fractal and fractional

MDPI

Article

Multifractal-State-Augmented Machine Learning for One-Trading-Day-Ahead USD/KRW Absolute-Return Forecasting

Dongxue Wang [1,2] and Yugang He [3,*]

[1] Research Center of China-Pakistan Economic Corridor, Kashi University, Kashi 844000, China
[2] College of Economics and Management, Kashi University, Kashi 844000, China
[3] Department of Chinese Trade and Commerce, Sejong University, Seoul 05006, Republic of Korea
* Correspondence: 1293647581@jbnu.ac.kr

Abstract

Daily exchange-rate returns show weak average persistence, yet their multiscale organization may inform future movement magnitude. This study evaluates whether time-varying multifractal states improve one-trading-day-ahead forecasts of USD/KRW absolute returns. The raw dataset contains 6520 Federal Reserve H.10 observations from 1999 to 2024. After initialization of the 252-day rolling MFDFA window, 6268 feature-aligned observations remain for model estimation and evaluation. Seven scaling features augment Extra Trees, and forecasts are routed through pre-test terciles of singularity-spectrum width. Shuffled, phase-randomized, and IAAFT surrogates, matched-feature ablations, temporal-shift placebos, specification variants, and expanding-window thresholds assess robustness. In the locked 2020–2024 test of 1249 observations, the routed model yields an RMSE of 0.367 and an MAE of 0.271, improving on market-only Extra Trees by 1.627% and 2.245%, respectively. Temporal displacement produces upper-tail p-values of 0.020 for RMSE and 0.010 for MAE. The model's R^2 is 0.045 and its forecast–realization correlation is 0.218. Validation-calibrated 95% and 99% bands reduce interval scores by 2.160% and 5.730%, although coverage remains below nominal. External-currency gains are mixed and do not survive multiplicity adjustment. The results indicate modest, time-aligned, USD/KRW-specific predictive information; they do not establish universal transferability, causal effects, directional prediction, or trading profitability.

Keywords: multifractal detrended fluctuation analysis; singularity-spectrum width; USD/KRW; absolute-return forecasting; machine learning; Extra Trees; regime adaptation; exchange-rate risk

Check for updates

Academic Editor: Sun-Yong Choi

Received: 23 July 2026
Revised: 5 September 2026
Accepted: 8 September 2026
Published: 15 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Foreign exchange fluctuations affect inflation, trade competitiveness, external financing, and financial stability. Their effects are especially important in economies that combine deep global integration with substantial exposure to dollar funding. Korea provides a useful setting. The won responds to global risk sentiment, cross-border capital flows, changes in interest-rate expectations, and shifts in export conditions. These forces do not operate at one fixed horizon. Some are absorbed gradually, while others produce sharp daily adjustments. Reliable forecasts of near-term exchange-rate risk are therefore valuable to banks, firms, asset managers, and public authorities. A central debate concerns the meaning of weak average return persistence. Under the conventional efficiency view, a

Hurst exponent near 0.5 suggests that past returns contain little useful information. The fractal view offers a different interpretation. It allows small and large fluctuations to follow different scaling laws, even when average second-order persistence appears random. Grobys [1] developed a multifractal model of realized asset variances and showed how power-law behavior and long dependence can arise from multiplicative cascades across time scales. This evidence suggests that one Hurst exponent may conceal economically relevant variation in market organization.

Evidence of multifractality, however, is not sufficient by itself. A wide spectrum may reflect heavy tails, nonlinear temporal dependence, finite samples, or some combination of these forces. It may also change when market conditions shift. Caporin et al. [2] show that exchange-rate volatility contains both global and intermittent factors. Their effects vary across trading intervals and changed unevenly during the COVID-19 period. This finding reinforces a broader point. Currency volatility is generated by overlapping mechanisms that become active at different times. A credible analysis should therefore identify the sources of multifractality and allow the estimated scaling state to evolve. The forecasting literature faces a related disagreement. Flexible algorithms can approximate nonlinear interactions without imposing a fixed functional form. Yet additional complexity does not guarantee genuine predictive improvement. Exchange-rate predictors often change across policy regimes, and highly persistent price levels can create misleading fit. Yu et al. [3] show that the important predictors of the RMB differ across reform periods. They also find that predictors selected by several machine-learning models improve forecasts more than those selected by one method. Their results favor state-aware information design, but they also warn against assuming that one predictor set remains stable. Recent deep-learning studies provide further evidence of nonlinear predictability. Arslan et al. [4] compare support vector regression, recurrent neural networks, and long short-term memory models for intraday exchange rates. Their results favor the long short-term memory specification across several currencies and sampling frequencies. Such findings demonstrate the value of flexible temporal learning. They do not resolve whether the improvement comes from economically meaningful state variables or from a larger statistical architecture. This distinction matters for interpretation, robustness, and risk management. Dollar-related information shocks add another layer of instability. Lee and Wang [5] document pronounced return reversals after Federal Open Market Committee announcements. They connect these reversals to signed jump volatility, uncertainty resolution, and illiquid market conditions. The mechanism is highly relevant to dollar-linked currencies. For USD/KRW, monetary news can interact with regional trading hours, foreign portfolio flows, and changing dollar liquidity. A fixed forecasting relationship may therefore perform differently before, during, and after periods of global stress.

These strands leave an important gap. Fractal studies often establish scale dependence without testing whether it improves real-time forecasts. Machine-learning studies often emphasize forecast accuracy without identifying whether multiscale market structure supplies incremental information. Evidence for USD/KRW remains particularly limited. Existing work rarely combines rolling multifractal features, surrogate tests, matched feature ablation, chronological validation, and formal predictive-accuracy testing within one design. As a result, it remains unclear whether changing fractal states are merely descriptive or useful for daily currency-risk forecasting. The present study addresses this question using 6520 daily USD/KRW observations from January 1999 through December 2024. The target is the next-day absolute log return, which provides an observable measure of daily exchange-rate risk. A 252-day rolling multifractal detrended fluctuation analysis produces generalized Hurst exponents and three summaries of scaling heterogeneity. The generalized-Hurst spread, ΔH, measures the difference between small- and large-

fluctuation persistence. The singularity-spectrum width, $\Delta\alpha$, measures the range of local scaling behavior. Spectrum asymmetry describes the shape of that range. Shuffled and phase-randomized surrogates distinguish distributional effects from temporal organization. The predictive experiment compares linear regression, random forest, histogram gradient boosting, and Extra Trees. Each algorithm receives the same lag structure. Matched specifications with and without multifractal features isolate their incremental value. The proposed multifractal regime-adaptive Extra Trees model uses ex ante terciles of $\Delta\alpha$ to route forecasts across low-, medium-, and high-complexity states. Model choices are made in the training and validation samples. The period from 2020 through 2024 remains untouched until final evaluation. This design reduces look-ahead bias and prevents crisis information from entering model selection.

This study contributes in five respects. First, it documents the multiscale structure of USD/KRW returns while separating heavy-tailed marginal effects from nonlinear temporal ordering. Second, it evaluates multifractal predictors through matched market-only and fractal-augmented learners. The algorithm, lag structure, estimation sample, and forecast dates are held constant. Third, it converts the singularity-spectrum width, $\Delta\alpha$, into an observable forecasting state without using test-period information to determine the state thresholds. Fourth, it evaluates the stability of the USD/KRW result through annual rolling-origin estimation, alternative MFDFA windows, higher-order detrending, grouped feature ablations, dependence-robust inference, and a model-universe predictive-ability test. To assess decision relevance beyond point accuracy, the same locked forecasts are converted into validation-quantile-calibrated 95% and 99% return bands. Coverage, mean width, a proper interval score, and conditional breach magnitude are used to evaluate calibration, sharpness, and tail severity. This extension preserves the original forecast target and does not use test-period observations for interval calibration. Fifth, it subjects the complete frozen forecasting protocol to a no-retuning transfer test across eight additional currencies. It also examines whether the larger 2022–2024 point gains survive ex ante volatility scaling, common volatility-state reweighting, and conditional loss-differential controls. The empirical claim is deliberately bounded. MFRA-ET produces the lowest RMSE and MAE in the locked USD/KRW evaluation, but the improvement is modest and sensitive to the MFDFA horizon. The cross-currency evidence is mixed and does not support universal transferability. The later-period point advantage also becomes statistically indistinguishable after the volatility-state composition is equalized. The subperiod analysis therefore evaluates forecast stability. It does not identify a causal effect of monetary tightening, market stress, or structural change. The model forecasts the magnitude of the next observed reference-rate change. It does not forecast return direction or intraday integrated variance. Its contribution concerns weak incremental information about daily USD/KRW displacement. It does not establish a general law across currencies, a causal effect of structural change, directional trading value, or transaction-cost-adjusted profitability.

Three theory-based hypotheses guide the analysis. First, USD/KRW returns exhibit scale-dependent heterogeneity that cannot be reproduced fully by shuffled or phase-randomized surrogates. Second, rolling multifractal variables contain incremental information about next-day return magnitude after conventional market predictors and the learning algorithm are held constant. Third, the singularity-spectrum state changes the conditional mapping from observable market variables to subsequent return magnitude. This hypothesis predicts state-dependent interactions and threshold effects. It does not require a monotonic increase in next-day absolute returns as $\Delta\alpha$ rises. Its direct empirical test is whether ex ante $\Delta\alpha$-based routing improves forecast loss beyond a global learner containing the same multifractal inputs. The expected gain may vary across loss functions and MFDFA measurement horizons. The remainder of the article is organized as follows.

Section 2 develops the theoretical and empirical motivation. Section 3 presents the data, MFDFA procedure, predictive models, and evaluation design. Section 4 reports the scaling evidence and out-of-sample results. Section 5 discusses the conclusions, implications, and limitations.

2. Literature Review and Theoretical Framework

2.1. Multiscale Risk, State Dependence, and Forecasting

Recent research views financial risk as a multiscale process rather than a sequence governed by one stable memory parameter. Grobys [6] develops a multifractal model of realized variances and links their power-law behavior to multiplicative cascades across time horizons. This result provides a theoretical basis for scale-dependent volatility. Fathi et al. [7] further show that power laws often describe realized foreign-exchange variance more accurately than the conventional lognormal model. Their evidence also points to greater tail risk than standard variance analysis suggests. These studies challenge Gaussian representations of currency risk. However, heavy tails alone can widen a multifractal spectrum. A credible analysis must therefore separate distributional effects from genuine temporal organization. Foreign-exchange evidence also indicates that fractal properties change over time. Ünal [8] finds marked variation in the multifractality and cross-correlation of EUR/TRY and USD/TRY rates. This variation implies that one full-sample Hurst exponent may conceal shifts in market efficiency. Sandubete et al. [9] reach a related conclusion from a chaos perspective. Nonlinear dependence becomes clearer at high frequencies but weakens when observations are aggregated. Apparent randomness at the daily horizon may therefore reflect information loss rather than the absence of structure. These findings are relevant to USD/KRW. Its daily returns can show weak average persistence even when smaller and larger movements follow different scaling relations. Structural changes create a further measurement problem. Saâdaoui [10] combined change-point detection with MFDFA and found that multifractality differs across market segments. This approach shows why estimates taken over a long sample may average across distinct states. Ali et al. [11] obtained similar evidence from high-frequency foreign-exchange and cryptocurrency jumps. Their rolling estimates reveal persistent but evolving multifractal behavior. Jumps also display stronger memory than a serially independent process would imply. Together, these studies support a rolling design. They also suggest that spectrum width may serve as a dynamic indicator of market complexity. Multifractal estimates can also depend on how common market information is treated. Lee et al. [12] showed that filtering aggregate market influence reveals asymmetric volatility and multifractality that remain hidden in raw cryptocurrency returns. Ke et al. [13] reached a complementary result by combining wavelet denoising with multifractal cross-correlation analysis. Their estimates reveal time-varying links among financial markets, technological innovation, and the real economy. These findings warn against treating all estimated complexity as intrinsic to one return series. Noise, shared factors, and temporal dependence may generate different portions of the observed spectrum. Surrogate tests are therefore necessary when multifractal variables are later used as forecasting inputs. Other complexity measures reinforce the importance of temporal scale. Mariani et al. [14] developed a multifractal conditional diffusion entropy method and showed that market states appear differently across time lags. Yang et al. [15] used detrended cross-correlation networks to document increasingly dense interactions among currencies linked to the U.S. dollar. A disturbance in one currency can therefore spread through a changing network rather than remaining isolated. The implication is important for USD/KRW. Its fractal state may reflect domestic trading conditions, global dollar forces, and regional currency adjustments. Rolling MFDFA offers

a concise way to summarize this evolving structure without assuming that one channel remains dominant.

The state dependence of currency volatility is also supported by recent econometric evidence. Morales-Zumaquero and Sosvilla-Rivero [16] identified global and intermittent factors in hourly exchange-rate volatility. Global factors operate throughout the day, while intermittent factors become active during particular trading sessions. Their effects also changed unevenly during the COVID-19 period. Such variation is especially relevant to the Korean won. USD/KRW responds to Asian trading, European risk reassessment, and U.S. monetary news. Peng et al. [17] further showed that climate-policy uncertainty has asymmetric predictive effects on exchange-rate volatility. Forecast performance is stronger in calm markets than in high-volatility states. Static predictor relationships are therefore unlikely to remain equally informative across regimes. Institutional and policy conditions provide additional sources of instability. Aftab and Mehmood [18] found that central-bank transparency reduces exchange-rate volatility in Asian emerging markets. They also showed that U.S. monetary-policy uncertainty raises currency risk. Wang and Wang [19] used high-frequency data and a directional-change volatility measure to show that exchange rates respond more strongly to expansionary FOMC announcements than to tightening or unchanged-policy announcements. This asymmetric response is directly relevant to a highly dollar-sensitive Asian currency such as the Korean won. Federal Reserve shocks can alter dollar funding, foreign portfolio flows, and interest-rate expectations in Korea. The resulting adjustment may involve an immediate jump, a delayed reversal, or persistent volatility. A regime-sensitive model is better suited to this sequence than a fixed linear relation. High-frequency evidence suggests that some extreme currency movements are conditionally predictable. Uzun et al. [20] found that machine-learning models can forecast the arrival and direction of intraday foreign-exchange jumps using liquidity measures and technical indicators. Their results remain meaningful during the COVID-19 period. Yet jump predictability does not imply that every crisis follows the same process. Sattarhoff and Lux [21] show that multifractal realized-volatility models perform differently in tranquil and turbulent samples. Their Markov-switching specification is particularly competitive at short horizons during unstable periods. Multifractal information may therefore be most useful when it identifies the prevailing risk state. Cross-market transmission adds another layer to won volatility. Ahmed et al. [22] documented time-varying multifractal cross-correlations between financial stress and major commodity indices. The financial-stress–Brent-oil pair exhibits the highest level of multifractal cross-correlation, while small fluctuations generally display greater persistence than large fluctuations. This channel matters for Korea because imported energy costs affect trade conditions and inflation expectations. Zhang and Huang [23] identified persistent multifractal cross-correlations among RMB exchange rates, including the KRW/CNY rate. Their surrogate analysis attributes the observed structure to both heavy tails and nonlinear dependence. Regional currencies are thus connected through scale-dependent mechanisms that can intensify under adverse shocks. Recent forecasting studies respond to these patterns by separating information across frequencies. Alexandridis et al. [24] combined wavelet components, macro-financial variables, machine learning, and forecast aggregation. Their results favor an amalgamation approach over reliance on one model or one frequency. Tang and Xie [25] used adaptive signal decomposition and dynamically optimized model weights for exchange-rate forecasting. Their framework adjusts the contribution of several neural models as market conditions change. Both studies show that multiscale information has predictive value. However, decomposition-based level forecasts may be difficult to interpret and can benefit from price persistence. Forecasting next-observation absolute returns avoids the mechanical persistence of exchange-rate levels and directly targets conditional displacement. Absolute

returns also have an established role in financial-risk measurement. Ding et al. [26] showed that absolute returns display stronger and more persistent dependence than signed returns. Granger and Ding [27] developed expected absolute return as an alternative measure of risk. Forsberg and Ghysels [28] further explained why absolute-return measures can contain useful information about future variation. This literature justifies absolute return as a daily conditional-magnitude target. It does not make it equivalent to realized volatility. Realized volatility instead uses intraday observations to estimate quadratic variation [29].

