

Artificial Intelligence in International Finance:**An Empirical Comparison of Traditional and Machine-Learning Models for EUR/USD Forecasting**Author: Elie Esper¹*Affiliation:*

1. *Business School, Holy Spirit University of Kaslik, P.O. Box 446, Jounieh, Lebanon*
2. *CIRAME Research Center, Holy Spirit University of Kaslik, P.O. Box 446, Jounieh, Lebanon*

Abstract:

This study investigates the use of artificial intelligence and machine learning in international-finance exchange-rate forecasting. The research compares a traditional ARIMA model with Random Forest and multilayer perceptron (MLP) neural-network models for forecasting the monthly EUR/USD exchange rate. The dataset contains 192 monthly observations from January 2010 to December 2025 obtained from the Federal Reserve Board's foreign-exchange series distributed through FRED. The sample is divided chronologically into a training period covering January 2010 to September 2022 and an out-of-sample test period covering October 2022 to December 2025. Machine-learning models use the previous twelve monthly exchange-rate observations as predictors, while ARIMA is selected using AIC from a small grid of candidate specifications. Forecast performance is evaluated using MAE, RMSE, MAPE, and directional accuracy. The results show that the Random Forest model achieved an RMSE of 0.0231 and MAPE of 1.76%, while the MLP achieved an RMSE of 0.0229 and MAPE of 1.75%. These errors were lower than those of the ARIMA benchmark, which produced an RMSE of 0.1156 and MAPE of 9.71%. The findings indicate that nonlinear machine-learning models can provide useful forecasting information for EUR/USD in the examined sample, while also highlighting the importance of out-of-sample validation and the limitations of relying on historical patterns alone.

Keywords:

Exchange-Rate Forecasting, EUR/USD, Artificial Intelligence, Machine Learning, ARIMA, Random Forest, Neural Networks, Time-Series Analysis

Introduction:

Exchange rates are a central variable in international finance because they affect the value of cross-border trade, foreign investments, multinational cash flows, and the measurement of financial risk. The EUR/USD exchange rate is particularly relevant because the euro and the U.S. dollar are two major international currencies and their exchange-rate relationship is continuously influenced by monetary conditions, economic news, expectations, risk sentiment, and global financial conditions. Forecasting this exchange rate is therefore relevant to firms that invoice internationally, investors who hold foreign assets, and financial institutions that manage currency exposure.

At the same time, exchange-rate forecasting has a long history of difficult out-of-sample performance. Meese and Rogoff (1983) famously reported that a random-walk benchmark was difficult for conventional estimated exchange-rate models to beat at one- to twelve-month horizons. This result created an important benchmark problem for international finance: a model can fit historical data and still fail to produce useful forecasts outside the estimation sample.

Recent developments in artificial intelligence and machine learning create a new research opportunity. Instead of relying only on a pre-specified linear relationship, machine-learning models can use lagged observations and nonlinear transformations to approximate complex patterns. Recent exchange-rate research has applied temporal machine-learning methods to EUR/USD and has also compared traditional time-series models with deep-learning approaches (Alexandria Engineering Journal, 2026; Scientific African, 2025).

Despite this development, the practical question for a finance student is not whether artificial intelligence is fashionable, but whether it improves measurable forecasting accuracy relative to established benchmarks. This study therefore evaluates EUR/USD monthly forecasting performance using a traditional ARIMA model, a Random Forest model, and a multilayer perceptron (MLP) neural network, while also reporting a random-walk benchmark.

The main research question is: How does the out-of-sample forecasting accuracy of machine-learning models compare with a traditional ARIMA model and a random-walk benchmark for the EUR/USD exchange rate?

The objectives are to: (1) construct a reproducible monthly EUR/USD forecasting dataset; (2) estimate a traditional ARIMA benchmark; (3) estimate Random Forest and MLP neural-network models using lagged exchange-rate information; (4) compare the models using MAE, RMSE, MAPE, and directional accuracy; and (5) discuss the implications and limitations of AI-assisted forecasting for international financial decision-making.

