotes the gold price at time t , while $X^{(DXY)}$ and $X^{(OIL)}$ represent the DXY index and oil price variables, respectively. The terms $w_{1,k}(\nu_1)$ and $w_{2,k}(\nu_2)$ denote the lag weights in the MIDAS model.

In this study, the lag weights are estimated using the Exponential Almon Lag Polynomial. The estimated parameters of the MIDAS model that provides the best forecasting performance are presented in Appendix A.1.

Therefore, in the next section, the MIDAS-AR model is developed by incorporating an autoregressive component into the model structure for forecasting gold prices.

4.4 Results of the MIDAS-AR Model

This section presents the forecasting results of the MIDAS-AR model, which is an extension of the MIDAS model obtained by incorporating an autoregressive component into the MIDAS model structure.

4.4.1 Analysis of AIC and BIC Values for the MIDAS-AR Model

Table 4.9 shows that for the Train:Test ratios of 70:30 and 75:25, the MIDAS-AR model with $K = 4$ yields the lowest AIC and BIC values. In contrast, for the Train:Test

Table 4.9 AIC and BIC Values of the MIDAS-AR Model.

Train:Test	$K = 4$		$K = 8$		$K = 12$	
	AIC	BIC	AIC	BIC	AIC	BIC
70:30	1688.220	1714.960	1877.386	1904.124	1874.809	1901.560
75:25	1803.926	1831.219	1928.653	1955.946	2010.140	2037.397
80:20	1936.791	1964.603	2166.121	2193.933	1930.380	1958.158

ratio of 80:20, the model with $K = 12$ produces the lowest AIC and BIC values.

4.4.2 Residual Diagnostic Checking

Table 4.10 Ljung–Box Test Results for the MIDAS-AR Model.

Train:Test	$K = 4$	$K = 8$	$K = 12$
	Ljung–Box p-value	Ljung–Box p-value	Ljung–Box p-value
70:30	0.019	0.021	0.004
75:25	0.025	8.77×10^{-7}	1.8×10^{-4}
80:20	0.036	9.76×10^{-8}	0.055

Table 4.10 shows that although the MIDAS-AR model with $K = 4$ under the Train:Test ratios of 70:30 and 75:25 yields the lowest AIC and BIC values, the Ljung–Box test indicates that the residuals still exhibit autocorrelation, since the p-values are less than 0.05. In contrast, for the Train:Test ratio of 80:20 with $K = 12$, the p-value of the Ljung–Box statistic is greater than 0.05, indicating that the residuals do not exhibit statistically significant autocorrelation.

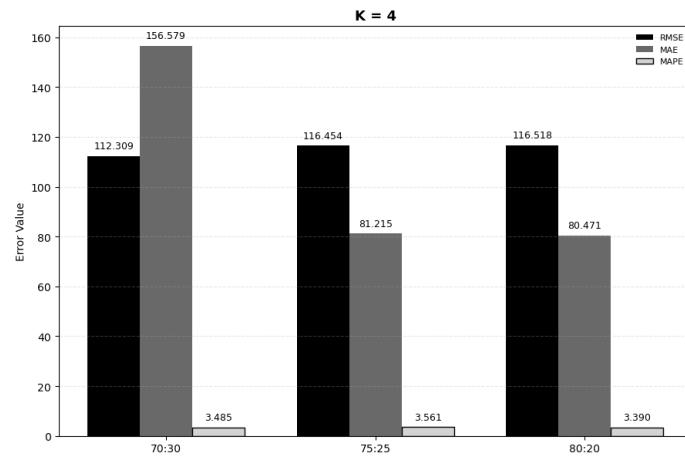
4.4.3 Performance Evaluation of the Model on the Test Dataset

Table 4.11 Performance Evaluation Results of the MIDAS-AR Model under Different Train:Test Ratios.

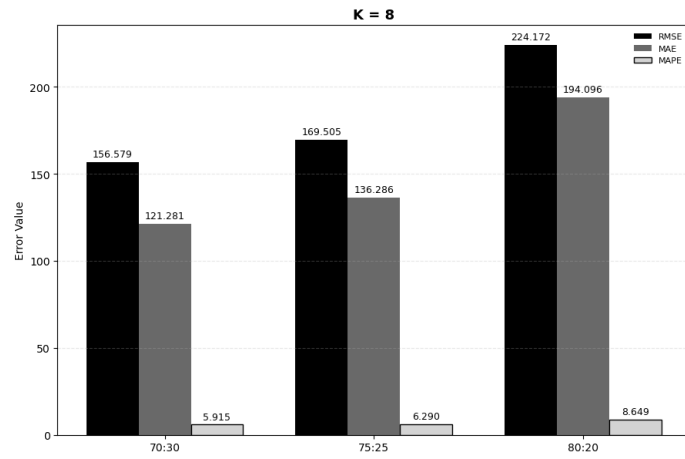
Train:Test	$K = 4$			$K = 8$			$K = 12$		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
70:30	112.309	156.579	3.485	156.579	121.281	5.915	181.826	143.373	6.769
75:25	116.454	81.215	3.561	169.505	136.286	6.290	204.205	168.575	7.647
80:20	116.518	80.471	3.390	224.172	194.096	8.649	114.239	79.739	3.383

The performance evaluation results indicate that the MIDAS-AR model substantially reduces forecasting errors compared with the MIDAS model. In particular, the case