Machine learning offers a flexible way to model nonlinear volatility dynamics. Christensen et al. [30] showed that machine-learning methods can capture long persistence and interactions among daily, weekly, and monthly volatility measures. Their gains are strongest when additional predictors contain information not represented by standard lag structures. Flexibility alone, however, does not establish reliable predictability. Hewamalage et al. [31] demonstrated that random data partitions, unsuitable loss functions, and weak benchmarks can produce misleading forecasting claims. Financial time series are especially vulnerable because dependence and structural breaks violate common cross-validation assumptions. Chronological validation and an untouched test sample are therefore essential. Predictor design is as important as algorithm choice. Pabuccu and Barbu [32] showed that feature selection can improve financial forecasts while reducing a large and noisy input space. Their evidence favors structured selection over indiscriminate expansion of technical variables. Pirayesh Neghab et al. [33] combined macroeconomic fundamentals with interpretable machine learning and show that the contributions of exchange-rate drivers can vary over time in both magnitude and sign. In their CAD/USD application, crude oil is the leading driver, followed by gold and the TSX stock index. This evidence supports the use of a compact, economically grounded, and interpretable information set. Rolling multifractal measures meet this requirement because they summarize economically meaningful differences across fluctuation sizes rather than add numerous unrelated indicators. Deep-learning studies provide strong evidence of nonlinear currency dynamics. Mao et al. [34] showed that CNN–LSTM achieves the highest predictive accuracy among the exchange-rate forecasting architectures they evaluate. They also found that appropriately lagged news sentiment can further improve forecasting performance. Barthélémy et al. [35] showed that LSTM and gated recurrent unit models also improve early warnings of currency crises. These models can represent long temporal dependence and complex interactions. Their architecture, however, does not identify the source of predictability. A forecasting gain may arise from additional parameters, latent regime learning, or meaningful economic states. This distinction matters when forecasts are intended for risk monitoring rather than purely statistical competition. Hybrid feature engineering offers one response. Wang et al. [36] combine economic predictors, two-stage feature extraction, clustering, and bidirectional neural networks. Their clustering structure allows separate models to represent different exchange-rate observations. Zhou and Zhu [37] instead decompose USD/RMB rates into high-frequency, low-frequency, and trend components before applying a CNN-LSTM model. Both approaches improve predictive accuracy. Yet their elaborate preprocessing makes it difficult to isolate which information creates the gain. The use of exchange-rate levels can also obscure performance because levels are highly persistent. Matched feature ablation on return-based risk offers a clearer test. Recent comparisons also caution against assuming that the most complex model is always superior. Date and Maunthrooa [38] showed that an interpretable bivariate Kalman filter can outperform univariate models and the random walk across many currency pairs. Souto et al. [39] found that TSMixer performs well when large cross-sectional information sets are available, but its advantage weakens for smaller foreign-exchange datasets. Model performance therefore depends on the information environment. A long daily series for one currency pair may favor robust

ensembles over parameter-intensive neural architectures. State conditioning can improve volatility models when it is supported by informative covariates. Nguyen et al. [40] embedded exogenous variables within a recurrent volatility model and allowed abrupt changes in volatility levels. Mishra et al. [41] combined Transformer mechanisms with GARCH and LSTM models. Their hybrid specifications remain competitive during the pandemic. These results show that flexible models can adapt to changing conditions. They also raise a practical question: whether a transparent state variable can deliver similar adaptation with less architectural complexity. Rolling spectrum width is a suitable candidate because it is observable before each forecast and has a direct interpretation as scaling heterogeneity. Tree ensembles provide a useful middle ground. Jazayeri [42] reports that a stacked random-forest and support-vector model improves short-horizon currency forecasts across several pairs. Tree methods capture thresholds and interactions without requiring large neural networks. They are also compatible with lagged volatility, semivolatility, and multifractal inputs. Extra Trees adds strong split randomization, which can reduce unstable searches among correlated predictors. These properties make tree ensembles appropriate for testing whether MFDFA variables add information to conventional market features. The strongest comparison must hold the learner, lag structure, and evaluation sample constant.

A smaller body of literature directly integrates fractal structure with machine learning. Zhu and Diao [43] combineed rough-volatility models with gated recurrent networks for option hedging. Their results show that domain-specific roughness can improve learning in a non-Markovian environment. Chen et al. [44] introduced multifractal variables into boosted-tree models and reported substantial gains during major economic and geopolitical events. These studies establish an important principle. Fractal information becomes operational when it enters a flexible learner. However, existing evidence remains concentrated in derivatives, freight, energy, and equity markets rather than daily exchange-rate risk. Recent work also treats fractal measures as regime indicators. Lee and Cho [45] used the Hurst exponent to update a graph-based volatility model as market memory changes. Their adaptive specification outperforms alternatives that ignore regime variation. Cho and Lee [46] used asymmetric fractality to calibrate the uncertainty of machine-learning views in Black–Litterman portfolios. The resulting portfolios perform especially well in declining markets. These findings suggest that fractal measures may be more valuable as conditioning states than as isolated linear predictors. They also support separate treatment of calm and turbulent market environments. Interpretability remains central to the practical use of such models. Ji et al. [47] combined multiscale decomposition, CatBoost, and SHAP analysis to distinguish short-, medium-, and long-horizon drivers of the RMB exchange rate. Zhu and Li [48] show that fractal network factors have different pricing effects across frequency bands. Their surrogate and out-of-sample tests reduce the risk that the observed structure is a statistical artifact. These studies demonstrate that multiscale prediction should identify which features matter and when they matter. Grouped permutation importance provides a suitable test when lagged predictors are highly correlated. Liu et al. [49] provide further evidence that regime-aware multifractal models improve volatility forecasts across currency, commodity, and stock markets. Their bivariate Markov-switching multifractal model captures both volatility dependence and changing correlations. It is particularly competitive over longer horizons. This result confirms that multifractality and regime variation are related but distinct. A daily forecasting design must determine whether fractal variables add information, and whether routing by a fractal state creates a further gain. These questions require matched ablations, formal predictive-accuracy tests, and regime-specific evaluation.

2.2. Economic Mechanism and Testable Implications

The forecasting mechanism can be stated more precisely through a heterogeneous-horizon representation. At forecast origin t, let the centered aggregate return be written as $r_{t+1} - u_t = \sum_{h\in\mathcal{H}} w_{h,t}u_{h,t}$, $E_t(u_{h,t+1}) = 0$, where $u_{h,t+1}$ is the return component associated with horizon h, and $w_{h,t}$ is its time-varying contribution to the observed exchange-rate movement. If Σ_t denotes the conditional covariance matrix of these components, then $\sigma^2_{t+1|t} = \mathbf{w}_t^{'}\sum_t \mathbf{w}_t$. For a conditionally symmetric scale family, and given the negligible daily conditional mean, the expected return magnitude satisfies $m_{t+|t} = E_t(|r_{t+1}|) \approx k_t\sqrt{\mathbf{w}_t^{'}\sum_t \mathbf{w}_t}$, where k_t depends on the shape of the standardized innovation distribution. This expression provides the economic link between heterogeneous horizons and next-day risk. Dealers, commercial firms, asset managers, and macro-oriented investors adjust positions at different speeds. Their relative influence changes with liquidity, dollar funding, policy news, and portfolio rebalancing. Evidence that volatility components behave differently across temporal resolutions is central to the heterogeneous-market hypothesis [50]. Multifractal asset-pricing models formalize a related mechanism through persistent volatility components with heterogeneous durations [51–53]. MFDFA provides a reduced-form measure of this latent cross-scale configuration. Curvature in $\tau_t(q) = qH_t(q) - 1$ implies that small and large fluctuations do not follow one common scaling law. Positive q emphasizes large movements, whereas negative q emphasizes small movements. The spectrum width, $\Delta\alpha_t = \alpha_{max,t} - \alpha_{min,t}$, therefore measures the dispersion of local scaling strengths. A wider spectrum indicates a larger contrast between the scaling behavior of small and large fluctuations. It does not imply that the next return must be larger. Nor does it identify a causal shock. In the forecasting design, $\Delta\alpha_t$ is treated as a noisy and predetermined proxy for a latent market state, $Z_t = \phi(\Delta\alpha_t)$. The relevant conditional restriction is therefore $m_{t+|t} = g(\mathbf{X}_t, Z_t)$, where $\mathbf{X}_t$ contains conventional market predictors. An illustrative parametric representation is $\log\sigma^2_{t+1|t} = \alpha + \mathbf{b}^{'}\mathbf{X}_t + cZ_t + \mathbf{d}^{'}(\mathbf{X}_t \otimes Z_t) + \varepsilon_{t+1}$. The central prediction is $\mathbf{d} \neq \mathbf{0}$, rather than $c > 0$. Thus, the same lagged return or historical-volatility signal may have different implications in different multifractal states. This restriction motivates the routing architecture. It also explains why a monotonic ordering of unconditional return magnitudes across $\Delta\alpha_t$ terciles is not required. The one-day lead follows from state persistence. Cross-horizon weights, covariance structures, liquidity conditions, and cascade components do not reset at the daily close. A state estimated using information through day t can therefore remain relevant for r_{t+1}. The 252-day window is the measurement horizon used to estimate that state. It is not the forecast horizon. Overlap between adjacent windows makes the estimated state smoother, but overlap alone cannot establish incremental predictability. The latter must survive comparisons with matched market predictors, a global MFDFA-augmented learner, alternative windows, alternative scale grids, and temporally displaced forecasts.

The preceding mechanism yields three testable implications and two attribution restrictions. First, genuine temporal multifractality should not be reproduced fully by surrogates that preserve the marginal distribution or approximate linear spectral dependence. Second, correctly aligned MFDFA variables should reduce forecast loss when they are added to an otherwise identical market-information model. This implication is stronger than a generic feature-count argument. An equal-sized block of non-MFDFA distributional summaries should not reproduce the same gain. Third, if $\Delta\alpha_t$ is informative as a conditional state descriptor, routing by thresholds fixed before the test period should improve matched loss relative to a global learner with the same inputs. However, any state partition can create local response surfaces. We therefore compare $\Delta\alpha_t$ with gates based on historical volatility, moving-average displacement, rolling kurtosis, and the principal component of the conventional market variables. A 2×2 design further separates the contribution

of MFDFA inputs from the contribution of $\Delta\alpha_t$-based routing. Finally, jointly delayed MFDFA blocks preserve feature dimension and internal dependence while breaking their alignment with the current market state. These restrictions make the competing explanations empirically falsifiable. They do not identify a causal economic mechanism, but they distinguish MFDFA-derived information from generic dimensional expansion and generic state partitioning.

3. Data and Methods

3.1. Robustness, Data-Snooping Adjustment, and Quantile-Calibrated Prediction-Band Evaluation

The raw sample begins on 4 January 1999 and ends on 31 December 2024. Missing releases and non-trading days are removed. No calendar interpolation is used. The resulting dataset contains 6520 observations. Because the rolling MFDFA measures require a complete 252-day initialization window, the first 252 observations serve only as the historical input for constructing the rolling features and do not enter the feature-aligned sample. Table 1 summarizes the variables, transformations, and forecasting roles used in the analysis. TThe resulting estimation and evaluation sample therefore contains 6268 observations (6520-252), which also serves as the common feature-aligned sample for the descriptive analysis. Daily log returns are measured in percentage points. Let C_t denote the USD/KRW rate on trading day t. The return is defined by Equation (1).

$$r_t = 100[\ln(C_t) - \ln(C_{t-1})]. \quad (1)$$

Table 1. Variables, transformations, and forecasting roles.

Variable	Symbol	Construction	Forecasting Role
Daily USD/KRW reference rate	C_t	Federal Reserve H.10 noon buying rate, quoted as KRW per U.S. dollar	Source series
Log return	r_t	100 × log difference of C_t	Market input
Absolute return	α_t	$\lvert r_t \rvert$	Input; α_{t+1} is target
Rolling historical volatility	$HV_{k,t}$	$\sqrt{(252/k)\sum_{j=0}^{k-1} r_{t-j}^2}$, for k = 5, 20, 60	Market input
Mean absolute return	$MA_{k,t}$	Mean $\lvert r \rvert$ over k = 5, 20, 60 days	Market input
Upside/downside historical variation	$HV^{+}_{20,t}$; $HV^{-}_{20,t}$	Annualized 20-day variation constructed separately from positive and negative daily returns	Market input
Generalized Hurst exponents	$H(q)$	$q \in \{-4, \ldots, 4\}$	Fractal input
Multifractal spread	ΔH_t	$H_t(-4) - H_t(4)$	Fractal input
Spectrum width	$\Delta\alpha_t$	$\alpha_{max} - \alpha_{min}$	Fractal input and state
Spectrum asymmetry	$A\alpha_t$	Normalized difference between right and left widths	Fractal input

Note: All predictors are observed at time t or earlier. Each market and multifractal variable enters the forecasting model with lags 0–4. The dependent variable is the next-observation absolute return, α_{t+1}. Rolling multifractal variables are estimated using only the 252 returns available through forecast origin t. The HV variables are constructed from daily returns and are therefore historical-volatility predictors. They are not realized-volatility measures based on intraday observations.

The forecasting target is the magnitude of the next observed change in the daily H.10 reference rate. Let $\alpha_t = |r_t|$ denote the observed absolute return at time t. The one-observation-ahead target is therefore $\alpha_{t+1} = |r_{t+1}|$. The associated estimand is the conditional expected magnitude $m_{t+1|t} = E(|r_{t+1}| \,|\, \mathcal{F}_t)$, where $\mathcal{F}_t$ contains only information available at the forecast origin. Under the conventional conditional-scale representation $r_{t+1} - u_t = \sigma_{t+1} z_{t+1}$, a stable standardized innovation distribution implies $E_t(|r_{t+1} - u_t|) = \sigma_{t+1} E_t(z_{t+1})$. This relation provides a conditional-scale interpretation for absolute return. It is especially relevant in the present sample because the mean daily return is 0.004%, compared with a standard deviation of 0.656%. The absolute transformation deliberately removes the sign. It therefore addresses the size of the next reference-rate change, not whether the won will appreciate or depreciate. Directional information is still available among the predictors through signed return lags and separate upside and downside historical-variation measures. The target also differs from realized volatility. One H.10 observation per day cannot recover intraday quadratic variation, temporary price jumps, microstructure noise, or within-day liquidity stress. Accordingly, α_{t+1} is described throughout the paper as an observable daily displacement measure. It is not described as an estimator of intraday integrated variance. Section 3.7 provides directional and alternative-target robustness tests. The training sample ends on 30 December 2016. The validation sample runs from 3 January 2017 to 31 December 2019. The untouched test sample runs from 2 January 2020 to 31 December 2024 and contains 1249 targets. Hyperparameters and feature decisions are determined before the test sample is evaluated. After selection, each model is refitted on the combined training and validation data. Standardization, where required, uses estimation-sample moments only. To clarify the information available at each forecast origin, the definitions, transformations, and forecasting roles of all variables are summarized in Table 1.

3.2. Multifractal Detrended Fluctuation Analysis

MFDFA is applied to returns because nonstationary price levels can produce spurious scaling. At each forecast origin t, the baseline calculation uses the trailing N = 252 returns available through that date. This one-trading-year window balances estimation precision against responsiveness to changing market conditions. A 252-day rolling multifractal window is also used by Zeng et al. [54] for daily financial-risk forecasting, providing a direct precedent for the annual measurement horizon. A 126-day window is more adaptive but provides fewer observations for negative-order moments. A 504-day window is more stable but can average across distinct market states. These alternatives are evaluated without changing the learner. The baseline scales are $s \in \{10, 15, 20, 30, 40, 60\}$. They are approximately logarithmically spaced and cover about two trading weeks to one quarter. Logarithmic spacing prevents the scaling regression from being dominated by densely clustered long scales. The lower bound reduces instability in local detrending, especially when negative q emphasizes small fluctuations. The upper bound satisfies $s_{max} < N/4$. At s = 60, the two-ended segmentation retains $2\lfloor 252/60 \rfloor = 8$ segments. Scales above 60 are therefore excluded from the rolling calculation. These scales are statistical aggregation levels and are not assigned mechanically to particular investor groups. The wider full-sample scale range is feasible because the complete series contains substantially more observations. Rolling estimates use the symmetric moment range $q \in \{-4, -3, \ldots, 4\}$ and first-order detrending. This parameterization is consistent with the rolling exchange-rate MFDFA design of Ünal [8], which also uses first-order detrending, a symmetric integer-q grid, and scales beginning at 10 observations. Negative moments emphasize small fluctuations, whereas positive moments assign greater weight to large movements. Restricting |q| to 4 limits the numerical influence of very small detrended variances and rare extreme

observations. Accordingly, the selected window, moment range, and scale grid follow established financial MFDFA practice but are adapted to the finite length of each rolling sample. They are not presented as a uniquely optimal configuration. The baseline window, scale grid, moment range, and detrending order were fixed before evaluation of the final test period. Section 3.5 examines windows of 126, 252, and 504 observations; polynomial orders $m \in \{1, 2, 3\}$; three neighboring scale grids; and symmetric moment ranges with $q_{max} \in \{2, 3, 4, 5\}$. These alternatives are diagnostic specifications. None is used to redefine the main model after observing its locked-test performance. The procedure then follows the five steps of Kantelhardt et al. [55]. First, the demeaned return profile is constructed within each rolling window of length N:

$$Y(k) = \sum_{i=1}^{k} [r_i - \check{r}], k = 1, \ldots, N. \quad (2)$$

Second, the profile is divided into non-overlapping segments of length s from both ends. This produces $2N_s$ segments and reduces the loss of observations at the boundary. A first-order polynomial y_v is fitted in each segment. The detrended variance is

$$F^2(v, s) = (\frac{1}{s}) \sum_{i=1}^{s} \{Y[(v-1)s + i] - y_v(i)\}^2. \quad (3)$$