Literature Review:

Exchange-rate forecasting is one of the classic empirical problems in international finance. Traditional exchange-rate models are often motivated by macroeconomic relationships involving prices, interest rates, money, income, and external balances. However, the forecasting evidence is more complicated than the theoretical relationships suggest. Meese and Rogoff (1983) compared several structural and time-series models and found that a random-walk model performed at least as well as the estimated alternatives over several short and medium forecasting horizons. This finding does not mean that exchange rates are economically meaningless; rather, it highlights how difficult it is to translate economic relationships into stable out-of-sample forecasts.

Other international-finance research demonstrates that exchange rates also contain time-varying risk and predictable patterns in returns and volatility. Menkhoff et al. (2012), for example, document a relationship between global foreign-exchange volatility risk and carry-trade returns. Such evidence reinforces the idea that the foreign-exchange market is affected by changing risk conditions rather than by a single stable linear mechanism.

Machine learning provides a different modeling philosophy. Random Forests aggregate many decision trees and can represent nonlinear interactions without requiring the researcher to specify a particular functional form (Breiman, 2001). Neural networks similarly allow nonlinear mappings between inputs and outputs. The MLP used in this study is a feed-forward neural network, while the broader deep-learning literature includes recurrent architectures such as LSTM networks designed to model sequential information (Hochreiter & Schmidhuber, 1997).

Recent literature has increasingly applied these tools to exchange-rate prediction. A 2025 study in Scientific African compared ARIMA, SARIMA, LSTM, RNN, GRU, and a SARIMA-LSTM hybrid and reported that the hybrid framework produced strong forecasting performance for several exchange-rate series. A 2026 study in the Alexandria Engineering Journal focused specifically on EUR/USD and used lagged factors with Random Forest, Support Vector Regression, and k-nearest-neighbor models, emphasizing the importance of temporal feature engineering and variable selection. Focusing on AI, literature focus on the enhancement of economic quad prediction capabilities (Nasser et al., 2025), increase in firm performance (Subrahmanyam et al., 2024) using blockchain technology, especially SMEs (Khalife et al., 2024) by allowing smarter decision making (Subrahmanyam & Beainy, 2027).

The literature therefore provides a clear motivation for comparing conventional and AI-based approaches, but results can vary substantially with the currency, frequency, sample period, predictors, and evaluation design. The present study addresses this issue through a deliberately simple and reproducible design: it uses only historical EUR/USD information, separates the sample chronologically into training and test periods, and evaluates forecasts on observations that were not used to fit the models. The contribution is consequently empirical and educational rather than a claim that one algorithm is universally superior.

Methodology:

This study uses a quantitative, out-of-sample forecasting design. The dependent variable is the monthly EUR/USD exchange rate, defined as U.S. dollars per one euro. The source is the Federal Reserve Board's H.10 foreign-exchange series as distributed through the Federal Reserve Bank of St. Louis FRED database. FRED describes the series as the U.S. dollars-to-one-euro spot exchange rate and reports the monthly series as averages of available daily noon buying rates. The study period is January 2010 through December 2025, producing 192 monthly observations.

The sample is divided chronologically rather than randomly. The first 80% of observations, January 2010 through September 2022, form the training sample (153 observations). The remaining 20%, October 2022 through December 2025, form the out-of-sample test sample (39

observations). This design prevents future observations from being used to train the models and is more appropriate for forecasting than random cross-validation.

For the machine-learning models, twelve lagged monthly exchange-rate observations are used as predictors. Thus, the feature vector for month t contains EUR/USD at $t-1$ through $t-12$. The Random Forest model contains 500 trees, a maximum tree depth of 6, and a minimum of two observations per leaf. The MLP neural network contains two hidden layers with 32 and 16 neurons, standardized inputs, and a fixed random seed to support reproducibility. The models are trained on the historical training observations and generate one-step-ahead predictions for each month in the test period using information available through the preceding month.