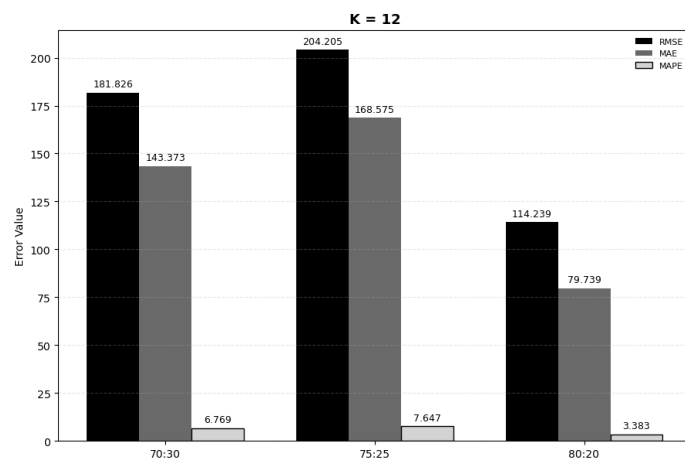
with a Train:Test ratio of 80:20 and $K = 12$ yields lower RMSE, MAE, and MAPE values.



(a) $K = 4$



(b) $K = 8$



(c) $K = 12$

Figure 4.3 Forecasting Performance of the MIDAS-AR Model under Different K Values and Train:Test Ratios.

4.4.4 The Selected MIDAS-AR Model

Based on the AIC, BIC, and forecasting error measures, including RMSE, MAE, and MAPE, the MIDAS-AR model with $K = 12$ under the Train:Test ratio of 80:20 provides the most appropriate forecasting performance in this study. Therefore, the MIDAS-AR model with $K = 12$ under this Train:Test ratio is selected as the representative model for presenting the structure of the MIDAS-AR model.

The general form of the MIDAS-AR model can be expressed as follows:

$$Y_t = \beta_0 + \phi_1 Y_{t-1} + \beta_1 \sum_{k=1}^{12} w_{1,k}(\nu_1) X_{t-k/m}^{(DXY)} + \beta_2 \sum_{k=1}^{12} w_{2,k}(\nu_2) X_{t-k/m}^{(OIL)} + \varepsilon_t,$$

where Y_t denotes the gold price at time t , and Y_{t-1} denotes the lagged gold price. The variables $X^{(DXY)}$ and $X^{(OIL)}$ represent the DXY index and oil price, respectively. The terms $w_{1,k}(\nu_1)$ and $w_{2,k}(\nu_2)$ denote the lag weights in the MIDAS-AR model, and ε_t represents the model error term. In this study, the lag weights are estimated using the Exponential Almon Lag Polynomial. The estimated parameters of the MIDAS-AR model that provides the best forecasting performance are presented in Appendix A.2.

In the next section, the MIDAS-AR-Threshold model is developed to investigate the performance of the model under regime changes.

4.5 Results of the MIDAS-AR-Threshold Model

This section presents the forecasting results of the MIDAS-AR-Threshold model, which incorporates regime-switching behavior by using a gold price threshold to divide the data into different regimes.

4.5.1 Analysis of AIC and BIC Values for the MIDAS-AR-Threshold Model

The information criteria results for the MIDAS-AR-Threshold model are presented in 4.12.

Table 4.12 AIC and BIC Values of the MIDAS-AR-Threshold Model.

Train:Test	$K = 4$		$K = 8$		$K = 12$	
	AIC	BIC	AIC	BIC	AIC	BIC
70:30	2293.925	2350.745	2293.184	2350.003	2286.862	2343.600
75:25	2452.783	2510.781	2453.081	2511.079	2443.262	2501.184
80:20	2628.081	2687.181	2631.032	2690.132	2620.818	2679.847

Table 4.12 shows that for the Train:Test ratios of 70:30, 75:25, and 80:20, the MIDAS-AR-Threshold model with $K = 12$ yields the lowest AIC and BIC values.

4.5.2 Residual Diagnostic Checking

Table 4.13 Ljung–Box Test Results for the MIDAS-AR-Threshold Model.

Train:Test	$K = 4$	$K = 8$	$K = 12$
	Ljung–Box p-value	Ljung–Box p-value	Ljung–Box p-value
70:30	0.089	0.072	0.039
75:25	0.067	0.051	0.039
80:20	0.029	0.038	0.056

Table 4.13 shows that although the MIDAS-AR-Threshold model with $K = 12$ under the Train:Test ratios of 70:30 and 75:25 yields the lowest AIC and BIC values, the Ljung–Box test indicates that the residuals still exhibit autocorrelation, since the p-value is 0.039, which is less than 0.05.

In contrast, for the Train:Test ratio of 80:20 with $K = 12$, the Ljung–Box statistic yields a p-value of 0.056, which is greater than 0.05, indicating that the residuals do not exhibit autocorrelation.

4.5.3 Performance Evaluation of the Model on the Test Dataset

Table 4.14 Performance Evaluation Results of the MIDAS-AR-Threshold Model under Different Train:Test Ratios.