Third, the q-order fluctuation function is computed. Positive q emphasizes large movements. Negative q emphasizes small movements.

$$F_{q(s)} = \{(\frac{1}{2N_s}) \sum_{v=1}^{2N_s} \left[F^2(v, s)\right]^{\frac{q}{2}}\}^{\frac{1}{q}}, q \neq 0. \quad (4)$$

For q = 0, the logarithmic average is used. Fourth, the slope of log $F_{q(s)}$ on log s yields the generalized Hurst exponent:

$$F_{0(s)} = \exp\{\frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln\left[F^2(v, s)\right]\}. \quad (5)$$

A constant H(q) indicates monofractality. Variation across q indicates multifractality. The mass exponent, singularity strength, and singularity spectrum are

$$\tau(q) = qH(q) - 1; \alpha(q) = \frac{d\tau(q)}{dq}; f(\alpha) = q\alpha - \tau(q); \Delta\alpha_t = \alpha_{max,t} - \alpha_{min,t} \quad (6)$$

The derivative of $\tau(q)$ is evaluated numerically using centered finite differences for interior moment orders and one-sided differences at the endpoints. The predictive state vector contains $H(-4)$, $H(-2)$, $H(2)$, $H(4)$, $\Delta H = H(-4) - H(4)$, $\Delta\alpha = \alpha_{max} - \alpha_{min}$, and spectrum asymmetry A_α. Let α_0 denote the singularity strength at the maximum of $f(\alpha)$, $w_L = \alpha_0 - \alpha_{min}$, and $w_R = \alpha_{max} - \alpha_0$. Spectrum asymmetry is defined as $A_\alpha = \frac{w_R - w_L}{w_R + w_L}$. Positive values indicate a longer right branch and greater scaling heterogeneity among small fluctuations, whereas negative values indicate a longer left branch and greater heterogeneity among large fluctuations. A value close to zero indicates an approximately symmetric spectrum. A larger ΔH or $\Delta\alpha$ indicates stronger heterogeneity across fluctuation scales. Full-sample surrogate diagnostics use the wider scale set $s \in \{16, 24, 32, 48, 64, 96, 128, 192, 256\}$. The observed statistics are evaluated against 999 shuffled surrogates, 999 phase-randomized surrogates, and 199 iterative amplitude-adjusted Fourier-transform surrogates. Shuffling preserves the empirical marginal distribution but removes temporal ordering. Phase randomization preserves the Fourier amplitudes and hence the linear power spectrum, while disrupting nonlinear phase organization. It does not preserve the empirical marginal

distribution exactly. IAAFT surrogates preserve the empirical marginal distribution and approximate the original power spectrum. They therefore provide a stricter joint control for heavy tails and linear dependence. The Monte Carlo reference distribution is discrete. With B = 999, the smallest attainable plus-one *p*-value is 1/1000 = 0.001. With B = 199, the corresponding minimum is 1/200 = 0.005. A reported value at either minimum therefore means that no surrogate statistic reached or exceeded the observed statistic, that is, b = 0. It does not imply that the underlying tail probability is zero or that it has been estimated without simulation uncertainty. For transparency, the empirical results report B and the exact exceedance counts for both width statistics. Upper-tail Monte Carlo *p*-values are calculated as (1 + b)/(1 + B), where b is the number of surrogate statistics at least as large as the observed statistic and B is the number of surrogates. These tests evaluate whether the estimated scaling widths are unusual under specific statistical nulls. They do not establish predictive or economic relevance, which is assessed separately through locked out-of-sample comparisons.

3.3. Predictive Models and Multifractal State Adaptation

The market-only information set contains 10 variables and 50 lagged predictors. The augmented information set adds seven rolling MFDFA variables and contains 85 predictors. Each variable enters at lags 0–4. Linear regression requires no hyperparameter tuning. Extra Trees [56] uses 600 trees and searches over $\max_{\text{features}} \in \{0.5, 0.7, 1.0\}$, and $\min_{\text{features}} \in \{4, 6, 10\}$. Random forest [57] uses 500 trees and searches over $\max_{\text{features}} \in \{0.5, 0.8\}$ and $\min_{\text{features,leaf}} \in \{4, 8, 12\}$. Histogram gradient boosting [58] uses 350 iterations and searches over learning rates {0.03, 0.05}, maximum leaf nodes {7, 15}, and L2 regularization values {1, 5}. Hyperparameters are selected using validation RMSE only. Each selected model is subsequently refitted on the combined training and validation sample. Negative forecasts are truncated at zero because the target is nonnegative. No future observation enters feature construction, and no test observation enters parameter selection, threshold estimation, or model tuning. The complete forecasting workflow, from daily USD/KRW returns and rolling MFDFA state construction to lagged feature formation and regime-adaptive Extra Trees estimation, is summarized in Figure 1.

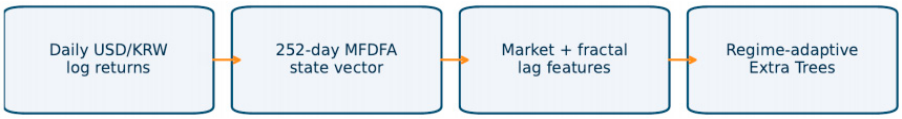

Figure 1. Fractal-state-augmented forecasting architecture. Note: one-trading-day-ahead absolute-return risk.

The proposed MFRA-ET model converts $\Delta\alpha$ into an observable routing state. The baseline partition uses the first and second development-sample terciles, which equal 0.243 and 0.430 before the final test. Quantile routing does not use the forecasting target and therefore avoids outcome-based threshold selection. Terciles were chosen to retain approximately balanced estimation samples while allowing a nonlinear conditional mapping through only three state-specific learners. The development sample contains 1671, 1671, and 1672 observations in the three states. The thresholds are not estimated structural breaks or economic risk boundaries. The state labels refer only to relative spectrum width. They do not imply a monotonic ordering of next-day absolute returns. Separate Extra Trees ensembles are estimated for the three complexity states. Let $S_t \in \{1, 2, 3\}$ denote the complexity state observed at forecast origin t, and $\mathbf{Z}_t$ denote the market-plus-fractal feature vector available through time t. The MFRA-ET forecast is

$$\breve{\alpha}_{t+1|t} = \sum_{k=1}^{3} \mathbf{1}(S_t = k) f_k(\mathbf{Z}_t; \theta_k). \quad (7)$$

Accordingly, $\breve{\alpha}_{t+1|t}$ denotes the forecast of the next-observation absolute return conditional on the information available through time t. MFRA-ET uses three ex ante complexity states defined by the first and second terciles of $\Delta\alpha_t$. The final development-sample thresholds are 0.243 and 0.430. State-specific models use 800 Extra Trees, $\max_{features} = 0.7$, and $\min_{samples,leaf} = 10$. A global random seed of 2024 is fixed before evaluation. The market-only and global fractal-augmented Extra Trees models use seed 2024. The low-, medium-, and high-complexity MFRA-ET ensembles use seeds 2024, 2025, and 2026, respectively. These values remain unchanged in the locked test, annual rolling-origin, feature-ablation, and MFDFA-sensitivity experiments. The routing state is observed at time t, and each forecast is generated only by the model associated with that state. The thresholds are estimated from the pre-2020 development sample and remain fixed in the principal 2020–2024 experiment. This restriction preserves the locked-test design and prevents subsequent structural changes from influencing the baseline partition. It does not assume that the distribution of $\Delta\alpha_t$ remains unchanged. Section 3.5 therefore introduces a pseudo-real-time expanding-window experiment in which both routing thresholds and all three local learners are re-estimated before each evaluation year. Additional partition and random-seed diagnostics examine sensitivity to routing granularity and algorithmic randomization. First, equal-frequency partitions with $K \in \{2, 3, 4, 5\}$ states are evaluated. For the validation experiment, all thresholds are estimated from the training sample. For the locked test, they are estimated from the complete pre-2020 development sample. Each routed model is compared with a global Extra Trees model containing the same market and MFDFA variables. The learner, hyperparameters, dates, and feature definitions remain fixed. These alternative partitions are reported as reviewer-requested sensitivity checks and do not alter the baseline tercile model. Second, the matched Extra Trees and random forest comparisons are repeated using 30 consecutive random seeds from 2000 through 2029. Hyperparameters and sample partitions remain unchanged. Seed repetitions assess algorithmic stability and are not treated as independent observations for statistical inference.

3.4. Forecast Evaluation and Statistical Inference

Forecasts are evaluated in the original percentage-return scale. RMSE penalizes large misses. MAE is less sensitive to extremes. MASE divides the model MAE by the MAE of the persistence benchmark $\breve{\alpha}^{persistence}_{t+1|t} = \alpha_t$. Values below one indicate improvement over persistence. R^2 is calculated against the test-sample mean. Forecast–realization correlation evaluates co-movement without rewarding a level bias. These two statistics answer different questions. Out-of-sample R^2 measures the reduction in squared forecast error relative to the unconditional test-sample mean. Correlation measures linear co-movement between forecasts and realizations. Neither statistic, considered alone, provides a formal test against accidental predictability. Section 3.6 therefore evaluates both their absolute levels and their increments over a matched market-only benchmark.

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\alpha_{t+1} - \breve{\alpha}_{t+1}\right)^2}; \\ MAE &= \frac{1}{N}\sum_{t=1}^{N}\left|\alpha_{t+1} - \breve{\alpha}_{t+1}\right|; \\ MASE &= \frac{\frac{1}{N}\sum_{t=1}^{N}\left|\alpha_{t+1} - \breve{\alpha}_{t+1}\right|}{\frac{1}{N}\sum_{t=1}^{N}\left|\alpha_{t+1} - \alpha_t\right|} = \frac{MAE_{model}}{MAE_{persistence}}; \\ R^2 &= 1 - \frac{\sum_{t=1}^{N}\left(\alpha_{t+1} - \breve{\alpha}_{t+1}\right)^2}{\sum_{t=1}^{N}\left(\alpha_{t+1} - \alpha_t\right)^2} \end{aligned} \tag{8}$$

Here, $\alpha_{test} = \frac{1}{N}\sum_{t=1}^{N} \alpha_{t+1}$ denotes the mean realized absolute return in the test sample. Equal predictive accuracy is evaluated using the Diebold–Mariano test (Diebold and Mariano [59]). Let $\check{\alpha}^{C}_{t+1|t}$ denote the forecast produced by a comparator model and let $\check{\alpha}^{M}_{t+1|t}$ denote the corresponding MFRA-ET forecast. The realized forecasting target is α_{t+1}. Under squared loss, the loss differential is defined as

$$d^{(2)}_{t+1} = (\alpha_{t+1} - \check{\alpha}^{C}_{t+1|t})^2 - (\alpha_{t+1} - \check{\alpha}^{M}_{t+1|t})^2. \tag{9}$$

Under absolute loss, the loss differential is

$$d^{(1)}_{t+1} = |\alpha_{t+1} - \check{\alpha}^{C}_{t+1|t}| - |\alpha_{t+1} - \check{\alpha}^{M}_{t+1|t}|. \tag{10}$$

In both definitions, the comparator loss is placed first and the MFRA-ET loss is placed second. Therefore, a positive average loss differential indicates that MFRA-ET produces a lower forecast loss. For loss function $\updownarrow \in \{1, 2\}$, the sample mean loss differential is

$$\check{d}^{(\updownarrow)} = \frac{1}{N}\sum_{t=1}^{N} d^{(\updownarrow)}_{t+1}. \tag{11}$$

The Diebold–Mariano statistic is calculated as

$$DM^{(\updownarrow)} = \frac{\sqrt{N}\check{d}^{(\updownarrow)}}{\sqrt{\check{\Omega}^{(\updownarrow)}}}. \tag{12}$$

where $\check{\Omega}^{(\updownarrow)}$ is the Newey–West estimate of the long-run variance of the loss differential. It is given by

$$\check{\Omega}^{(\updownarrow)} = \check{\gamma}^{(\updownarrow)}_0 + 2\sum_{j=1}^{L}(1 - \frac{j}{L+1})\check{\gamma}^{(\updownarrow)}_j. \tag{13}$$

where $\check{\gamma}^{(\updownarrow)}_0$ denotes the sample autocovariance of $d^{(\updownarrow)}_{t+1}$ at lag j, and L = 5. The five-lag Newey–West correction allows for short-run serial dependence in daily forecast losses. All tests are two-sided. A positive $DM^{(\updownarrow)}$ statistic favors MFRA-ET, whereas a negative statistic favors the comparator. The null hypothesis is

$$H_0 : E[d^{(\updownarrow)}_{t+1}] = 0, \tag{14}$$

which states that the comparator and MFRA-ET have equal expected predictive accuracy. The alternative hypothesis is

$$H_1 : E[d^{(\updownarrow)}_{t+1}] \neq 0. \tag{15}$$

Point-forecast performance is assessed using the loss measures defined above. Rejection of H_0 with a positive test statistic indicates that MFRA-ET achieves a statistically significant reduction in forecast loss. The ten pairwise comparisons—five comparators evaluated under squared and absolute loss—are treated as one inferential family. Let $p_{(1)} \leq \cdots \leq p_{(10)}$ denote the ordered raw *p*-values. For the hypothesis at rank k, the Holm-adjusted value is calculated as $p^{Holm}_{(k)} = \min\{1, \max_{j \leq k}[(10 - j + 1)p_{(j)}\}$. The Holm step-down procedure is used as the primary family-wise error-rate adjustment. Bonferroni-adjusted values, $p^{Bonf}_i = \min(1, 10p_i)$, are also reported as a conservative sensitivity check. Both procedures remain valid under arbitrary dependence among the comparisons. Statistical conclusions are assessed at $\alpha = 0.050$, and the text distinguishes nominal evidence from evidence that survives multiplicity adjustment.

3.5. Robustness, Data-Snooping Adjustment, and Risk-Band Evaluation

The principal evaluation uses one locked 2020–2024 test period. The annual rolling-origin exercise and the original specification checks were defined without altering the baseline forecast rule. Additional moment-range, partition, and random-seed diagnostics were introduced in response to the review. They are reported in full and are not used to select a new preferred model from the locked test. Together, these procedures examine whether the ranking survives changes in forecast origin, MFDFA construction, routing granularity, feature composition, and algorithmic randomization. First, a pseudo-real-time annual rolling-origin design is implemented. For each evaluation year $y \in \{2020, \ldots, 2024\}$, the information set ends on 31 December of year y-1. The lower and upper routing thresholds are recomputed as the first and second terciles of all $\Delta\alpha_t$ values available at that origin. Formally, $q_{k,y} = Q_{k/3}\{\Delta\alpha_t; t < y\}$ for $k \in \{1, 2\}$. The three state-specific Extra Trees models are then re-estimated under the updated partition. The algorithm, hyperparameters, feature definitions, lag structure, tree counts, and random seeds remain fixed at their pre-2020 values. No observation from the evaluation year enters threshold construction or model fitting. This procedure permits gradual threshold adaptation while preserving strict temporal ordering. Threshold drift is evaluated in two complementary ways. First, the annual values of $q_{1,y}$ and $q_{2,y}$ are compared with the fixed pre-2020 thresholds of 0.243 and 0.430. Second, the reassignment rate is calculated as the proportion of annual observations whose state under the expanding thresholds differs from their state under the fixed thresholds. State shares are also reported. They are not constrained to equal one third within an evaluation year. Consequently, the procedure allows persistent clustering in one region of the evolving predictor space. Second, MFDFA is repeated with rolling windows of 126, 252, and 504 observations. A separate pre-test audit evaluates these windows over the 2017–2019 validation period. The learner, feature definitions, and random seed remain fixed. The 2020–2024 observations do not enter this audit. Validation RMSE remains the declared selection criterion, while validation MAE is reported to show whether the ranking depends on the loss function. Third, scale sensitivity is examined using the short grid {8, 12, 16, 24, 32, 48}, the baseline grid {10, 15, 20, 30, 40, 60}, and the shifted grid {12, 18, 24, 36, 48, 60}. Each comparison uses a 252-day window, $q = -4, \ldots, 4$, first-order detrending, and the same forecasting protocol. The alternative grids are diagnostic specifications. They are not used to redefine the preferred model after observing the locked test results. Fourth, polynomial detrending orders $m = 1, 2, 3$ are examined for the 252-day baseline grid. Fifth, moment-range sensitivity is evaluated using $q_{max} \in \{2, 3, 4, 5\}$. Each grid contains every integer moment from $-q_{max}$ to q_{max}. For $q_{max} = 2$, the four generalized-Hurst inputs are $H(-2)$, $H(-1)$, $H(1)$, and $H(2)$. For $q_{max} \geq 3$, they are $H(-q_{max})$, $H(-2)$, $H(2)$, and $H(q_{max})$. The corresponding ΔH, $\Delta\alpha$, and A_α values are recomputed from each complete moment grid. Sixth, routing sensitivity is assessed with two, three, four, and five equal-frequency $\Delta\alpha$ states. Thresholds are estimated without the evaluation target and remain fixed within each forecast experiment. Seventh, mean next-day absolute returns are reported for the low-, medium-, and high-Δ_α states in both the development and test samples. This diagnostic tests whether the data support a monotonic level interpretation. It does not enter model selection. Eighth, five restricted feature sets are compared with market-only Extra Trees: the four generalized Hurst exponents; the dispersion group $\{\Delta H, \Delta_\alpha, A_\alpha\}$; the pair $\{\Delta H, \Delta_\alpha\}$; Δ_α alone; and all seven MFDFA variables.