For the traditional model, ARIMA is estimated using the training sample. Candidate specifications with p and q values from 0 to 2 and differencing d from 0 to 1 are compared using the Akaike Information Criterion (AIC); the lowest-AIC specification is then used to forecast the test period. The resulting selected specification for this sample is ARIMA(1,1,2). A random-walk benchmark is also included because of its importance in the exchange-rate forecasting literature (Meese & Rogoff, 1983).

Forecast accuracy is assessed with four measures. Mean Absolute Error (MAE) measures the average absolute forecast error. Root Mean Squared Error (RMSE) gives greater weight to larger errors. Mean Absolute Percentage Error (MAPE) expresses the absolute error as a percentage of the observed value. Directional accuracy measures the percentage of test observations for which the model correctly predicts whether the exchange rate moves up or down relative to the previous observation. No claim of causal inference is made: the analysis evaluates predictive performance, not the causal determinants of exchange rates.

Results:

The descriptive statistics show that the 192 monthly EUR/USD observations have a mean of 1.1864, a standard deviation of 0.1119, a minimum of 0.9853, and a maximum of 1.4460. The lowest observation in the sample is associated with the pronounced euro weakness visible in 2022, while the series subsequently recovered during 2023–2025.

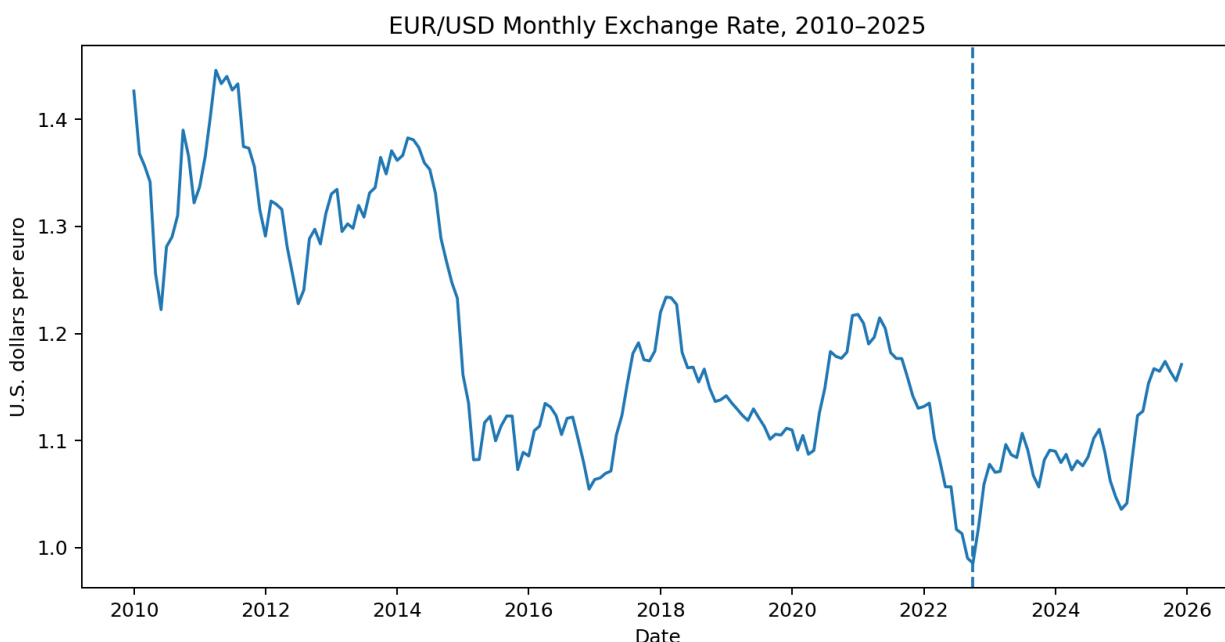


Figure 1. Monthly EUR/USD exchange rate, 2010–2025. The vertical marker separates the training and test samples.