Train:Test	$K = 4$			$K = 8$			$K = 12$		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
70:30	156.643	106.211	4.574	158.018	106.949	4.598	135.126	90.721	3.964
75:25	149.778	104.857	4.375	140.800	98.276	4.127	136.735	95.454	4.029
80:20	119.018	82.845	3.467	119.025	82.845	3.467	119.924	83.354	3.478

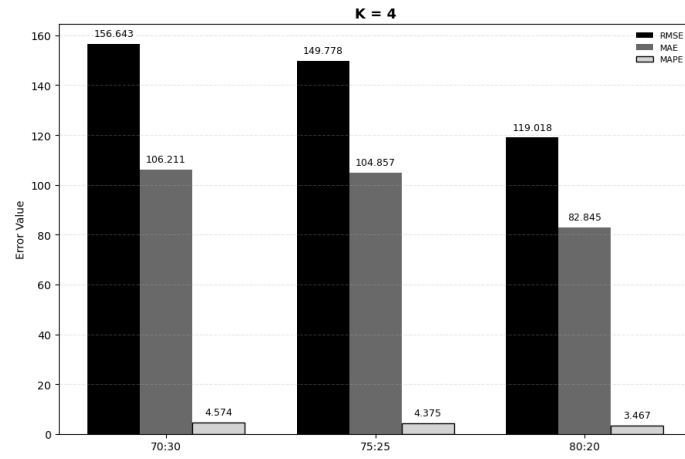
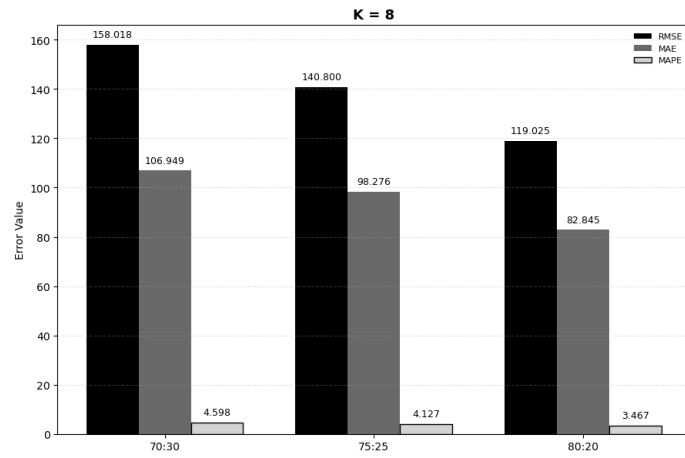
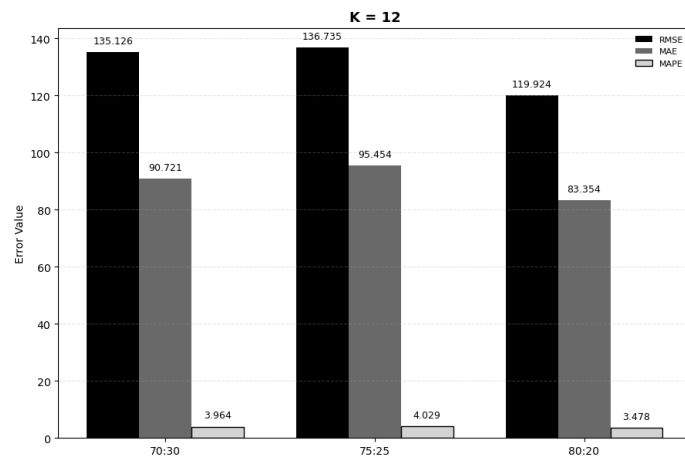
(a) $K = 4$ (b) $K = 8$ (c) $K = 12$

Figure 4.4 Forecasting Performance of the MIDAS-AR-Threshold Model under Different K Values and Train:Test Ratios.

4.5.4 Regime Determination and the Selected MIDAS-AR-Threshold Model

Based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Ljung–Box test results, and forecasting error measures, including RMSE, MAE, and MAPE, the MIDAS-AR-Threshold model with $K = 12$ under the Train:Test ratio of 80:20 provides the most appropriate forecasting performance. This model yields the lowest AIC and BIC values, does not exhibit autocorrelation in the residuals, and produces relatively low forecasting errors compared with the other cases.

In this study, the MIDAS-AR-Threshold model uses the one-period lagged gold price (Y_{t-1}) as the threshold variable for regime classification, defined as

$$\text{Regime}_t = \begin{cases} 1, & Y_{t-1} \leq c, \\ 2, & Y_{t-1} > c, \end{cases}$$

where c denotes the threshold parameter estimated from the model estimation process.

Table 4.15 Threshold Value of the MIDAS-AR-Threshold Model.

Train:Test	K	Threshold	Regime 1	Regime 2
80:20	12	1171.3	$Y_{t-1} \leq c$	$Y_{t-1} > c$

The threshold value reported in Table 4.15 was selected as the value that minimizes the residual sum of squares (RSS) during the threshold estimation procedure.

The MIDAS-AR-Threshold model can then be expressed in the following general form:

$$Y_t = \begin{cases} \beta_{0,1} + \phi_1 Y_{t-1} + \beta_{1,1} \sum_{k=1}^{12} w_{1,k}(\nu_{1,1}) X_{t-k/m}^{DXY} + \beta_{2,1} \sum_{k=1}^{12} w_{2,k}(\nu_{2,1}) X_{t-k/m}^{OIL} + \varepsilon_t, & Y_{t-1} \leq c, \\ \beta_{0,2} + \phi_2 Y_{t-1} + \beta_{1,2} \sum_{k=1}^{12} w_{1,k}(\nu_{1,2}) X_{t-k/m}^{DXY} + \beta_{2,2} \sum_{k=1}^{12} w_{2,k}(\nu_{2,2}) X_{t-k/m}^{OIL} + \varepsilon_t, & Y_{t-1} > c, \end{cases}$$

where Y_t denotes the gold price at time t , and Y_{t-1} denotes the lagged gold price. The variables $X^{(DXY)}$ and $X^{(OIL)}$ represent the DXY index and oil price, respectively. The

terms $w_{1,k}$ and $w_{2,k}$ denote the lag weights in the MIDAS-AR-Threshold model ($i = 1, 2$), and ε_t represents the model error term. The estimated parameters of the MIDAS-AR-Threshold model that provides the best forecasting performance are presented in Appendix A.3.