Ninth, four attribution diagnostics distinguish MFDFA content from generic feature expansion and generic regime adaptation. First, a capacity-matched control block contains seven non-MFDFA summaries calculated from the same 252-day return window: skewness, excess kurtosis, the 5th and 95th percentiles, the interquartile range, the maximum absolute return, and the share of positive returns. Each variable enters at lags 0–4. The

block therefore contributes 35 columns, exactly matching the seven-variable MFDFA block and its lag structure. The Extra Trees specification, estimation dates, and target remain unchanged. Second, a 2×2 design crosses the input set—market variables alone or market variables plus MFDFA variables—with the forecasting architecture—one global learner or $\Delta\alpha_t$-based routing. In the routed market-input cell, $\Delta\alpha_t$ is used only to assign the observation to a state. It is excluded from the local learners' predictor matrices. This design identifies the incremental effect of MFDFA inputs at a fixed routing rule and the incremental effect of routing at a fixed input set. Third, the same market-plus-MFDFA predictor set is routed separately by development-sample terciles of five alternative state variables: 20-day historical volatility, 60-day historical volatility, 20-day moving-average displacement, 252-day excess kurtosis, and the first principal component of the ten contemporaneous market variables. All local models retain the baseline Extra Trees specification. Pairwise comparisons with $\Delta\alpha_t$ routing use two-sided Diebold–Mariano tests with five-lag Newey–West covariance estimates. Fourth, the 35 MFDFA columns are delayed jointly by 99 displacement values spanning 60–750 trading days. Joint displacement preserves the number of predictors, their cross-sectional organization, and their serial dependence, while severing their alignment with the current scaling state. Correctly aligned and displaced models use a common development sample of 4264 observations and the same 1249 locked test observations. For gain G, the upper-tail placebo probability is calculated as $p = [1 + \sum_{b=1}^{99G} I(G_b \geq G_0)]/100$, where G_0 is the gain from the aligned block. These reviewer-requested diagnostics do not alter the baseline model or its pre-test hyperparameters.

Tenth, grouped permutation importance is evaluated with a dependence-aware procedure. The diagnostic uses the fixed global Extra Trees model containing all market and MFDFA variables. The model is not re-estimated during permutation. For feature group g, its five lagged columns are permuted jointly. Test observations are divided into nonoverlapping blocks of 20 trading days, and the blocks are randomly reordered. Joint block permutation preserves the internal alignment of the five lags and retains short-run dependence within each block. Within each replication, the same block-order map is applied to all 17 feature groups. This paired construction prevents unequal permutation severity from affecting feature comparisons. For permutation replication r, the importance of group g is $I_{g,r} = RMSE_{g,r}^{pern} - RMSE^{base}$, where $RMSE^{base}$ is the locked-sample RMSE of the fitted global MFDFA-augmented Extra Trees model. Positive values indicate that permutation worsens forecast accuracy. Negative values indicate redundancy, substitution among correlated variables, or unstable marginal reliance. They do not imply a negative economic effect. The expected permuted squared error is estimated from 200 block permutations. Sampling uncertainty is then evaluated with 5000 paired circular moving-block replications of the 1249 locked test dates. The bootstrap block length is 20 trading days. For each bootstrap sample, the baseline and all permuted losses use identical date indices. The 2.5th and 97.5th percentiles form the 95% interval for each importance estimate. Rank stability is evaluated within every moving-block replication. For feature g, the analysis reports the proportion of replications in which it ranks first and the proportion in which it enters the top three. Pairwise rank claims concerning $\Delta\alpha$ are tested from the paired differences $D_{g,b} = I_{\Delta\alpha,b} - I_{g,b}$. The one-sided null is that $\Delta\alpha$ is no more important than feature g. The plus-one bootstrap probability is adjusted by the Holm procedure across the 16 comparisons. Permutation importance is interpreted as model reliance within the fitted global Extra Trees architecture. It is not a causal effect, a feature-selection criterion,

or evidence that one predictor is uniquely dominant. For loss measure L, the relative improvement of candidate model j over market-only Extra Trees is

$$G_{L,j} = 100\left(\frac{L_{ET} - L_j}{L_{ET}}\right). \tag{16}$$

Positive values indicate lower forecast loss. Because daily loss differentials are serially dependent, uncertainty around the rolling-origin gains is estimated with 5000 circular moving-block bootstrap replications. Block lengths of 5, 20, and 60 trading days are considered. The 20-day specification is primary and the other lengths assess sensitivity to the assumed dependence horizon. Pairwise Diebold–Mariano tests do not account for selection across several candidate feature sets. A studentized Superior Predictive Ability procedure based on Hansen [60] is therefore applied to the full prespecified feature universe. Market-only Extra Trees is the benchmark. Candidate losses include the five restricted feature specifications, global Extra Trees with all MFDFA variables, and MFRA-ET. The maximum studentized loss gain is evaluated with 5000 circular moving-block replications. Results are reported for absolute and squared loss. RMSE and MAE summarize point-forecast accuracy, but they do not evaluate conditional tail exposure. Transaction costs cannot be subtracted directly because the target is absolute-return magnitude rather than directional portfolio profit and loss. Decision relevance is therefore assessed through validation-quantile-calibrated two-sided return bands with nominal coverage levels of 95% and 99%. For model j, the calibration multiplier is obtained exclusively from the 2017–2019 validation sample:

$$c_{1-\eta,j} = Q^{val}_{1-\eta}\left(\frac{\alpha_{t+1}}{\check{\alpha}^{(j)}_{t+1|t}}\right). \eta \in \{0.05, 0.01\}. \tag{17}$$

The corresponding test-period band for the signed return is

$$[L^{(j)}_{t+1}, U^{(j)}_{t+1}] = [-c_{1-\eta,j}\check{\alpha}^{(j)}_{t+1|t}, c_{1-\eta,j}\check{\alpha}^{(j)}_{t+1|t}]. \tag{18}$$

Equation (17) estimates the empirical $(1-\eta)$ quantile of the realized-to-predicted absolute-move ratio using only the 2017–2019 validation sample. Equation (18) combines this fixed multiplier with each time-varying magnitude forecast. It therefore produces an ex ante two-sided prediction band for the signed daily return. No observation from the locked 2020–2024 test period enters the calibration step. This design prevents test-period information from determining either the band width or the nominal tail threshold. The prediction bands are evaluated using coverage, mean width, conditional mean breach magnitude, and the proper interval score. Coverage is the sample mean of $\mathbb{I}(L^{(j)}_{t+1} \le r_{t+1} \le U^{(j)}_{t+1})$. Mean width is the average value of $U^{(j)}_{t+1} - L^{(j)}_{t+1}$. For observations outside the band, breach magnitude is defined as $\max\{L^{(j)}_{t+1} - r_{t+1},\ r_{t+1} - U^{(j)}_{t+1},\ 0\}$. Its conditional mean measures the severity of tail exceedances rather than only their frequency. Following Gneiting and Raftery [61], the interval score is

$$IS^{(j)}_{\eta,t+1} = U^{(j)}_{t+1} - L^{(j)}_{t+1} + \frac{2}{\eta}(L^{(j)}_{t+1} - r_{t+1})\mathbf{1}(r_{t+1} < L^{(j)}_{t+1}) + \frac{2}{\eta}(r_{t+1} - U^{(j)}_{t+1})\mathbf{1}(r_{t+1} > U^{(j)}_{t+1}). \tag{19}$$

A lower interval score indicates a better joint balance between sharpness and miscoverage. The band endpoints are quantile-calibrated daily reference-rate thresholds. They are not separately estimated asymmetric quantile-regression forecasts. Mean breach magnitude is a tail-severity diagnostic rather than portfolio Expected Shortfall. Moreover, the bands should not be interpreted as fully calibrated portfolio VaR because the model contains

neither a position-specific profit-and-loss mapping nor an exposure rule. The exercise evaluates supplementary risk-monitoring value, not trading profitability.

3.6. Diagnostic Design for Incremental Predictability and Temporal Alignment

A low out-of-sample R^2 can arise for two distinct reasons. The forecastable component may be small, or the observed fit may result from accidental temporal alignment. We distinguish these possibilities explicitly. Let $\mathcal{F}_t$ denote the information available when the forecast is formed. The realized absolute return can be written as

$$\alpha_{t+1} = m_t + u_{t+1}, m_t = E(\alpha_{t+1}|\mathcal{F}_t),\ E(u_{t+1}|\mathcal{F}_t) = 0. \tag{20}$$

Equation (20) does not require m_t to account for most daily variation. Surprise policy decisions, geopolitical events, liquidity shocks, and other information arriving after time u_{t+1}. Multifractal states are therefore evaluated as predictors of a potentially small conditional component. They are not treated as sufficient statistics for the next day's realized risk. The first diagnostic compares MFRA-ET with market-only Extra Trees. This benchmark holds the learner, market variables, lag structure, estimation sample, and forecast dates constant. The incremental statistics are

$$\Delta R^2 = R^2_{MFRA} - R^2_{ET}, \Delta\rho = \rho(\alpha, \breve{\alpha}^{MFRA}) - \rho(\alpha, \breve{\alpha}^{ET}). \tag{21}$$

Sampling uncertainty is estimated with 10,000 paired circular moving-block bootstrap replications. In each replication, the same date indices are applied to the realizations and both sets of forecasts. This preserves the matched forecast comparison. The primary block length is 20 trading days. Blocks of 5 and 60 days assess sensitivity to shorter and longer dependence horizons. The second diagnostic is a temporal-alignment placebo. The fixed MFRA-ET forecast sequence is circularly shifted relative to the realized series. Shifts within 59 trading days of the correct alignment are excluded to reduce contamination from short-run volatility dependence. The placebo distribution therefore uses all admissible shifts $s = 60, \ldots, N - 60$. A one-sided empirical *p*-value is calculated as the proportion of shifted statistics at least as large as the correctly aligned statistic, with the standard finite-sample adjustment. This procedure preserves the forecast distribution and its serial dependence while removing its calendar-time alignment with the target. Third, we estimate a Mincer–Zarnowitz slope regression [62]:

$$\alpha_{t+1} = \alpha + b\breve{\alpha}^{MFRA}_{t+1|t} + e_{t+1}. \tag{22}$$

The null hypothesis H_0: $b = 0$ tests whether the forecast is unrelated to the realization. We use this slope test as an alignment diagnostic. We do not interpret it as a joint test of perfect calibration. Finally, incremental information is examined through the forecast-encompassing regression

$$\alpha_{t+1} = \alpha + b\breve{\alpha}^{ET}_{t+1|t} + c(\breve{\alpha}^{MFRA}_{t+1|t} - \breve{\alpha}^{ET}_{t+1|t}) + e_{t+1}. \tag{23}$$

The null hypothesis H_0: $c = 0$ states that the MFRA-ET increment contains no information beyond the market-only forecast. Equations (22) and (23) use heteroskedasticity- and autocorrelation-consistent standard errors with five lags. We also recompute R^2 and correlation from the pooled annual rolling-origin forecasts. All diagnostics use forecasts generated by the previously fixed models. No test-period result is used to select features, hyperparameters, or routing thresholds.

3.7. Target-Definition and Directional Robustness

The target analysis addresses two distinct concerns. The first is whether pooling positive and negative returns conceals directional heterogeneity. The second is whether the forecast ranking depends on the absolute-value transformation. For the first test, the locked forecasts are evaluated separately over the two ex post sign sets

$$\mathcal{D}^{+} = \{t : r_{t+1} > 0\}, \mathcal{D}^{-} = \{t : r_{t+1} < 0\}. \tag{24}$$

Because USD/KRW is quoted as Korean won per U.S. dollar, $\mathcal{D}^{+}$ represents won-depreciation days and $\mathcal{D}^{-}$ represents won-appreciation days. The split is used only for forecast evaluation. It is not available when the forecast is formed and does not define a directional trading rule. Thirteen zero-return observations are excluded from this diagnostic. For direction $d \in \{+, -\}$, the loss gain over market-only Extra Trees is

$$G_{L}^{(d)} = 100\left[\frac{L_{ET}^{(d)} - L_{MFRA}^{(d)}}{L_{ET}^{(d)}}\right], L \in \{RMSE, MAE\}. \tag{25}$$

Uncertainty is estimated with 10,000 paired circular moving-block bootstrap replications. The primary block length is 20 trading days. Blocks of 5 and 60 days assess sensitivity to the dependence horizon. Dates are resampled before the sign subsets are reconstructed. This procedure preserves the paired comparison between the target and the two forecasts. The second test replaces the absolute-return target with the one-day squared return:

$$\beta_{t+1} = r_{t+1}^{2}. \tag{26}$$

The market-only Extra Trees and MFRA-ET models are then refitted to β_{t+1}. The predictors, sample partitions, tree architecture, hyperparameters, routing rule, and random seeds remain fixed. This test determines whether the incremental result is specific to the absolute-value transformation. The squared daily return remains a noisy daily magnitude measure. It is not interpreted as high-frequency realized volatility. For the alternative target, the relative loss gain is

$$G_{L}^{(2)} = 100\left[\frac{L_{ET}^{(2)} - L_{MFRA}^{(2)}}{L_{ET}^{(2)}}\right]. \tag{27}$$

The same paired moving-block procedure is used. No result from either robustness test is used to alter the primary forecasting specification.

3.8. Cross-Currency Transferability and Volatility-Adjusted Subperiod Inference

External transferability is evaluated using eight additional Federal Reserve H.10 exchange-rate series: USD/JPY (DEXJPUS), USD/CNY (DEXCHUS), USD/INR (DEXINUS), USD/SGD (DEXSIUS), USD/THB (DEXTHUS), USD/MYR (DEXMAUS), EUR/USD (DEXUSEU), and GBP/USD (DEXUSUK). The panel contains six Asian currencies and two major developed-market comparators. Each series uses the same 1999–2024 sampling framework, removal of missing releases and non-trading days, and one-observation-ahead absolute-log-return target. No calendar interpolation is applied. For EUR/USD and GBP/USD, the quotation direction differs from that of the Asian series. This distinction does not affect the target because

$$\left|\Delta \log\left(C_{t}^{-1}\right)\right| = \left|\Delta \log\left(C_{t}\right)\right|. \tag{28}$$

The transferability test freezes all choices made for USD/KRW. The rolling window, moment orders, scale grid, detrending order, predictors, lag structure, tree architecture, hyperparameters, and random seeds remain unchanged. Each currency retains the same training, validation, and 2020–2024 test dates. The two $\Delta\alpha$ thresholds are estimated separately from that currency's pre-2020 development sample because the level and dispersion of spectrum width differ across series. No test observation enters model fitting, threshold construction, or specification selection. Two transported models are evaluated. The first is global Extra Trees augmented with all MFDFA variables. The second is the three-state MFRA-ET specification. Both are compared with the same market-only Extra Trees benchmark. Relative RMSE and MAE gains are calculated using Equation (16). Pairwise Diebold–Mariano tests use the five-lag Newey–West variance estimator defined in Section 3.4. To account for testing eight external currencies, Holm-adjusted conclusions are reported separately for each model-loss family. The exercise is a frozen-specification portability test. It does not estimate the best pair-specific MFDFA model. The 2020–2021 and 2022–2024 USD/KRW comparisons are treated as ex post descriptive subperiods. Let $G_{L,E}$ and $G_{L,L}$ denote the MFRA-ET loss gains in the earlier and later subperiods. Their difference is

$$D_L = G_{L,L} - G_{L,E}, L \in \{\mathrm{RMSE}, \mathrm{MAE}\}. \tag{29}$$

Uncertainty in D_L is estimated with 10,000 paired circular moving-block replications. The primary block length is 20 trading days. Dates are resampled separately within each subperiod, and the same indices are applied to realizations and competing forecasts. Two-sided empirical *p*-values test H_0: $D_L = 0$). This procedure determines whether the later-period forecasting gain is statistically larger. It does not identify an endogenous structural break or a causal effect of monetary tightening, geopolitical conflict, or any other event.