Table 1. Descriptive statistics

Statistic	EUR/USD
Observations	192
Mean	1.1864
Std. Dev.	0.1119
Minimum	0.9853
25th Percentile	1.1007
Median	1.1541
75th Percentile	1.2886
Maximum	1.4460

The out-of-sample comparison is reported in Table 2. The Random Forest and MLP models produced substantially lower absolute forecast errors than both the random-walk and ARIMA benchmarks over the selected test period. The Random Forest recorded an MAE of 0.0190, RMSE of 0.0231, and MAPE of 1.76%. The MLP recorded an MAE of 0.0190, RMSE of 0.0229, and MAPE of 1.75%.

The ARIMA(1, 1, 2) model produced an MAE of 0.1075, RMSE of 0.1156, and MAPE of 9.71%. The random-walk benchmark produced an MAE of 0.1022, RMSE of 0.1105, and MAPE of 9.23%.

Table 2. Out-of-sample forecasting performance

Model	MAE	RMSE	MAPE (%)	Directional Accuracy (%)
Random Walk	0.1022	0.1105	9.23	46.15

ARIMA(1, 1, 2)	0.1075	0.1156	9.71	46.15
Random Forest	0.0190	0.0231	1.76	48.72
MLP Neural Network	0.0190	0.0229	1.75	51.28

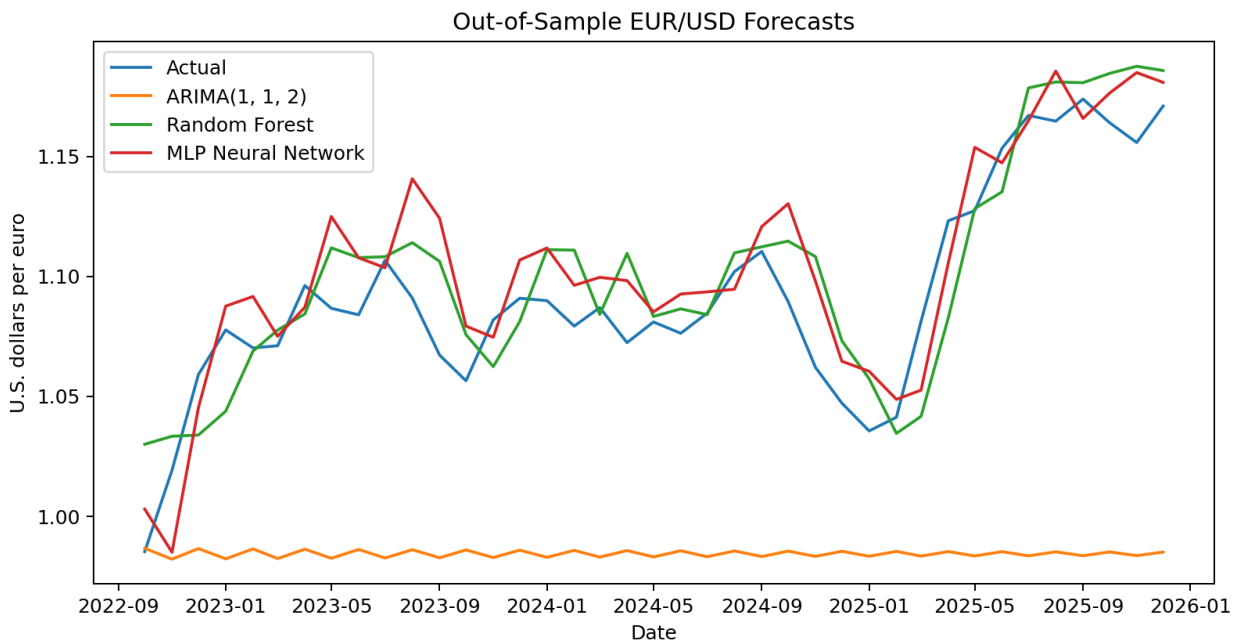


Figure 2. Actual and out-of-sample model forecasts, October 2022–December 2025.