4.6 Results of the MIDAS-AR-HMM Model

This section presents the forecasting results of the MIDAS-AR-HMM model, which is an extension of the MIDAS-AR model by incorporating a Hidden Markov Model (HMM) to classify market regimes before estimating the MIDAS-AR model within each regime. In this study, regime classification is based on gold returns and the volatility of gold returns as observed variables, in order to capture the changing characteristics of the gold market over time.

4.6.1 Regime Classification Using the Hidden Markov Model

In this study, market regimes are not determined by a fixed threshold as in the MIDAS-AR-Threshold model. Instead, a Hidden Markov Model (HMM) is used to learn market states from the statistical characteristics of gold prices. The observed variables consist of gold returns and volatility, defined as

$$D_t = (r_t, \sigma_t),$$

where $r_t = \ln \left(\frac{Y_t}{Y_{t-1}} \right)$ denotes the logarithmic return of the gold price at time t , where Y_t represents the gold price at time t .

The variable σ_t denotes volatility, which is calculated using the rolling standard deviation of gold returns as follows:

$$\sigma_t = \sqrt{\frac{1}{m-1} \sum_{i=0}^{m-1} (r_{t-i} - \bar{r})^2},$$

where m denotes the rolling window, which is set to $m = 3$ in this study.

In this study, the Gaussian Hidden Markov Model is specified with two hidden states to represent different market regimes. The HMM learns the patterns in the data from the distribution of gold returns and volatility while considering the temporal continuity of regimes.

In addition, the HMM uses a Transition Probability Matrix to describe the probabilities of remaining in or switching between regimes over time, which can be expressed as follows:

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix},$$

where p_{ij} denotes the probability that the market transitions from regime i to regime j in the next period.

For this study, the estimated transition probability matrix, whose parameters are estimated from the training data, is given by

$$P = \begin{pmatrix} 0.977 & 0.023 \\ 0.257 & 0.743 \end{pmatrix}.$$

The estimated transition probability matrix shows that the probabilities on the main diagonal (diagonal entries) are relatively high, particularly $p_{11} = 0.977$, indicating that once the market is in Regime 1, it is highly likely to remain in the same regime in the following period. Similarly, $p_{22} = 0.743$ indicates that Regime 2 also exhibits persistence over time, although it has a higher probability of switching back to Regime 1 compared with Regime 1.

4.6.2 Analysis of AIC and BIC Values for the MIDAS-AR-HMM Model

The information criteria results for the MIDAS-AR-HMM model are presented in Table 4.16.

Table 4.16 AIC and BIC Values of the MIDAS-AR-HMM Model.

Train:Test	$K = 4$		$K = 8$		$K = 12$	
	AIC	BIC	AIC	BIC	AIC	BIC
70:30	1668.592	1721.915	1697.447	1750.771	1665.408	1708.731
75:25	1783.724	1838.167	1813.046	1867.489	1771.150	1825.593
80:20	1907.645	1963.066	1934.989	1990.410	1894.908	1950.330

Table 4.16 shows that for the Train:Test ratios of 70:30, 75:25, and 80:20, the MIDAS-AR-HMM model with $K = 12$ yields the lowest AIC and BIC values.

4.6.3 Residual Diagnostic Checking

Table 4.17 Ljung–Box Test Results for the MIDAS-AR-HMM Model.

Train:Test	$K = 4$	$K = 8$	$K = 12$
	Ljung–Box p-value	Ljung–Box p-value	Ljung–Box p-value
70:30	0.960	0.954	0.950
75:25	0.041	0.056	0.099
80:20	0.714	0.651	0.645

Table 4.17 shows that for the Train:Test ratio of 75:25 with $K = 4$, the Ljung–Box test indicates that the residuals exhibit autocorrelation, since the p-value is less than 0.05.

4.6.4 Performance Evaluation of the Model on the Test Dataset

Table 4.18 Performance Evaluation Results of the MIDAS-AR-HMM Model under Different Train:Test Ratios.

Train:Test	$K = 4$			$K = 8$			$K = 12$		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
70:30	96.557	70.935	3.423	99.446	72.093	3.470	101.186	73.457	3.543
75:25	103.617	76.949	3.496	104.654	77.255	3.512	106.471	78.866	3.592
80:20	110.816	80.304	3.464	109.375	78.0039	3.375	111.089	79.396	3.443