A larger percentage gain during a volatile subperiod need not imply stronger multifractal predictability. The relative gain in Equation (29) is invariant to a common positive rescaling of both model losses. However, time-varying volatility may affect the two forecast errors differently. Three further diagnostics therefore separate a mechanical scale effect from changes in volatility-state composition. First, forecast errors are normalized by two scale variables observed at forecast origin t. These are the 20-day mean absolute return, $\mathrm{MA20}_t$, and 20-day historical volatility, $\mathrm{HV20}_t$. For model $j \in \{\mathrm{ET}, \mathrm{MFRA}\}$ define $e_{j,t+1} = \alpha_{t+1} - \check{\alpha}^{(j)}_{t+1|t}$. For subperiod P and scale $s_t \in \{\mathrm{MA20}_t, \mathrm{HV20}_t\}$, the normalized losses are

$$\mathrm{NRMSE}_{j,P(s)} = \left[\frac{1}{np}\sum_{t\in P}\left(\frac{e_{j,t+1}}{s_t}\right)^2\right]^{\frac{1}{2}}, \mathrm{NMAE}_{j,P(s)} = \frac{1}{np}\sum_{t\in P}\frac{|e_{j,t+1}|}{s_t}. \tag{30}$$

Relative gains are calculated from these normalized losses using Equation (16). Both denominators are known when the forecast is formed. The procedure therefore introduces no future information. Second, volatility-composition effects are removed through direct standardization. Low-, medium-, and high-volatility states are defined by the pre-2020 development-sample terciles of $\mathrm{HV20}_t$. The fixed thresholds are 6.098 and 8.832. Each subperiod is then assigned the same one-third weight in every state. For $k \in \{1, 2, 3\}$,

$$\mathrm{MSE}^{\mathrm{bal}}_{j,P} = \frac{1}{3}\sum_{k=1}^{3}\left[\frac{1}{np,k}\sum_{t\in P} e^2_{j,t+1}\right], \mathrm{MAE}^{\mathrm{bal}}_{j,P} = \frac{1}{3}\sum_{k=1}^{3}\left[\frac{1}{np,k}\sum_{t\in P}|e_{j,t+1}|\right], \mathrm{RMSE}^{\mathrm{bal}}_{j,P} = \sqrt{\mathrm{RMSE}^{\mathrm{bal}}_{j,P}}. \tag{31}$$

This construction holds the volatility-state distribution constant. It therefore prevents a larger share of high-volatility observations from mechanically determining the subperiod comparison. Uncertainty is estimated with 10,000 paired circular moving-block replications.

Dates are sampled separately within each subperiod, and identical indices are applied to the target and both forecast sequences. The balanced intervals use the 9972 replications in which all three fixed volatility states appear in both subperiods. Third, conditional loss-differential regressions test whether a discrete later-period effect remains after observable market conditions are controlled. The dependent variables are $d_{r+1}^{(1)} = \frac{|e_{ET,t+1}| - |e_{MFRA,t+1}|}{MA20_t}$, $d_{r+1}^{(2)} = \frac{e_{ET,t+1}^2 - e_{MFRA,t+1}^2}{MA20_t^2}$. They are estimated from

$$d_{r+1}^{(q)} = \alpha_q + b_q Late_t + \gamma_q^\top \mathbf{W}_t + u_{t+1}^{(q)}, q \in \{1,2\}, \quad (32)$$

where $Late_t = 1$ for 2022–2024. The control vector $\mathbf{W}_t$ contains standardized values of log $MA20_t$, log 60-day historical volatility, the log ratio of 20-day to 60-day historical volatility, the current absolute return, and upside–downside semivolatility imbalance. All controls are observed at time t. Newey–West covariance estimates use 20 lags. A positive b_q indicates a larger conditional loss advantage during 2022–2024. It does not identify a causal effect of monetary tightening because the later-period indicator is not an exogenous treatment.

4. Results and Discussion

4.1. Distributional and Multifractal Evidence

The Korean foreign exchange market offers a demanding setting for evaluating scale-dependent dynamics. The won is traded, yet its behavior remains sensitive to global dollar cycles, regional capital flows, geopolitical risk, and shifts in domestic monetary conditions. These forces can produce abrupt movements alongside periods of relative calm. Accordingly, average persistence alone may conceal variation across fluctuation sizes and market states. The analysis therefore begins by examining both conventional distributional properties and rolling multifractal measures. Particular attention is given to H(2), the generalized-Hurst spread ΔH, and the singularity-spectrum width $\Delta\alpha$. To sum up, these statistics clarify whether near-random second-order scaling coexists with richer temporal heterogeneity. The evidence is summarized in Table 2.

Table 2. Descriptive statistics for market and rolling multifractal variables.

Variable	N	Mean	SD	Min	Median	Max	JB *p*
Daily USD/KRW reference rate	6268	1151.781	114.451	903.200	1141.840	1570.100	0.000
Return (%)	6268	0.004	0.656	−13.222	0.000	10.135	0.000
Absolute return (%)	6268	0.413	0.510	0.000	0.286	13.222	0.000
HV_{20}	6268	8.532	5.973	1.330	7.494	82.553	0.000
H(2)	6268	0.493	0.077	0.215	0.486	0.777	0.000
ΔH	6268	0.205	0.163	−0.324	0.200	0.984	0.000
$\Delta\alpha$	6268	0.368	0.215	0.007	0.339	1.482	0.000
Spectrum asymmetry	6268	0.062	0.466	−1.000	0.099	1.000	0.000

Note: N = 6268 for every variable. The cleaned raw series contains 6520 daily observations. The first 252 observations provide the initialization history required for the rolling MFDFA features and are therefore not included in the feature-aligned descriptive sample. Thus, 6520 − 252 = 6268. Jarque–Bera *p*-values test normality.

Table 2 shows that the USD/KRW rate averaged 1151.781 won per dollar, with observations ranging from 903.200 to 1570.100. The median, 1141.840, remains slightly below the mean. This broad interval reflects repeated episodes of won appreciation and depreciation across global and domestic stress cycles. Daily returns remain centered near zero, but their standard deviation is 0.656%. The extreme losses and gains also reveal sizable tail risk. Absolute returns are right-skewed: their mean of 0.413% exceeds the median

of 0.286%. Likewise, HV_{20} reaches 82.553, far above its mean of 8.532. These patterns are consistent with volatility clustering and short episodes of intense currency-market adjustment. Jarque–Bera tests reject normality for every variable. The rolling H(2) mean is 0.493, close to the benchmark of 0.5. Its range from 0.215 to 0.777, however, indicates alternating anti-persistent and persistent states. The positive mean values of ΔH and $\Delta\alpha$ further reveal heterogeneous scaling across fluctuation sizes. Their broad ranges show that multifractal complexity changes materially over time. Spectrum asymmetry averages only 0.062, yet its dispersion and full bounded range indicate pronounced shifts in spectrum shape. Descriptive evidence alone cannot separate temporal dependence from heavy-tailed distributional effects. That distinction requires the surrogate comparisons reported in Table 3.

Table 3. Full-sample MFDFA and expanded-surrogate evidence.

Series or Surrogate Null	B	H(2)	ΔH	Standardized Gap	p_{MC}	Δα	Standardized Gap	p_{MC}
Observed USD/KRW returns		0.489	0.296			0.578		
Shuffled surrogates	999	0.501	0.186	2.979	0.001	0.380	2.901	0.002
Phase-randomized surrogates	999	0.493	0.010	14.522	0.001	0.036	25.238	0.001
IAAFT surrogates	199	0.492	0.180	3.676	0.005	0.371	3.409	0.005

Note: Surrogate entries are simulation means. The standardized gap is the observed statistic minus the surrogate mean, divided by the surrogate standard deviation. Upper-tail Monte Carlo *p*-values are calculated as $p_{MC} = (1 + b)/(1 + B)$, where b is the number of surrogate statistics at least as large as the observed statistic. The exceedance counts $(b_{\Delta H}, b_{\Delta\alpha})$ are (0, 1) for the shuffled surrogates, (0, 0) for the phase-randomized surrogates, and (0, 0) for the IAAFT surrogates. Thus, $p_{MC} = 0.001$ with B = 999 and $p_{MC} = 0.005$ with B = 199 are the minimum attainable values and indicate zero surrogate exceedances. They should not be interpreted as zero population tail probabilities. Shuffling preserves the empirical marginal distribution but removes temporal ordering. Phase randomization preserves Fourier amplitudes but not the exact marginal distribution. IAAFT preserves the empirical marginal distribution and approximates the original power spectrum. These tests evaluate statistical scaling structure. They do not establish forecasting value, economic significance, or a causal mechanism.

Table 3 separates average second-order scaling from heterogeneity across fluctuation sizes. The observed H(2) is 0.489. Its proximity to 0.500 provides little support for strong average persistence. The width statistics provide a different result. The observed ΔH of 0.296 exceeds the shuffled mean of 0.186 by 2.979 surrogate standard deviations. None of the 999 shuffled realizations reaches the observed value, yielding b = 0 and $p_{MC} = 0.001$. The observed $\Delta\alpha$ of 0.578 exceeds the shuffled mean of 0.380 by 2.901 standard deviations. One shuffled realization reaches or exceeds the observed value, giving b = 1 and $p_{MC} = 0.002$. The shuffled results show that the heavy-tailed marginal distribution reproduces part, but not all, of the empirical spectrum width. Phase-randomized surrogates produce mean values of 0.010 for ΔH and 0.036 for $\Delta\alpha$. Neither observed width is exceeded by any of the 999 realizations. Both comparisons therefore have b = 0 and attain the Monte Carlo floor of $p_{MC} = 0.001$. The observed widths also exceed the IAAFT means of 0.180 and 0.371 by 3.676 and 3.409 surrogate standard deviations, respectively. No IAAFT realization exceeds either observed statistic, giving b = 0 and $p_{MC} = 0.005$ in both cases. The values of 0.001 and 0.005 are finite-simulation minima rather than exact population tail probabilities. Taken together, the results show that neither the empirical marginal distribution nor its combination with approximate linear spectral dependence reproduces the full observed width. This conclusion concerns statistical scaling structure only. Economic and predictive relevance are evaluated separately through feature ablation, locked forecasts, temporal-alignment placebos, and model-universe inference. Figure 2 displays the associated H(q) curves and singularity spectra.

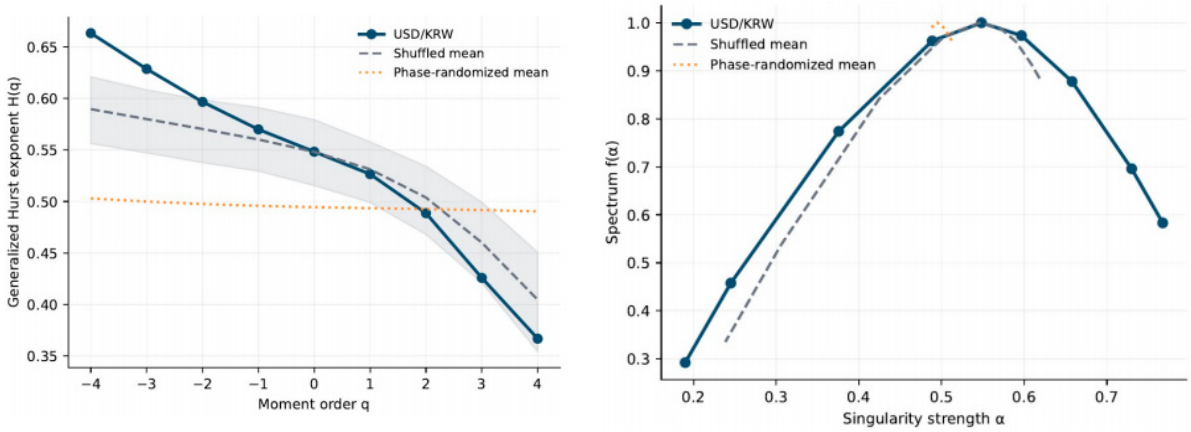

Figure 2. Generalized Hurst exponents and singularity spectra for USD/KRW returns and surrogate series. Panel (**a**) reports the empirical H(q) and the surrogate results. The gray shaded band represents the pointwise 90% simulation interval, defined by the 5th and 95th percentiles of 999 shuffled-surrogate realizations. The gray dashed line denotes the shuffled-surrogate mean, whereas the orange dotted line denotes the phase-randomized mean. Panel (**b**) reports the corresponding mean singularity spectra. The surrogate means are based on 999 shuffled and 999 phase-randomized realizations. Table 3 additionally reports results from 199 IAAFT realizations, together with the finite-simulation Monte Carlo *p*-values and their exceedance counts.

Figure 2 displays the full-sample multifractal structure. In panel (a), H(q) for USD/KRW declines from 0.66 at q = −4 to 0.37 at q = 4. Negative orders emphasize small fluctuations, whereas positive orders place weight on large movements. The slope implies that small changes are relatively persistent, whereas large shocks are anti-persistent and quickly corrected. This contrast is consistent with routine flows interrupted by adjustments to global dollar conditions, portfolio reallocations, and policy news. The shuffled curve slopes downward, but less sharply. Heavy tails explain part of the scale dependence. The flat phase-randomized curve remains close to 0.5, showing that linear dependence alone cannot reproduce the empirical pattern. Panel (b) reinforces this conclusion. The observed singularity spectrum is broad and wider than the phase-randomized spectrum. It extends beyond the shuffled spectrum, especially away from the peak. The common peak near f(α) = 1 indicates comparable dominant scaling behavior, but the different widths reveal sharply different local regularity. The broad empirical support reflects a richer mixture of scaling states. It is compatible with heterogeneous trading horizons and episodic stress in the Korean foreign exchange market. Taken together, both panels show that the won is near-random only in an average second-order sense. Its small and large fluctuations follow distinct scaling laws. Full-sample curves, however, cannot reveal when those states emerge or disappear. The time variation in H(2) and Δα is therefore examined in Figure 3.

Figure 3 shows that the scaling structure of USD/KRW returns is time-varying. Panel (a) reveals repeated crossings of the H(2) = 0.5 benchmark. Persistence alternates with anti-persistent phases rather than forming a stable long-memory regime. The sharp decline during the global financial crisis is followed by a rebound above 0.5. Similar reversals appear around COVID-19 and the 2022–2024 tightening cycle. These movements show how shocks propagate and dissipate differently. Panel (b) reveals greater instability in Δα. Spectrum width rises during the global financial crisis, when dollar funding stress and capital-flow reversals pressured the won. Smaller increases appear near the pandemic and later geopolitical and monetary disturbances. However, Δα does not move mechanically with the exchange-rate level or H(2). It measures dispersion across scaling behaviors, not the direction of won appreciation or depreciation. A market can display near-random

second-order persistence while remaining complex across fluctuation sizes. Recurring shifts in both panels caution against imposing one fractal relationship on the full sample. They support a forecasting design in which the MFDFA features are updated at every origin and the conditional response may vary across complexity states. Visual variation in $\Delta\alpha$, however, does not by itself establish drift in its tercile boundaries. A series may move frequently across relatively stable cutoffs. Section 4.5 therefore separates changes in annual state prevalence from changes in the expanding-sample threshold locations.

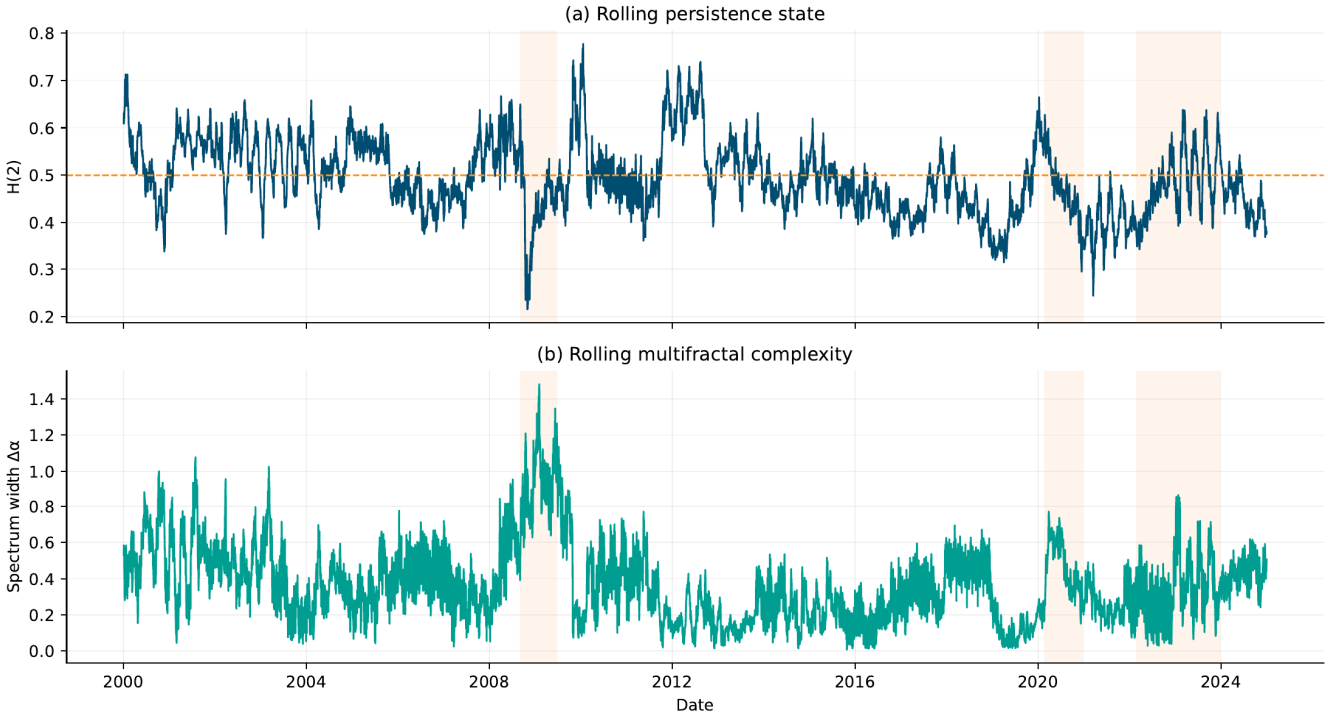

Figure 3. Rolling H(2) and singularity-spectrum width for USD/KRW returns. Note: The orange dashed horizontal line in panel (**a**) marks the H(2) = 0.5 benchmark associated with uncorrelated behavior; values above and below this line indicate persistent and anti-persistent behavior, respectively. The shaded regions denote the global financial crisis, the COVID-19 disruption, and the war-and-monetary-tightening period. These intervals are descriptive event windows and are not interpreted as causal treatment periods.