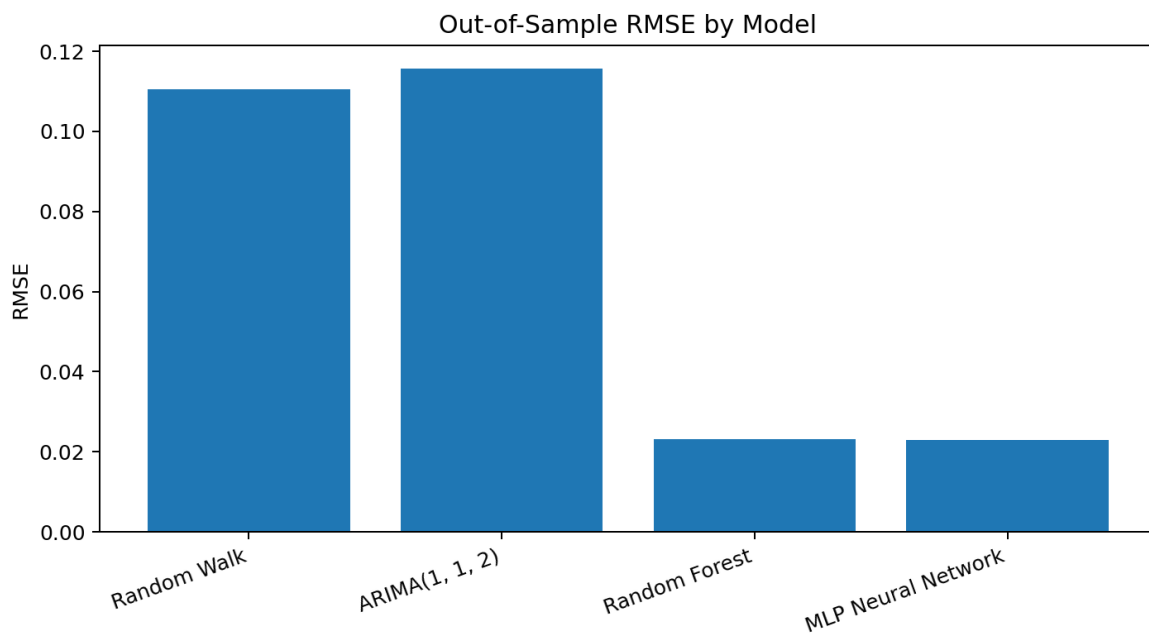


Figure 3. RMSE comparison across forecasting models.

Table 3 provides the first ten observations in the test period as a transparent check of the forecast construction. The full 39-month forecast file is supplied separately in the accompanying Excel workbook.

Table 3. First ten out-of-sample forecasts

Date	Actual	ARIMA	Random Forest	MLP
2022-10	0.9853	0.9867	1.0300	1.0030
2022-11	1.0192	0.9822	1.0334	0.9851
2022-12	1.0591	0.9866	1.0339	1.0453
2023-01	1.0777	0.9823	1.0438	1.0876
2023-02	1.0702	0.9864	1.0688	1.0916
2023-03	1.0711	0.9824	1.0776	1.0750
2023-04	1.0962	0.9863	1.0844	1.0872
2023-05	1.0867	0.9825	1.1119	1.1250
2023-06	1.0840	0.9862	1.1078	1.1078
2023-07	1.1067	0.9826	1.1082	1.1036

Discussion and Conclusion:

This study examined whether AI-oriented machine-learning models can improve the out-of-sample forecasting of the EUR/USD exchange rate relative to traditional and benchmark approaches. Using 192 monthly observations from January 2010 to December 2025, the analysis compared ARIMA, Random Forest, MLP neural network, and a random-walk benchmark on a chronologically separated test sample.

The empirical results show materially different forecast errors across the models. In the selected test period, the Random Forest and MLP models generated substantially lower MAE, RMSE, and MAPE values than the ARIMA and random-walk benchmarks. These results provide evidence that lag-based nonlinear models can capture useful predictive patterns in this particular sample. However, the evidence should be interpreted as sample-specific forecasting evidence rather than proof that artificial intelligence will consistently outperform conventional models in every currency market or period.

For international-finance practice, the findings suggest that AI-based forecasts may be useful as an additional information source for currency-risk monitoring, scenario analysis, and budgeting. At the same time, forecast models should not replace risk controls, hedging policies, sensitivity analysis, or professional judgment. Exchange-rate markets are exposed to structural breaks, policy announcements, geopolitical shocks, and changing relationships that may cause historical patterns to weaken.