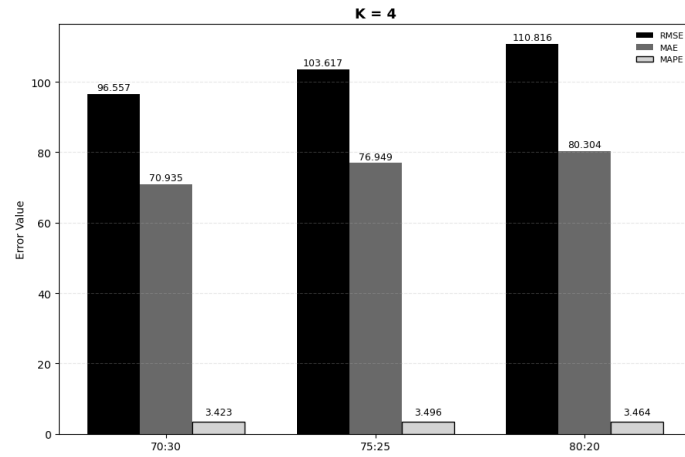
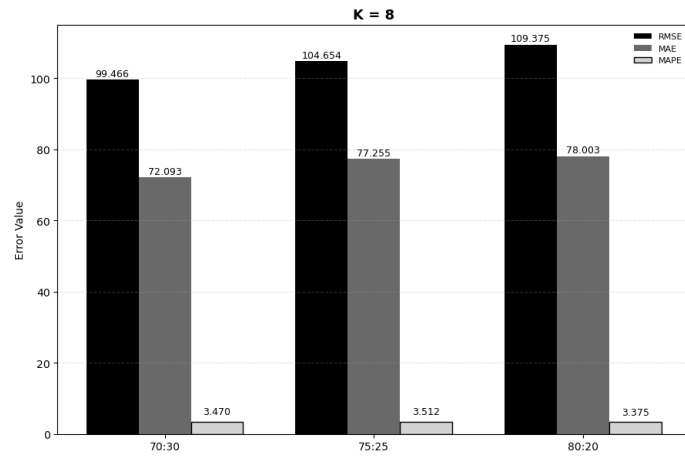
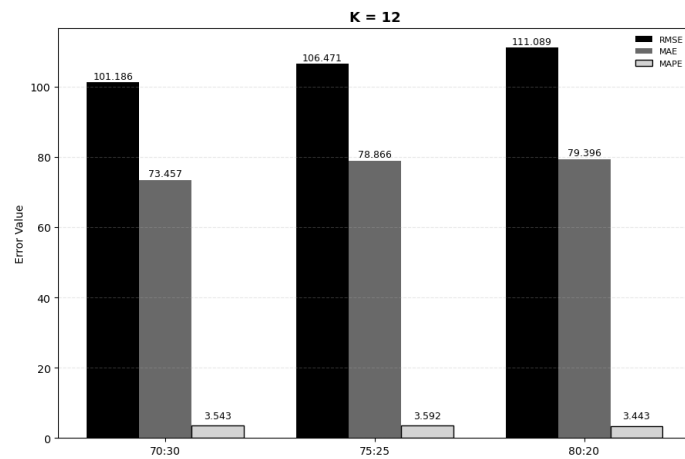
(a) $K = 4$ (b) $K = 8$ (c) $K = 12$

Figure 4.5 Forecasting Performance of the MIDAS-AR-HMM Model under Different K Values and Train:Test Ratios.

4.6.5 The Selected MIDAS-AR-HMM Model

Based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and forecasting error measures, including RMSE, MAE, and MAPE, the MIDAS-AR-HMM model provides good forecasting performance in several cases. However, the model with $K = 4$ under the Train:Test ratio of 70:30 provides the best overall forecasting performance.

This model yields an RMSE of 96.557, an MAE of 70.935, and a MAPE of 3.423, which are the lowest forecasting errors compared with most of the other models. In addition, the Ljung–Box test indicates that the residuals do not exhibit autocorrelation, since the p-value is greater than 0.05.

Although the case with a Train:Test ratio of 80:20 and $K = 8$ yields the lowest MAPE value of 3.375, the overall forecasting errors in terms of RMSE and MAE remain higher than those of the model with $K = 4$ under the Train:Test ratio of 70:30. Therefore, the MIDAS-AR-HMM model with $K = 4$ under the Train:Test ratio of 70:30 is selected as the most appropriate model for representing the structure of the MIDAS-AR-HMM model in this study.

The general form of the MIDAS-AR-HMM model can be expressed as follows:

$$Y_t = \begin{cases} \beta_{0,1} + \phi_1 Y_{t-1} + \\ \beta_{1,1} \sum_{k=1}^4 w_{1,k}(\nu_{1,1}) X_{t-k/m}^{DXY} + \beta_{2,1} \sum_{k=1}^4 w_{2,k}(\nu_{2,1}) X_{t-k/m}^{OIL} + \varepsilon_t, & S_t = 1, \\ \beta_{0,2} + \phi_2 Y_{t-1} + \\ \beta_{1,2} \sum_{k=1}^4 w_{1,k}(\nu_{1,2}) X_{t-k/m}^{DXY} + \beta_{2,2} \sum_{k=1}^4 w_{2,k}(\nu_{2,2}) X_{t-k/m}^{OIL} + \varepsilon_t, & S_t = 2. \end{cases}$$

where S_t denotes the hidden regime at time t , ($S_t = 1, 2$), determined by the Hidden Markov Model. The terms $w_{1,k}$ and $w_{2,k}$ denote the lag weights under the Exponential Almon Lag Polynomial for the DXY index and oil price variables, respectively. The estimated parameters of the MIDAS-AR-HMM model that provides the best forecasting performance are presented in Appendix A.4.

4.7 Results of the MIDAS-AR-HMM-SOM Model

4.7.1 Regime Classification Using the Self-Organizing Map and Hidden Markov Model

In this study, the MIDAS-AR-HMM-SOM model is developed from the MIDAS-AR-HMM model by incorporating the Self-Organizing Map (SOM) technique to cluster data characteristics before applying the Hidden Markov Model (HMM), with the aim of improving the classification of gold market regimes.