4.2. Out-of-Sample Forecasting Performance

The evidence establishes that USD/KRW scaling behavior changes across time and fluctuation sizes. The next question is whether this structure improves real-time forecasts. The evaluation therefore uses the untouched 2020–2024 test period, which spans the COVID-19 disruption, global monetary tightening, geopolitical shocks, and won depreciation. This interval provides a demanding setting for assessing model robustness. Forecast accuracy is compared across statistical benchmarks, market-only machine-learning models, fractal-augmented specifications, and the proposed MFRA-ET model. RMSE and MAE measure losses in the original absolute-return scale. MASE evaluates performance relative to the persistence benchmark, while R^2 and forecast–realization correlation capture explanatory content and directional co-movement in risk magnitude. Because every hyperparameter is selected before the test period, the comparison reflects out-of-sample performance rather than in-sample fit. The complete results are reported in Table 4.

Table 4. Out-of-sample one-trading-day-ahead absolute-return forecasts, 2020–2024.

Model	RMSE	MAE	MASE	R^2	Corr.
MFRA-ET	0.367	0.271	0.747	0.045	0.218
Extra Trees + MFDFA	0.369	0.274	0.755	0.032	0.186
Random forest + MFDFA	0.370	0.276	0.759	0.027	0.181
Market-only histogram gradient boosting	0.371	0.274	0.755	0.023	0.183
20-day historical mean	0.372	0.278	0.765	0.016	0.202
Market-only random forest	0.372	0.278	0.766	0.015	0.172
Market-only Extra Trees	0.373	0.277	0.765	0.013	0.157
Histogram gradient boosting + MFDFA	0.373	0.274	0.756	0.009	0.157
Market-only linear regression	0.380	0.285	0.784	−0.028	0.171
Linear regression + MFDFA	0.382	0.287	0.791	−0.036	0.159
Persistence	0.501	0.363	1.000	−0.782	0.109

Note: Matched market-only and MFDFA-augmented specifications use identical algorithms, lag structures, sample partitions, and evaluation dates. Lower RMSE, MAE, and MASE are preferred. The test sample contains 1249 observations. All hyperparameters were selected before the test period. The reported percentage gains are relative reductions in forecast loss. They are not portfolio returns. The approximately 0.006-percentage-point reduction in each principal error measure equals about 0.6 basis points in forecast-error units and should not be compared directly with a bid–ask spread. R^2 is defined relative to the unconditional mean of the locked test sample. It measures total variation explained and is not treated as a stand-alone test of predictability. Paired incremental estimates, dependence-robust intervals, and temporal-alignment diagnostics are examined separately in the subsequent robustness analyses.

Table 4 establishes a modest out-of-sample loss advantage, not strong explanatory power. MFRA-ET records the lowest RMSE (0.367), MAE (0.271), and MASE (0.747). Its R^2 of 0.045 means that the model explains about 4.5% of test-period variation. Its forecast–realization correlation of 0.218 also indicates weak co-movement. Both statistics are limited in absolute terms. The more informative comparison is the matched market-only Extra Trees model. Relative to that benchmark, R^2 rises from 0.013 to 0.045, while correlation rises from 0.157 to 0.218. The MASE of 0.747 indicates lower MAE than persistence, but persistence is a weak benchmark. Market-only Extra Trees therefore remains the central comparator. Adding the seven MFDFA variables lowers RMSE from 0.373 to 0.370 and MAE from 0.277 to 0.274. Subsequent $\Delta\alpha_t$-based routing reduces the two losses to 0.367 and 0.271. Relative to market-only Extra Trees, the joint reductions are 1.627% for RMSE and 2.245% for MAE. The corresponding changes are approximately 0.006 percentage points in the original return scale and remain small in economic units. Although the Diebold–Mariano tests indicate that these loss reductions are unlikely to reflect sampling variation alone, statistical significance over 1249 test observations does not imply an economically large effect. The practical contribution of MFRA-ET is therefore a marginal reduction in average forecast loss, rather than a material improvement in point-prediction accuracy. Before multiplicity adjustment, routing beyond global MFDFA-augmented Extra Trees is nominally significant under absolute loss ($p = 0.026$) but not under squared loss ($p = 0.099$). The absolute-loss result does not survive either Holm ($p = 0.077$) or Bonferroni ($p = 0.257$) adjustment. The corresponding adjusted p-values under squared loss are 0.099 and 0.989. The lower point losses of MFRA-ET therefore indicate a small descriptive routing increment, not family-wise-error-controlled evidence of routing superiority. More importantly, this nested sequence alone cannot determine whether the improvement reflects MFDFA content, additional predictor capacity, $\Delta\alpha_t$-based state information, or generic local estimation. Section 4.5 therefore reports capacity-matched controls, a 2×2 decomposition, alternative routing variables, and MFDFA-block displacement placebos. Sampling uncertainty and calendar-time alignment are evaluated separately in Section 4.6. Because the target contains

no directional information, none of these loss reductions can be interpreted as a trading return or compared directly with transaction costs.

The matched learner comparisons reveal clear architecture dependence. Adding MFDFA variables to linear regression changes RMSE and MAE by −0.387% and −0.846%, respectively. The corresponding changes are 0.664% and 1.010% for Extra Trees, 0.825% and 1.090% for random forest, and −0.711% and −0.208% for histogram gradient boosting. MFDFA variables therefore do not provide a model-independent forecasting gain. Their positive contribution is confined to the two bagged-tree learners in the present experiment. Algorithmic randomization does not explain the positive bagged-tree results. Across 30 fixed seeds, Extra Trees produces positive RMSE and MAE gains in all 30 cases. Its median RMSE gain is 0.780%, with a range of 0.602–0.941%. Its median MAE gain is 1.100%, with a range of 0.839–1.267%. Random forest also produces positive gains in all 30 cases. Its median RMSE and MAE gains are 0.599% and 0.922%, with ranges of 0.332–0.849% and 0.599–1.269%. These repetitions assess computational stability rather than sampling uncertainty. The evidence supports architecture-conditional information for bagged trees. It does not support universal usefulness across forecasting algorithms. Explanatory-power and temporal-alignment diagnostics are reported in Section 4.6. Figure 4 first examines when the daily error differences arise within the locked test period. Aggregate metrics cannot reveal whether the advantage is concentrated around turbulence or accumulated across ordinary days. The temporal correspondence between realized risk and competing forecasts is therefore examined in Figure 4.

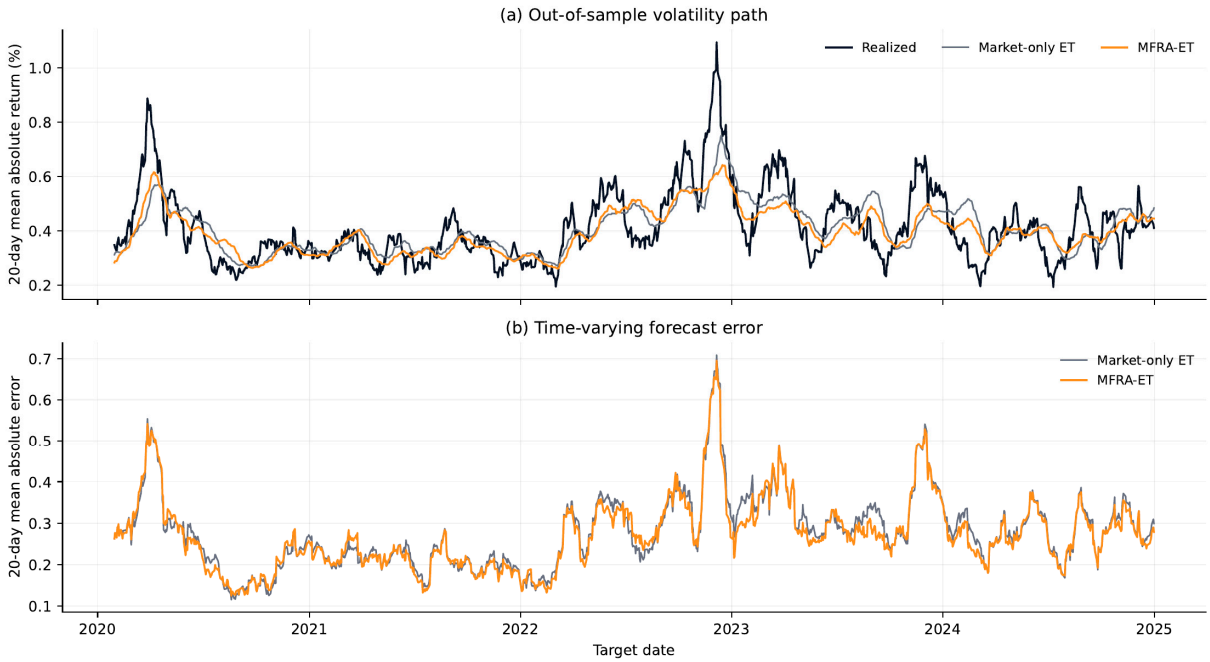

Figure 4. Out-of-sample observed and predicted daily absolute-return magnitude, 2020–2024. Note: All plotted series are 20-day moving averages used for visual clarity; formal evaluation is based on unsmoothed daily forecasts.

Figure 4 compares 20-day moving averages of observed absolute returns and competing forecasts. Panel (a) shows that both Extra Trees models capture broad changes in daily USD/KRW displacement. They identify the large reference-rate movements during the early COVID-19 period and the renewed instability during 2022–2024. Neither model reproduces every extreme. The forecasts are smoother than the observed magnitudes and understate several abrupt changes, especially in late 2022. This result is expected because the next daily observation contains substantial unanticipated information. The figure does

not describe intraday volatility. A temporary within-day jump that reverses before the next H.10 observation is absent from both the target and the plotted series. Within this daily-data scope, MFRA-ET adjusts differently when the estimated complexity state changes. It tracks several increases and reversals more closely than market-only Extra Trees, although the advantage is not uniform. Panel (b) shows that the rolling absolute-error paths overlap during stable periods. Wider gaps emerge near changes in the daily magnitude process. The aggregate advantage therefore accumulates through many small forecast differences rather than one extreme observation. The sign-stratified analysis reported in Section 4.6 further shows that this gain is stronger on won-appreciation days. Figure 4 consequently supports a limited state-dependent daily-magnitude result. It does not provide evidence about intraday integrated variance or directional profitability.

4.3. Predictive-Accuracy Tests

Point forecast rankings do not establish whether loss differences are statistically distinguishable from sampling variation. This issue is important for USD/KRW risk, where forecast errors are noisy and serially dependent. The analysis therefore applies Diebold–Mariano tests to each comparator and MFRA-ET. Squared loss emphasizes large misses, whereas absolute loss gives a robust assessment across ordinary and turbulent observations. The loss differential is defined as comparator loss minus MFRA-ET loss, so a positive statistic favors the model. Newey–West variance estimation with five lags accounts for short-run dependence in the differential. Testing both losses shows whether the conclusion depends on a few extreme exchange-rate movements. The results are presented in Table 5.

Table 5. Diebold–Mariano tests with family-wise multiplicity adjustments: comparator versus MFRA-ET.

Comparator	Loss	DM	Raw p	Holm-Adjusted p	Bonferroni-Adjusted p	Mean Loss Gain
Persistence	Squared	9.531	0.000	0.000	0.000	0.116
Persistence	Absolute	11.592	0.000	0.000	0.000	0.092
20-day historical mean	Squared	1.997	0.046	0.092	0.458	0.004
20-day historical mean	Absolute	2.576	0.010	0.040	0.100	0.007
Linear regression	Squared	3.905	0.000	0.000	0.000	0.010
Linear regression	Absolute	4.407	0.000	0.000	0.000	0.014
Extra Trees	Squared	2.761	0.006	0.029	0.058	0.004
Extra Trees	Absolute	3.301	0.001	0.006	0.010	0.007
Extra Trees + MFDFA	Squared	1.650	0.099	0.099	0.989	0.002
Extra Trees + MFDFA	Absolute	2.231	0.026	0.077	0.257	0.003

Note: The loss differential is defined as comparator loss minus MFRA-ET loss. A positive DM statistic and a positive mean loss gain therefore favor MFRA-ET. All tests are two-sided. The long-run variance of each loss differential is estimated using the Newey–West estimator with five lags. Squared and absolute losses correspond to Equations (9) and (10), respectively. The ten tests in this table are treated as one inferential family. Holm-adjusted p-values provide the primary family-wise error-rate correction, while Bonferroni-adjusted p-values provide a more conservative sensitivity check. Adjusted values are calculated from the unrounded p-values. Values below 0.001 are reported as $p < 0.001$. Mean loss gain is the average comparator loss minus the average MFRA-ET loss.

Table 5 reports positive DM statistics in all ten comparisons, but the multiplicity corrections narrow the inferential conclusions. Under the primary Holm procedure, seven comparisons remain significant at the 0.050 level. Five remain significant under the more conservative Bonferroni procedure. Equality of predictive accuracy is rejected against persistence and linear regression under both loss functions after either correction. Against the 20-day historical mean, the absolute-loss result survives Holm adjustment ($p = 0.040$), whereas the squared-loss result does not ($p = 0.092$). Neither comparison survives the

Bonferroni correction. Market-only Extra Trees provides the most informative benchmark because it holds the learner, lag structure, and conventional predictor set constant. Relative to this benchmark, MFRA-ET produces DM statistics of 2.761 under squared loss and 3.301 under absolute loss. The corresponding Holm-adjusted p-values are 0.029 and 0.006. Under the Bonferroni correction, the absolute-loss result remains significant ($p = 0.010$), whereas the squared-loss result becomes marginal ($p = 0.058$). The evidence for the joint MFRA-ET architecture is therefore strongest under absolute loss. The comparison with global Extra Trees + MFDFA addresses the narrower question of whether routing adds predictive accuracy after the multifractal inputs have already been included. Under squared loss, the raw and Holm-adjusted p-values are both 0.099, and the Bonferroni-adjusted value is 0.989. Under absolute loss, the raw p-value of 0.026 rises to 0.077 after Holm adjustment and to 0.257 after Bonferroni adjustment. Thus, neither routing comparison remains significant after family-wise error control. MFRA-ET has lower point losses than the non-routed MFDFA model under both loss functions, but these differences provide only descriptive evidence of a small and loss-dependent routing increment. They do not establish the statistical superiority of routing. The broader predictive conclusion therefore rests on the matched comparison with market-only Extra Trees and the subsequent model-universe inference, rather than on the routing contrast alone.

4.4. Descriptive Subperiod Performance and Feature Interpretation

Full-sample tests establish an average USD/KRW loss reduction, but they do not show whether its magnitude is stable through time. The locked sample is therefore divided into the ex post intervals 2020–2021 and 2022–2024. These periods coincide with different monetary, geopolitical, and volatility conditions. They are not estimated structural regimes or causal treatments. Panel A of Table 6 reports the unadjusted forecast losses, while Panel B tests the raw difference in percentage gains. Panel C documents the change in the volatility environment. Panel D reports scale-normalized and volatility-balanced comparisons. Panel E evaluates whether a later-period loss advantage remains after conventional market-state controls are included. The analysis distinguishes temporal heterogeneity from causal attribution.