The study has several limitations. First, it uses only the historical EUR/USD exchange rate and therefore does not include interest-rate differentials, inflation, oil prices, volatility indices, monetary-policy variables, or geopolitical-risk measures. Second, the sample is monthly, so the results cannot be generalized automatically to daily or intraday forecasting. Third, the test period is one historical period and the analysis does not establish statistical significance of forecast differences through a formal Diebold-Mariano test. Fourth, the machine-learning models involve

hyperparameter choices that could be optimized more extensively. Fifth is not focusing on the increasingly important subject of sustainability in finance (Kaspard et al., 2025 ; Kaspard et al., 2026), lastly additional focus on cryptocurrencies would have been useful because Beainy and Kamel (2026) argue that in the modern international financial system, cryptocurrencies construct pillars, especially in developed countries.

Future research can extend the framework by incorporating macroeconomic and financial variables, comparing additional currencies, using rolling-window or expanding-window evaluation, applying LSTM or Transformer architectures, and testing whether forecasting gains survive transaction costs and economic-value measures. Such extensions would provide a stronger basis for evaluating the practical role of AI in international financial forecasting.

References

- Alexandria Engineering Journal. (2026). Transitioning from static to temporal machine learning models: Forecasting the EUR/USD exchange rate utilizing lagged factors and feature significance. 135, 168–178. <https://doi.org/10.1016/j.aej.2025.12.051>
- Beainy, R., & Kamel, C. (2026). Gold and Bitcoin: A Quantitative Comparative Analysis. *The Banking and Finance Review*, 17(1).
- Board of Governors of the Federal Reserve System. (2026). U.S. dollars to euro spot exchange rate [DEXUSEU]. Federal Reserve Bank of St. Louis, FRED. <https://fred.stlouisfed.org/series/DEXUSEU>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Kaspard, J. A., Khalil, F. C., & Kamel, C. A. (2025). Factors contributing to non-financial performance of companies. In *Sustainable Economic Development: Utilizing Non-Financial Performance, Sustainability Reporting, and Corporate Governance* (pp. 413-442). Cham: Springer Nature Switzerland.
- Kaspard, J., Kamel, C., Khalil, F., & Beainy, R. (2026). Financial Performance, Risk, and Market Integration of Sustainability-Oriented Equity Indices: Implications for the Sustainability Transition (2010–2025). *Risks*, 14(5), 99.
- Khalife, D., Subrahmanyam, S., & Farah, A. (2024). A Sustainable Circular Business Model to Improve the Performance of Small and Medium-sized Enterprises Using Blockchain Technology. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, 15(2), 240-250.
- Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, 14(1–2), 3–24. [https://doi.org/10.1016/0022-1996\(83\)90017-X](https://doi.org/10.1016/0022-1996(83)90017-X)

- Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2012). Carry trades and global foreign exchange volatility. *The Journal of Finance*, 67(2), 681–718.
<https://doi.org/10.1111/j.1540-6261.2012.01728.x>
- Nasser, A., Yammine, J., Harouny, J., & Khalife, D. (2025). Predicting Economic Quads Through Asset Returns Using Ensemble Machine Learning. *IEEE Access*.
- Scientific African. (2025). A hybrid framework of deep learning and traditional time series models for exchange rate prediction. 29, e02818.
<https://doi.org/10.1016/j.sciaf.2025.e02818>
- Subrahmanyam, S., & Beainy, R. H. (2027). Smart Decision Support for Uncertainty and Disruptions in Industry 4.0 Operations. In *Smart Operational Research Solutions in the Era of Industry 4.0 and Sustainability* (pp. 167-198). IGI Global Scientific Publishing.
- Subrahmanyam, S., Aishwaryalaxmi, N. S., Khalife, D., Shaikh, I. A. K., Faldu, R., & Asthana, N. (2024, April). Impact of knowledge management and big data analytics capabilities on firm performance. In *2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)* (pp. 1-5). IEEE.