The variables used in the analysis consist of the logarithmic return of gold prices, volatility, and lagged return, which can be expressed as follows:

$$D_t = (r_t, \sigma_t, r_{t-1}),$$

where r_t denotes the logarithmic return of the gold price at time t , σ_t denotes volatility, which is calculated using the rolling standard deviation of logarithmic returns, and r_{t-1} denotes the lagged return of the gold price.

Before entering the SOM process, all data are standardized using the Standardization method to reduce the effect of differences in data scales. The standardized data are then processed by the Self-Organizing Map (SOM) to group observations with similar characteristics. The resulting clusters are subsequently used as inputs to the Gaussian Hidden Markov Model (HMM) for classifying the hidden regimes of the gold market.

In this study, the Gaussian Hidden Markov Model is specified with two hidden states to represent different market regimes. The HMM learns the probability distribution of the data under each regime while considering the temporal continuity of regimes. In addition, the HMM uses a Transition Probability Matrix to describe the probabilities of remaining in or switching between regimes over time, which can be expressed as follows:

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix},$$

where p_{ij} denotes the probability that the market transitions from regime i to regime j

in the next period.

For this study, the estimated Transition Probability Matrix is given by

$$P = \begin{pmatrix} 0.468 & 0.532 \\ 0.299 & 0.701 \end{pmatrix}.$$

4.7.2 Analysis of AIC and BIC Values for the MIDAS-AR-HMM-SOM Model

The information criteria results for the MIDAS-AR-HMM-SOM model are presented in Table 4.19.

Table 4.19 AIC and BIC Values of the MIDAS-AR-HMM-SOM Model.

Train:Test	$K = 4$		$K = 8$		$K = 12$	
	AIC	BIC	AIC	BIC	AIC	BIC
70:30	1515.225	1568.548	1509.521	1562.844	1512.470	1565.793
75:25	1622.531	1676.974	1617.403	1671.846	1620.268	1674.711
80:20	1737.543	1792.964	1737.445	1792.866	1740.144	1795.565

Table 4.19 shows that the MIDAS-AR-HMM-SOM model with $K = 8$ yields the lowest AIC and BIC values under all Train:Test ratios. This indicates that, based on the information criteria, the model with $K = 8$ is more appropriate than the other cases in terms of balancing goodness of fit and model complexity.

4.7.3 Residual Diagnostic Checking

Table 4.20 Ljung–Box Test Results for the MIDAS-AR-HMM-SOM Model.

Train:Test	$K = 4$	$K = 8$	$K = 12$
	Ljung–Box p-value	Ljung–Box p-value	Ljung–Box p-value
70:30	0.053	0.067	0.126
75:25	0.041	0.056	0.099
80:20	0.121	0.137	0.185

Table 4.20 shows that the Ljung–Box test results for the MIDAS-AR-HMM-SOM model generally produce p-values greater than 0.05, indicating that the residuals of most models do not exhibit statistically significant autocorrelation. In particular, the cases with

$K = 8$ and $K = 12$ satisfy the Ljung–Box test criterion under all Train:Test ratios. In contrast, the case with a Train:Test ratio of 75:25 and $K = 4$ yields a p-value of 0.041, which is slightly less than 0.05.

4.7.4 Performance Evaluation of the Model on the Test Dataset

Table 4.21 Performance Evaluation Results of the MIDAS-AR-HMM-SOM Model on the Test Dataset.

Train:Test	$K = 4$			$K = 8$			$K = 12$		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
70:30	61.899	43.973	2.133	69.601	49.840	2.379	66.677	49.931	2.417
75:25	64.553	46.790	2.129	71.837	52.559	2.351	68.986	52.340	2.373
80:20	65.832	46.750	1.979	70.106	49.365	2.053	67.355	48.695	2.049

Table 4.21 shows that the MIDAS-AR-HMM-SOM model produces low forecasting errors in all cases, with MAPE values below 3% for all Train:Test ratios. Based on the forecasting error measures, the model with $K = 4$ yields the lowest RMSE and MAE values for the Train:Test ratios of 70:30 and 75:25. In contrast, for the Train:Test ratio of 80:20, the model with $K = 4$ produces the lowest MAPE value of 1.979.