Table 6 confirms that the volatility environment changed materially after 2021. Mean daily absolute return rises from 0.361 to 0.443, an increase of 22.828%. Mean ex ante $HA20_t$ rises from 7.254 to 9.126, or 25.819%. More strikingly, the share of observations in the high-$HA20_t$ state increases from 12.625% to 46.933%. The reviewer's scale and composition concerns are therefore empirically relevant. The unadjusted comparison reproduces the original temporal pattern. MFRA-ET lowers RMSE by 0.423% and MAE by 0.547% during 2020–2021. The corresponding gains are 2.069% and 3.081% during 2022–2024. The raw later-minus-earlier differences are 1.646 and 2.535 percentage points. Their moving-block p-values are 0.147 and 0.103. Neither raw difference is statistically distinguishable from zero. A common multiplicative increase in volatility cannot mechanically raise the percentage gain because the scale factor cancels from both the numerator and denominator of Equation (16). The direct normalization results provide a stronger test. After errors are divided by $MA20_t$, the later-minus-earlier RMSE difference is 2.015 percentage points ($p = 0.047$). The corresponding MAE difference is not significant ($p = 0.134$). Scaling by $HA20_t$ produces the same loss-dependent pattern. The RMSE difference is significant at $p = 0.043$, whereas the MAE difference remains nonsignificant at $p = 0.129$. A purely mechanical scale effect therefore cannot account for the normalized RMSE result. The evidence is not uniform across loss measures. Volatility-composition balancing leads to a different conclusion. Once both subperiods receive the same one-third weight in each pre-specified $HA20_t$ state, the RMSE gains become 1.708% and 1.990%. Their difference falls from 1.646 to 0.282 percentage

points and is not significant ($p = 0.907$). The balanced MAE gains are 0.881% and 2.731%. Their difference is also nonsignificant ($p = 0.410$). Hence, the greater concentration of high-volatility observations explains a substantial portion of the raw calendar-period contrast. The conditional regressions reinforce this boundary. After observable market conditions are controlled, the 2022–2024 coefficient is 0.007 under scaled absolute loss, with $p = 0.516$. It is 0.027 under scaled squared loss, with $p = 0.139$. Neither coefficient is distinguishable from zero. These regressions cannot absorb every institutional or policy change. They nevertheless show that a discrete later-period advantage is not robust after conventional market-state variables are included. The combined evidence has a precise interpretation. Higher volatility does not inflate percentage gains through simple arithmetic. Yet the volatility-state composition accounts for much of the apparent RMSE contrast, and no robust later-period coefficient remains after observable controls. Federal Reserve tightening, dollar strength, geopolitical conflict, and capital-flow changes remain plausible competing explanations. They are historical context, not identified mechanisms. The 2022–2024 result is therefore retained only as a descriptive point pattern. It is not evidence that crises or structural shifts increase the predictive value of MFDFA variables.

Practical implementation entails additional but bounded computational and governance costs. Relative to market-only Extra Trees, the MFDFA layer adds 35 lagged columns and requires one rolling MFDFA update at each daily forecast origin. The routing architecture stores three 800-tree ensembles, or 2400 trees in total, compared with 600 trees in the global Extra Trees specification. Each local ensemble is trained on an approximately balanced subsample of 1671, 1671, or 1672 development observations. The online forecasting burden is smaller than the total stored model suggests. Only one 800-tree ensemble is traversed after the current state is assigned. The number of trees evaluated per forecast therefore rises from 600 to 800, an increase of 33.333%, rather than requiring all 2400 trees to be evaluated. The principal design also does not involve daily model refitting. Its recurring workload consists of one 252-observation MFDFA update, one state assignment, and one ensemble prediction. This additional workload must be considered relative to the modest forecast improvement. MFRA-ET reduces RMSE from 0.373 to 0.367 and MAE from 0.277 to 0.271 relative to market-only Extra Trees. The corresponding gains are 1.627% and 2.245%. In the original return scale, each reduction is approximately 0.006 percentage points, or 0.600 basis points, in forecast-error units. Moreover, the routing increment beyond global MFDFA-augmented Extra Trees does not survive Holm or Bonferroni adjustment. Computational feasibility therefore does not imply that the model delivers a positive net operational benefit for every user. The relevant institutional costs also include feature-pipeline validation, numerical monitoring of negative-order moments, threshold-drift surveillance, audit documentation, and periodic model review. For a central bank or financial institution with an established end-of-day analytics system, MFRA-ET may be operated as a supplementary surveillance layer. A shadow-deployment design would allow its forecasts and prediction bands to be compared continuously with those of the simpler market-only model. The model should be retained only if its incremental performance remains stable after implementation and governance costs are considered. For users without this infrastructure, or for applications requiring intraday decisions, market-only Extra Trees remains the more parsimonious specification. Figure 5 next examines the model's reliance on each feature group and the sampling stability of the resulting importance order.

Table 6. Forecast performance, volatility adjustment, and gain-difference inference across descriptive USD/KRW subperiods.

Panel A. Forecast Performance

Period	Model	RMSE	MAE	MASE	R^2	Corr.
2020–2021	Extra Trees	0.305	0.229	0.734	0.035	0.195
2020–2021	Extra Trees + MFDFA	0.304	0.229	0.733	0.044	0.209
2020–2021	MFRA-ET	0.304	0.228	0.729	0.044	0.218
2022–2024	Extra Trees	0.411	0.310	0.781	−0.011	0.102
2022–2024	Extra Trees + MFDFA	0.407	0.304	0.766	0.012	0.140
2022–2024	MFRA-ET	0.403	0.300	0.756	0.030	0.183

Panel B. Difference Between MFRA-ET Gains

Loss Measure	2020–2021 Gain	2022–2024 Gain	Later-Minus-Earlier Difference	20-day Block 95% CI	*p*-Value
RMSE	0.423%	2.069%	1.646 percentage points	[−0.524, 3.975]	0.147
MAE	0.547%	3.081%	2.535 percentage points	[−0.505, 5.561]	0.103

Panel C. Ex Ante Volatility Conditions

Statistic	2020–2021	2022–2024	Later-Period Change
Mean daily absolute return	0.361	0.443	22.828%
Mean 20-day historical volatility	7.254	9.126	25.819%
High-HV20 state share	12.625%	46.933%	34.308 percentage points

Panel D. Scale-Normalized and Volatility-Balanced Gain differences

Adjustment and Loss Measure	2020–2021 Gain (%)	2022–2024 Gain (%)	Later-Minus-Earlier Difference	20-day Block 95% CI	*p*-Value
$MA20_t$-scaled RMSE	−0.129	1.886	2.015	[0.037, 3.962]	0.047
$MA20_t$-scaled MAE	0.686	2.879	2.193	[−0.686, 5.040]	0.134
$HA20_t$-scaled RMSE	−0.147	1.890	2.036	[0.070, 3.990]	0.043
$HA20_t$-scaled MAE	0.610	2.856	2.246	[−0.649, 5.113]	0.129
Volatility-balanced RMSE	1.708	1.990	0.282	[−3.949, 3.602]	0.907
Volatility-balanced MAE	0.881	2.731	1.850	[−4.299, 5.434]	0.410

Panel E. Conditional Later-Period Coefficients from Loss-Differential Regressions

Dependent Variable	Later-Period Coefficient	Newey–West SE	*p*-Value
$MA20_t$-scaled absolute-loss differential	0.007	0.011	0.516
$HA20_t$-scaled squared-loss differential	0.027	0.018	0.139

Note: Panel A reports descriptive losses for the two ex post subperiods. RMSE and MAE are reported to three decimal places throughout Section 4.4, consistent with Table 4. All percentage gains, confidence intervals, and inferential statistics are computed from unrounded loss values. Panel B uses 10,000 paired circular moving-block replications. Panel C uses information observed at each forecast origin. The high-volatility state is defined by the upper development-sample tercile of $HA20_t$, whose fixed cutoff is 8.832. Panel D normalizes individual forecast errors by $MA20_t$ or $HA20_t$. Its volatility-balanced rows assign equal weight to the three development-sample $HA20_t$ states. The balanced confidence intervals use the 9972 moving-block replications that retain all three states in both subperiods. Panel E reports Newey–West estimates with 20 lags and 1249 observations. Positive gains and positive regression coefficients favor MFRA-ET. Neither subperiod is an estimated structural regime or a causal treatment period.

Figure 5 replaces the original 25-draw row-permutation ranking with a dependence-aware assessment. Panel (a) shows that ΔH has the largest point importance, at 4.904 in 10^{-4} RMSE units. Its 95% interval is [0.879, 8.903]. The point importance of $\Delta\alpha$ is 4.164, with an interval of [−0.550, 8.964]. The 20-day mean absolute return follows at 4.025, but its interval is much wider at [−8.739, 16.877]. The intervals overlap substantially. The point order therefore cannot be interpreted as a statistically identified ranking. Panel (b) makes this uncertainty explicit. Across 5000 moving-block replications, $\Delta\alpha$ ranks first in 12.040% of cases and enters the top three in 65.520%. Its median rank is 3.000, and its 95% rank range is 1.000–8.000. By comparison, ΔH ranks first in 25.840% of replications and enters

the top three in 83.700%. The 20-day mean absolute return ranks first in 36.420%, although its importance estimate is considerably less precise. These frequencies place ΔH and $\Delta\alpha$ in an upper importance tier. They do not establish a stable first-place variable. Formal paired comparisons lead to the same conclusion. The $\Delta\alpha - \Delta H$ importance difference is -0.740 in 10^{-4} RMSE units. Its 95% interval is $[-3.461, 1.981]$, and the Holm-adjusted p-value is 1.000. The difference relative to the 20-day mean absolute return is 0.140, with an interval of $[-11.653, 12.433]$ and an adjusted p-value of 1.000. Relative to upside RV20, the difference is 3.478, with an interval of $[-10.973, 17.406]$ and an adjusted p-value of 0.893. None of the 16 comparisons supports the superiority of $\Delta\alpha$ after multiplicity adjustment. Negative importance estimates remain possible because the predictors are correlated and Extra Trees can use substitute variables. The overlap is especially clear for the negatively ranked groups. H(4) enters the construction of ΔH directly, while spectrum asymmetry and $\Delta\alpha$ are derived from the same singularity spectrum. Absolute return and contemporaneous return are also likely to overlap with the information summarized by the RV20 measures. Their negative estimates therefore most plausibly reflect substitution among correlated predictors, rather than true unimportance. Permuting one group may remove noise or allow another group to operate more effectively. A negative value therefore indicates weak or unstable marginal reliance within this fitted model. It is not evidence of a negative economic effect. Figure 5 is consequently interpreted as a model-reliance diagnostic. It does not establish causal importance or a unique predictor hierarchy. The predictive relevance of $\Delta\alpha$ instead rests on prespecified loss comparisons. Subsequent matched forecast tests separately evaluate its incremental contribution as a predictor and as a routing variable. These tests support the use of $\Delta\alpha$ within the forecasting architecture, but they do not imply that it is uniquely dominant.

4.5. Rolling-Origin Robustness, Specification Sensitivity, and Decision Relevance

The locked test establishes one historical out-of-sample comparison. It does not show whether the ranking survives repeated forecast origins. Table 7 therefore reports annual expanding-window estimates. Model definitions and hyperparameters remain fixed, while the estimation sample and $\Delta\alpha$ thresholds use only information available before each evaluation year. The results are shown in Table 7.

Panel A of Table 7 shows that the expanding thresholds move much less than the underlying $\Delta\alpha_t$ sequence displayed in Figure 3. The lower boundary ranges from 0.241 to 0.248, while the upper boundary ranges from 0.422 to 0.434. Relative to the fixed thresholds of 0.243 and 0.430, only 22 of the 1249 test observations change state. The pooled reassignment rate is therefore 1.761%. Threshold stability does not imply constant state prevalence. The updated high-state share falls to 5.221% in 2021 and rises to 54.980% in 2024. The routing system consequently remains responsive to persistent movement in the realized $\Delta\alpha_t$ distribution. Panel B shows that updating the thresholds does not overturn the forecast ranking. MFRA-ET reduces RMSE in all five evaluation years. It reduces MAE in four years, with a deterioration of 0.400% in 2020. The pooled expanding-window gains are 1.470% for RMSE and 2.320% for MAE. Their 20-day moving-block 95% confidence intervals are [0.470%, 2.390%] and [0.890%, 3.850%], respectively. Both remain positive with 5-day and 60-day blocks. These results are close to the fixed-threshold gains of 1.627% and 2.245% in the principal locked experiment. The main finding is therefore not generated by stale pre-2020 cutoffs. The heterogeneous annual gains nevertheless show that the conditional relationship is not constant. They do not identify a crisis, tightening, or structural-shift effect. The stability of these findings is examined further in Table 8, which evaluates the sensitivity of the results to the MFDFA design and feature composition in both the pre-test validation sample and the locked test sample.

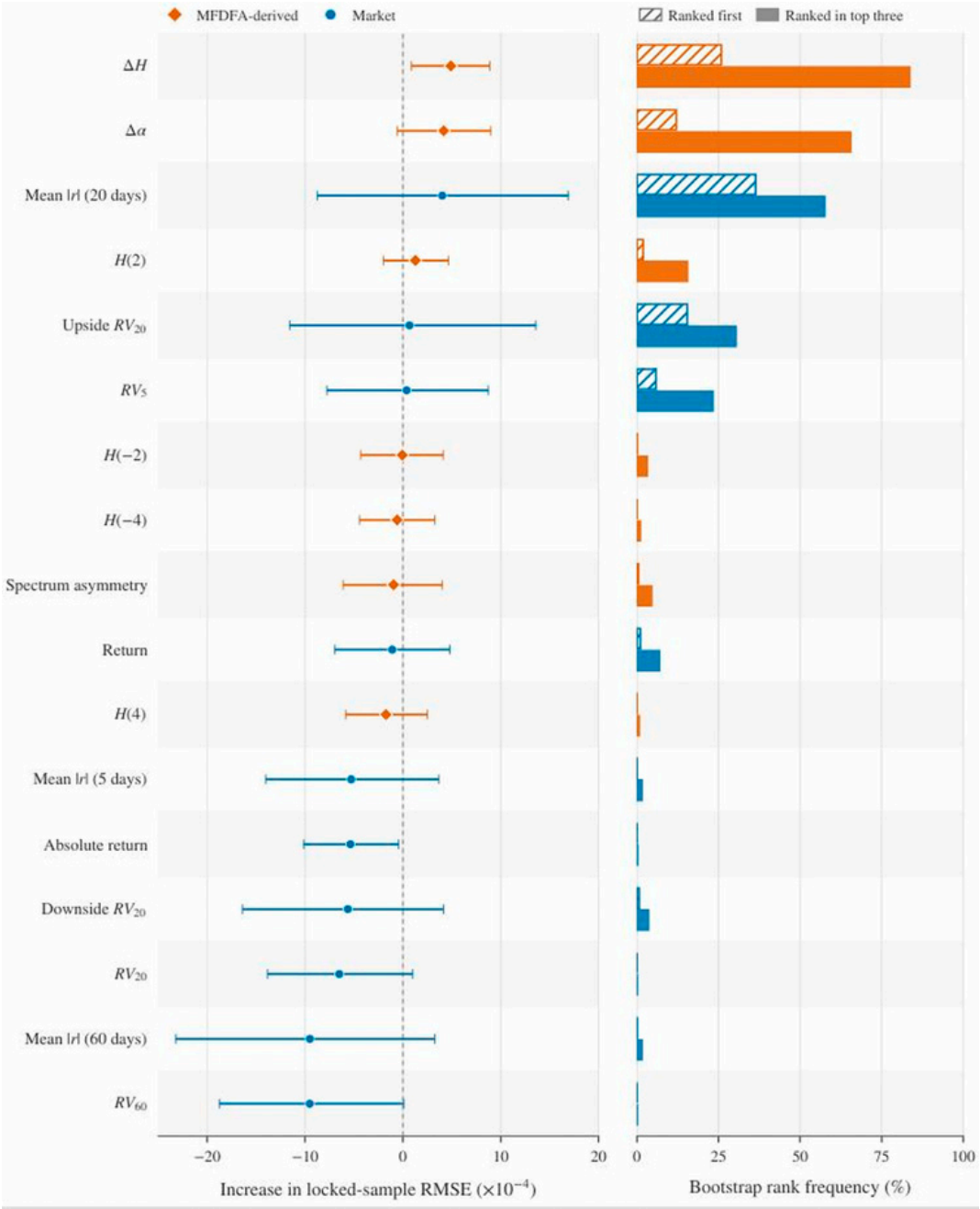

Figure 5. Dependence-aware grouped permutation importance and rank stability for global MFDFA-augmented Extra Trees. Panel (**a**) shows the increase in locked-sample RMSE, expressed in units of 10^{-4} percentage points of daily absolute return. Panel (**b**) shows the bootstrap frequency (%). Note: Panel (**a**) reports the mean change in locked-sample RMSE after 200 random reorderings of nonoverlapping 20-trading-day blocks. RMSE is measured on the same percentage-point scale as in Table 4. Thus, an axis value of 1.000 represents an RMSE increase of 0.0001 percentage points. All five lags of a feature are permuted jointly, and the same block-order map is applied to every feature group within each replication. Error bars are 95% percentile intervals from 5000 paired circular moving-block replications with a 20-day block length. Panel (**b**) reports the frequencies with which each feature ranks first or enters the top three. Negative importance means that permuting a feature group slightly lowers RMSE relative to the unpermuted model. With correlated predictors, this result can reflect redundancy, predictor substitution, or unstable marginal reliance; it does not establish intrinsic unimportance or a negative economic effect. Overlapping intervals and rank frequencies should be used to assess ordering uncertainty.