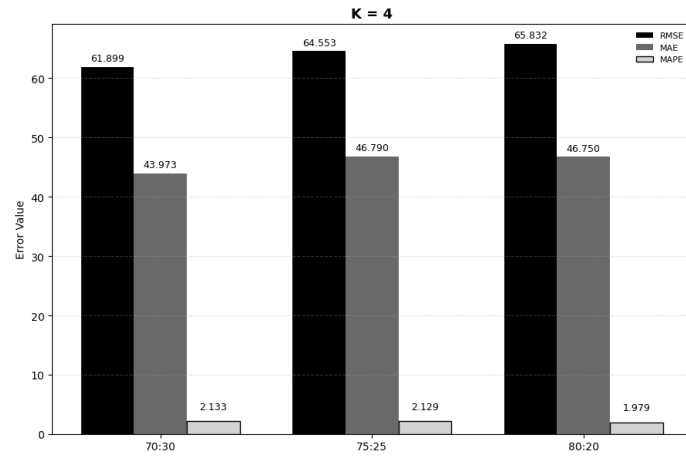
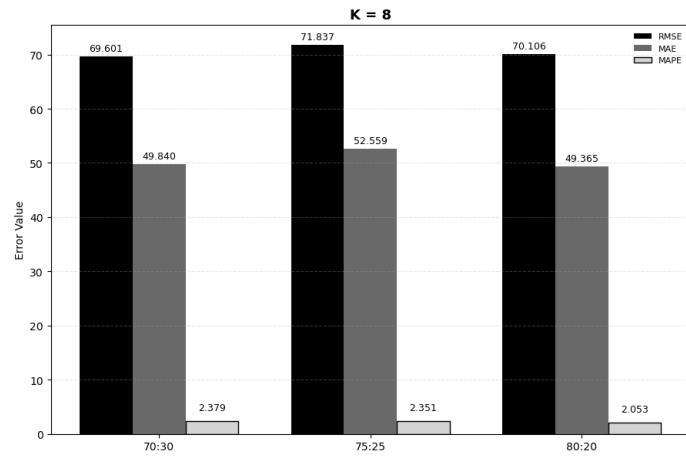
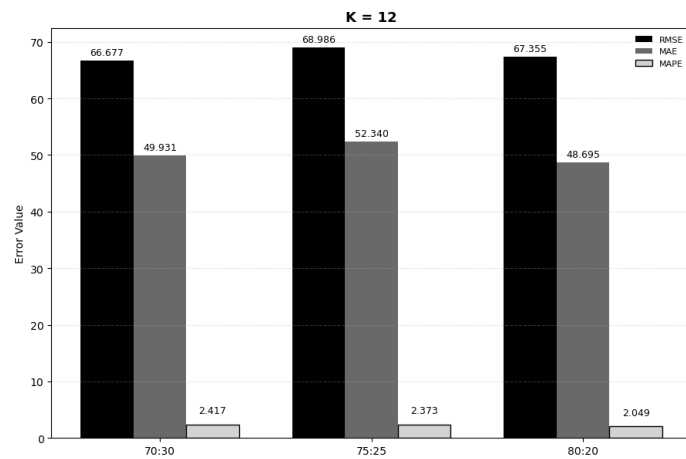
(a) $K = 4$ (b) $K = 8$ (c) $K = 12$

Figure 4.6 Forecasting Performance of the MIDAS-AR-HMM-SOM Model under Different K Values and Train:Test Ratios.

4.7.5 Selection of the Most Appropriate MIDAS-AR-HMM-SOM Model

Based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Ljung–Box test results, and forecasting error measures, including RMSE, MAE, and MAPE, the MIDAS-AR-HMM-SOM model provides good better forecasting performance under all data partition schemes.

The AIC and BIC results indicate that the model with $K = 8$ yields the lowest information criteria values under all Train:Test ratios. However, when the forecasting performance on the test dataset is also considered, the model with $K = 4$ produces lower forecasting errors in several cases. In particular, for the Train:Test ratio of 70:30, the model with $K = 4$ yields an RMSE of 61.899, an MAE of 43.973, and a MAPE of 2.133, which are the lowest RMSE and MAE values among all models. For the Train:Test ratio of 80:20, the model with $K = 4$ produces the lowest MAPE value of 1.979.

In addition, the Ljung–Box test for the case with a Train:Test ratio of 70:30 and $K = 4$ yields a p-value of 0.053, which is greater than 0.05, indicating that the residuals do not exhibit autocorrelation.

Therefore, the MIDAS-AR-HMM-SOM model with $K = 4$ under the Train:Test ratio of 70:30 is selected as the most appropriate model in this study, since it provides the lowest overall forecasting errors and satisfies the residual diagnostic criterion.

The general form of the MIDAS-AR-HMM-SOM model can be expressed as follows:

$$Y_t = \begin{cases} \beta_{0,1} + \phi_1 Y_{t-1} + \\ \beta_{1,1} \sum_{k=1}^4 w_{1,k}(\nu_{1,1}) X_{t-k/m}^{DXY} + \beta_{2,1} \sum_{k=1}^4 w_{2,k}(\nu_{2,1}) X_{t-k/m}^{OIL} + \varepsilon_t, & S_t = 1, \\ \beta_{0,2} + \phi_2 Y_{t-1} + \\ \beta_{1,2} \sum_{k=1}^4 w_{1,k}(\nu_{1,2}) X_{t-k/m}^{DXY} + \beta_{2,2} \sum_{k=1}^4 w_{2,k}(\nu_{2,2}) X_{t-k/m}^{OIL} + \varepsilon_t, & S_t = 2. \end{cases}$$

where S_t denotes the hidden regime at time t , ($S_t = 1, 2$), obtained from the Self-Organizing Map combined with the Hidden Markov Model. The terms $w_{1,k}$ and $w_{2,k}$ denote the lag weights under the Exponential Almon Lag Polynomial for the DXY index and oil price variables, respectively. The estimated parameters of the MIDAS-AR-HMM-

SOM model that provides the best forecasting performance are presented in Appendix A.5.

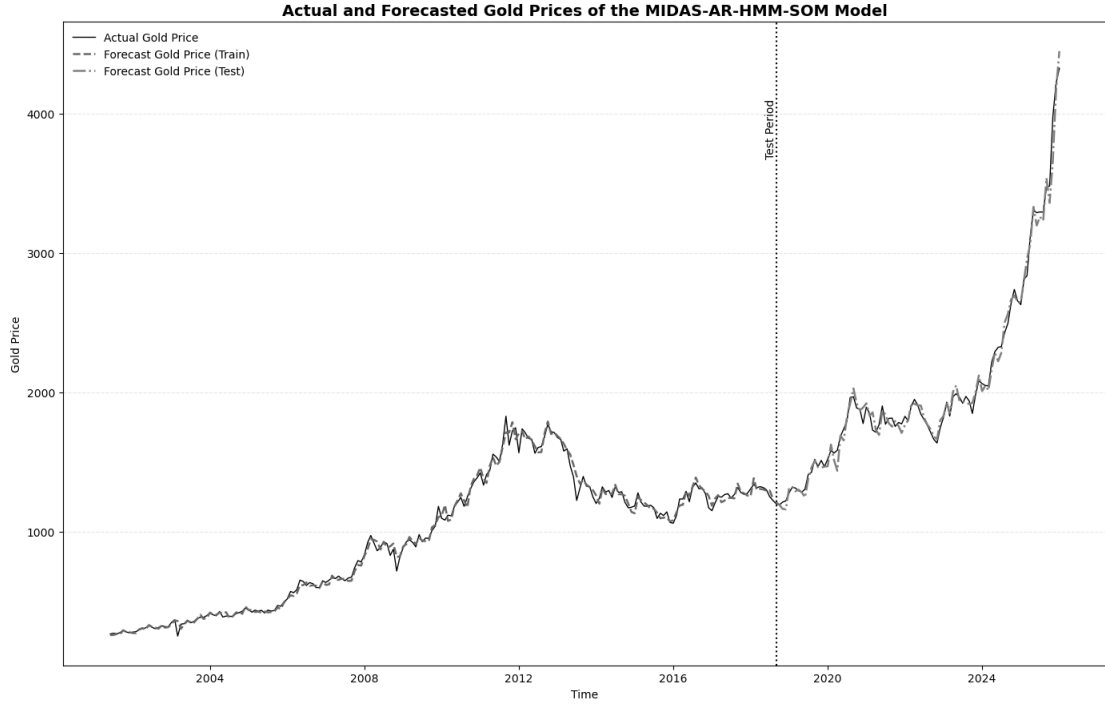


Figure 4.7 Actual and Forecasted Gold Prices of the MIDAS-AR-HMM-SOM Model during the Training and Testing Periods.

4.8 Comparison of Results under Different Train:Test Ratios

In this section, the forecasting performance of the models is compared under different Train:Test ratios, namely 80:20, 75:25, and 70:30. The Diebold–Mariano Test (DM Test) is employed to examine differences in forecasting accuracy between models. The test results are presented separately for each lag specification of the MIDAS models, namely $K = 4$, $K = 8$, and $K = 12$, to facilitate a clear comparison of model performance under different lag structures.

Note: A positive DM Statistic indicates that Model 1 provides more accurate forecasts than Model 2, whereas a negative DM Statistic indicates that Model 2 provides more accurate forecasts than Model 1. However, if the p-value is greater than 0.05, the difference is considered statistically insignificant.

Table 4.22 Diebold–Mariano Test Results for $K = 4$ (Train:Test = 70:30).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	4.553	0.000
MIDAS-AR-HMM-SOM	MIDAS	12.446	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	3.126	0.002
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	4.090	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	3.423	0.00

Table 4.23 Diebold–Mariano Test Results for $K = 8$ (Train:Test = 70:30).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	4.554	0.000
MIDAS-AR-HMM-SOM	MIDAS	12.450	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	2.581	0.011
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	3.747	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.557	0.012

Table 4.24 Diebold–Mariano Test Results for $K = 12$ (Train:Test = 70:30).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	4.549	0.00
MIDAS-AR-HMM-SOM	MIDAS	12.445	0.00
MIDAS-AR-HMM-SOM	MIDAS-AR	2.725	0.008
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	3.675	0.00
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.877	0.005

Table 4.25 Diebold–Mariano Test Results for $K = 4$ (Train:Test = 75:25).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	4.004	0.000
MIDAS-AR-HMM-SOM	MIDAS	12.491	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	3.115	0.003
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	3.845	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	3.389	0.001

Table 4.26 Diebold–Mariano Test Results for $K = 8$ (Train:Test = 75:25).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	4.004	0.000
MIDAS-AR-HMM-SOM	MIDAS	12.495	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	2.698	0.009
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	3.183	0.002
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.612	0.010

Table 4.27 Diebold–Mariano Test Results for $K = 12$ (Train:Test = 75:25).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	4.003	0.000
MIDAS-AR-HMM-SOM	MIDAS	12.490	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	2.701	0.009
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	3.442	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.888	0.005

Table 4.28 Diebold–Mariano Test Results for $K = 4$ (Train:Test = 80:20).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	3.353	0.001
MIDAS-AR-HMM-SOM	MIDAS	11.611	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	2.975	0.004
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	2.852	0.006
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.996	0.003

Table 4.29 Diebold–Mariano Test Results for $K = 8$ (Train:Test = 80:20).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	3.355	0.001
MIDAS-AR-HMM-SOM	MIDAS	11.614	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	2.660	0.010
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	2.700	0.009
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.385	0.020

Table 4.30 Diebold–Mariano Test Results for $K = 12$ (Train:Test = 80:20).

Model 1	Model 2	DM-Statistic	p-value
MIDAS-AR-HMM-SOM	ARIMA	3.353	0.001
MIDAS-AR-HMM-SOM	MIDAS	11.611	0.000
MIDAS-AR-HMM-SOM	MIDAS-AR	2.712	0.009
MIDAS-AR-HMM-SOM	MIDAS-AR-Threshold	2.734	0.008
MIDAS-AR-HMM-SOM	MIDAS-AR-HMM	2.568	0.013