Table 7. Annual expanding-window threshold updating and rolling-origin forecast performance.

Panel A. Expanding-Window $\Delta\alpha$ Thresholds and Annual State Assignments							
Evaluation Year	**NN**	**ET RMSE**	**MFRA-ET RMSE**	**RMSE Gain (%)**	**ET MAE**	**MFRA-ET MAE**	**MAE Gain (%)**
2020	250	0.343	0.340	0.77	0.249	0.250	−0.40
2021	249	0.261	0.261	0.33	0.210	0.206	1.74
2022	250	0.436	0.428	1.80	0.311	0.304	2.15
2023	249	0.439	0.434	1.22	0.340	0.327	3.74
2024	251	0.353	0.343	2.67	0.281	0.271	3.62
Pooled	1249	0.372	0.367	1.47	0.278	0.272	2.32
Panel B. Forecast Performance Under Annual Expanding-Window Threshold Updating							
Evaluation Year	**N**	**ET RMSE**	**MFRA-ET RMSE**	**RMSE Gain (%)**	**ET MAE**	**MFRA-ET MAE**	**MAE Gain (%)**
2020	250	0.343	0.340	0.77	0.249	0.250	−0.40
2021	249	0.261	0.261	0.33	0.210	0.206	1.74
2022	250	0.436	0.428	1.80	0.311	0.304	2.15
2023	249	0.439	0.434	1.22	0.340	0.327	3.74
2024	251	0.353	0.343	2.67	0.281	0.271	3.62
Pooled	1249	0.372	0.367	1.47	0.278	0.272	2.32

Note: At each annual origin, the estimation sample ends on 31 December of the preceding year. The lower and upper thresholds are the expanding-sample terciles of $\Delta\alpha_t$. The algorithm, feature definitions, lag structure, hyperparameters, tree counts, and random seeds remain fixed. "Reassigned" is the percentage of annual observations whose state under the expanding thresholds differs from that obtained using the fixed pre-2020 thresholds of 0.243 and 0.430. Annual state shares are not constrained to be equal. Percentage gains are calculated from unrounded losses, and positive values favor MFRA-ET. The 20-day moving-block 95% confidence intervals are [0.470%, 2.390%] for the pooled RMSE gain and [0.890%, 3.850%] for the pooled MAE gain. Both intervals remain positive under 5-day and 60-day block lengths.

Table 8. Pre-test and locked-test sensitivity to the MFDFA design and feature composition.

Panel A. Pre-Test Validation Audit Across Rolling Windows, 2017–2019						
Window	**ET RMSE**	**MFRA-ET RMSE**	**RMSE Gain**	**ET MAE**	**MFRA-ET MAE**	**MAE Gain**
126	0.309	0.308	0.136%	0.229	0.230	−0.456%
252	0.309	0.308	0.261%	0.229	0.230	−0.355%
504	0.308	0.313	−1.351%	0.229	0.235	−2.645%
Panel B. Locked-Test Rolling-Window and Detrending Sensitivity, 2020–2024							
Window	**Order**	**ET RMSE**	**MFRA-ET RMSE**	**RMSE Gain**	**ET MAE**	**MFRA-ET MAE**	**MAE Gain**
126	1	0.372	0.371	0.340%	0.277	0.275	0.590%
252	1	0.373	0.367	1.627%	0.277	0.271	2.245%
504	1	0.372	0.371	0.300%	0.277	0.277	0.060%
252	2	0.373	0.370	0.850%	0.277	0.273	1.650%
252	3	0.373	0.371	0.550%	0.277	0.273	1.590%
Panel C. Scale-Grid Sensitivity Under a 252-day Window							
Scale Grid		**Validation RMSE Gain**		**Validation MAE Gain**	**Test RMSE Gain**	**Test MAE Gain**	
{8, 12, 16, 24, 32, 48}		−0.104%		0.099%	0.213%	1.323%	
{10, 15, 20, 30, 40, 60}		0.261%		−0.355%	1.627%	2.245%	
{12, 18, 24, 36, 48, 60}		−0.319%		−1.932%	1.494%	1.727%	

Table 8. *Cont.*

Panel D. Next-Day Absolute Returns Across ex Ante Δ_α States			
Δ_α State	**Development N**	**Development Mean**	**Test Mean**
Low, $\Delta\alpha \leq 0.243$	1671	0.378	0.424
Medium, $0.243 < \Delta\alpha \leq 0.430$	1671	0.358	0.404
High, $\Delta\alpha > 0.430$	1672	0.503	0.407

Panel E. Feature Composition Under a Common Extra Trees Specification				
Feature Specification	**RMSE**	**RMSE Gain**	**MAE**	**MAE Gain**
Market only	0.373		0.277	
Market + H(q) exponents	0.370	0.680%	0.275	0.970%
Market + ΔH, $\Delta\alpha$, Aα	0.370	0.860%	0.274	1.360%
Market + ΔH, $\Delta\alpha$	0.370	0.730%	0.274	1.160%
Market + $\Delta\alpha$	0.370	0.670%	0.275	1.040%
Market + all MFDFA variables	0.370	0.810%	0.274	1.180%
MFRA-ET	0.367	1.627%	0.271	2.245%

Panel F. Symmetric Moment-Range Sensitivity Under the 252-day Baseline Design					
q_{max}	**Moment Grid**	**Validation RMSE Gain (%)**	**Validation MAE Gain (%)**	**Test RMSE Gain (%)**	**Test MAE Gain (%)**
2	$\{-2, \ldots, 2\}$	−0.722	−1.765	1.786	2.176
3	$\{-3, \ldots, 3\}$	−0.117	−0.887	1.718	2.142
4	$\{-4, \ldots, 4\}$	0.261	−0.355	1.627	2.245
5	$\{-5, \ldots, 5\}$	0.304	−0.114	1.370	2.226

Panel G. Equal-Frequency Routing Sensitivity Relative to Global Extra Trees with MFDFA				
Number of States	**Validation RMSE Gain (%)**	**Validation MAE Gain (%)**	**Test RMSE Gain (%)**	**Test MAE Gain (%)**
2	−0.200	−0.285	0.748	0.602
3	0.474	−0.070	0.824	1.082
4	−0.143	−0.592	0.078	0.188
5	0.428	0.224	−0.077	0.395

Note: Panel A uses only the 2017–2019 validation period. No locked-test observation enters that audit, and validation RMSE is the declared selection loss. Panel B changes the rolling window or detrending order. Panel C changes the scale grid under a 252-day window and first-order detrending. Panel D uses the development-sample thresholds of 0.243 and 0.430. Panel E changes only the MFDFA feature group. Panel F changes the symmetric moment range and recomputes all MFDFA quantities. Panel G compares equal-frequency routing schemes with a global Extra Trees model containing the same market and MFDFA predictors. The moment-range and partition experiments are reviewer-requested diagnostics. They do not redefine the baseline model. All comparisons retain the same dates, lag structure, learner protocol, and forecasting target. Percentage gains are calculated from unrounded losses. Positive values favor the candidate model.

Table 8 provides a deliberately demanding assessment of the MFDFA design. In the pre-test validation audit, the 252-day window produces the largest RMSE gain, at 0.261%. The corresponding gains are 0.136% for 126 days and −1.351% for 504 days. The validation MAE results are less favorable. None of the three windows improves that loss. Thus, the 252-day specification is supported by the declared validation RMSE criterion, but it is not uniformly superior across losses. The locked test produces a clearer ranking. The 252-day specification lowers RMSE by 1.627% and MAE by 2.245%. The gains fall to 0.340% and 0.590% with the 126-day window. They decline further to 0.300% and 0.060% with the 504-day window. Higher-order detrending preserves positive gains but reduces their size. These results support the proposed bias–variance tradeoff. They do not establish 252 days as a universally optimal horizon. Scale-grid sensitivity leads to the same qualified conclusion. Both alternative grids retain positive locked-test gains, ranging from 0.213% to 1.494% for RMSE and from 1.323% to 1.727% for MAE. The baseline grid produces the strongest test gains and is the only grid with a positive validation RMSE gain. Its validation

MAE gain remains negative. The direction of the locked-test result is therefore stable, but its magnitude depends on the scale design. Panel D rejects an interpretation of the terciles as ordered risk levels. Mean next-day absolute returns are 0.424, 0.404, and 0.407 in the low-, medium-, and high-$\Delta\alpha$ states. A Kruskal–Wallis test does not reject equality of the three unconditional distributions ($p = 0.578$). A median-centered Levene test also finds no variance separation ($p = 0.942$). The states should therefore be understood as regions of the predictor space rather than monotonic levels of expected risk. The relevant test concerns the conditional forecasting relationship. Relative to the global MFDFA model, state-specific routing lowers RMSE within each state by 0.418–1.557% and MAE by 0.916–1.249%. Its aggregate incremental effect is significant under absolute loss ($p = 0.026$) but not under squared loss ($p = 0.099$). This evidence supports a modest and loss-dependent change in the conditional response surface. Panel F shows that the locked-test result is not tied to $q_{max} = 4$. Across $q_{max} \in \{2, 3, 4, 5\}$, the RMSE gains range from 1.370% to 1.786%, while the MAE gains range from 2.142% to 2.245%. Every locked-test gain remains positive. The validation results are less uniform. This difference confirms that the sensitivity exercise should not be used to select a new specification after the test. Panel G provides a separate assessment of the routing rule. Two-, three-, and four-state routing improves both locked-test losses relative to the global MFDFA model. Five-state routing improves MAE but reduces RMSE by 0.077%. The three-state specification is the only partition that improves the declared validation RMSE and both locked-test losses. Terciles are therefore a parsimonious routing choice rather than uniquely identified economic boundaries. Table 9 separates three possible sources of the forecast gain. Panel A tests whether performance improves merely because the augmented model receives more predictors. It also decomposes feature augmentation from routing. Panel B asks whether $\Delta\alpha_t$ can be replaced by conventional market-state gates. Panel C tests whether any MFDFA block with the same dimension and internal dependence can reproduce the aligned forecast gain.

Table 9. Competing-explanations diagnostics.

Panel A. Matched Feature Capacity and 2×2 Feature–Routing Decomposition

Model Design	RMSE	MAE	RMSE Gain Relative to Market-Only ET (%)	MAE Gain Relative to Market-Only ET (%)
Global market-only ET	0.373	0.277		
Global market + seven non-MFDFA summaries	0.393	0.305	−5.476	−9.748
Global market + seven MFDFA variables	0.370	0.274	0.810	1.176
$\Delta\alpha_t$-routed market-input ET	0.368	0.274	1.211	1.309
MFRA-ET	0.367	0.271	1.627	2.245

Panel B. Alternative Routing Variables with Identical Market-Plus-MFDFA Inputs

Routing Variable	RMSE	MAE	RMSE Gain Relative to Global MFDFA ET (%)	MAE Gain Relative to Global MFDFA ET (%)	DM pp, Squared Loss	DM pp, Absolute Loss
$\Delta\alpha_t$	0.367	0.271	0.824	1.082	—	—
20-day historical volatility	0.369	0.274	0.137	0.157	0.173	0.096
60-day historical volatility	0.371	0.277	−0.452	−1.038	0.010	0.001
20-day moving-average displacement	0.368	0.273	0.414	0.565	0.361	0.342
252-day excess kurtosis	0.372	0.278	−0.616	−1.203	0.016	0.002
First market-feature principal component	0.370	0.274	−0.166	0.077	0.054	0.094

Table 9. *Cont.*

Displacement Statistic	Panel C. Joint Temporal Displacement of the Global MFDFA Feature Block	
	RMSE Gain Relative to Market-Only ET (%)	MAE Gain Relative to Market-Only ET (%)
Correctly aligned MFDFA block	0.836	1.452
Mean across 99 displaced blocks	0.090	−0.421
Median across 99 displaced blocks	0.097	−0.463
5th percentile	−0.458	−1.560
95th percentile	0.670	0.624
Maximum	0.860	1.086
Plus-one upper-tail pp-value	0.020	0.010

Note: In Panel A, the matched non-MFDFA block comprises rolling skewness, excess kurtosis, the 5th and 95th return percentiles, interquartile range, maximum absolute return, and the share of positive returns. Each variable enters at lags 0–4. The block therefore contains 35 columns, the same number as the MFDFA block. In the routed market-input model, $\Delta\alpha_t$ is used only as the state gate and is excluded from the local predictor matrices. Panel B holds the complete predictor set, learner specification, estimation sample, forecast dates, and tercile construction rule constant. The reported Diebold–Mariano *p*-values compare $\Delta\alpha_t$ routing with each alternative route. Panel C uses one global Extra Trees learner and a common development sample of 4264 observations. All 35 MFDFA columns are delayed jointly. The 99 displacements span 60–750 trading days. Percentage gains are calculated from unrounded losses. Positive values favor the candidate model.

The capacity-matched comparison does not support the view that performance improves simply because 35 columns are added to Extra Trees. The non-MFDFA block raises RMSE to 0.393 and MAE to 0.305. Its corresponding gains are −5.476% and −9.748%. By contrast, the same learner with the 35 MFDFA columns lowers the losses to 0.370 and 0.274, yielding gains of 0.810% and 1.176%. The feature ablations in Table 8 provide consistent evidence. $\Delta\alpha_t$ alone improves both losses, while the three-variable dispersion block performs slightly better than the complete seven-variable block. Forecast gain is therefore not monotonic in the number of added MFDFA variables. The factorial comparison shows that feature augmentation and routing are distinct, but complementary, channels. Without routing, MFDFA inputs reduce RMSE by 0.810% and MAE by 1.176%. When $\Delta\alpha_t$ is used only as a gate and the local learners receive market inputs alone, routing reduces the losses by 1.211% and 1.309%. Adding the MFDFA inputs within the fixed $\Delta\alpha_t$ route produces further reductions of 0.422% and 0.949%. Conversely, introducing $\Delta\alpha_t$ routing while holding the complete MFDFA input set fixed reduces RMSE by 0.824% and MAE by 1.082%. Thus, neither channel alone accounts for the full MFRA-ET gain. Alternative gates produce a more qualified result. The $\Delta\alpha_t$ route has the lowest point RMSE and MAE among all examined routing variables. It significantly outperforms 60-day historical-volatility routing under squared and absolute loss ($p = 0.010$) and $p = 0.001$). It also outperforms the rolling-kurtosis route ($p = 0.016$ and $p = 0.002$). However, the differences relative to 20-day historical volatility, 20-day moving-average displacement, and the market principal component do not reach the 5% significance level. Generic state adaptation can therefore explain part of the benefit. The evidence does not identify $\Delta\alpha_t$ as a unique or causal state variable. The displacement experiment addresses whether feature dimension and persistent internal organization are sufficient to generate the global MFDFA gain. On the common sample, the aligned block improves RMSE by 0.836% and MAE by 1.452%. The 99 delayed blocks have mean gains of only 0.090% and −0.421%. The aligned MAE gain exceeds every displaced value. One displaced block exceeds the aligned RMSE gain. The corresponding plus-one upper-tail probabilities are 0.020 and 0.010. Hence, the observed result depends

on contemporaneous alignment between the estimated scaling state and the forecast origin. It is not reproduced by an arbitrary MFDFA block of equal dimension. Taken together, the diagnostics reject the simplest feature-count explanation and show that the forecast gain is not generated solely by routing. They also set a clear boundary on interpretation. The evidence supports time-aligned information in MFDFA-derived state variables within this forecasting architecture. It does not prove that multifractality is the sole economic source of predictability. Data-snooping-adjusted inference and risk-band evaluation are reported in Table 10.

Table 10. Data-snooping-adjusted inference and quantile-calibrated prediction-band evaluation.

Panel A. Studentized Superior Predictive Ability Tests			
Loss Function		**Maximum Statistic**	**20-Day Block *p*-Value**
Absolute loss		4.559	<0.001
Squared loss		4.239	<0.001
Panel B. Validation-Quantile-Calibrated 95% and 99% Prediction Bands			
Nominal Band	**Model**	**Coverage (%)**	**Mean Width**	**Interval Score**	**Mean Tail-Breach Magnitude**
95%	Market-only ET	94.00	2.187	2.938	0.312
95%	MFRA-ET	93.76	2.131	2.874	0.298
99%	Market-only ET	98.40	3.127	4.019	0.279
99%	MFRA-ET	98.24	3.043	3.789	0.212

Note: Market-only Extra Trees is the benchmark. The candidate universe includes all prespecified feature groups, global MFDFA-augmented Extra Trees, and MFRA-ET. Inference uses 5000 circular moving-block replications. The *p*-values remain at or below 0.001 with 5-day and 60-day blocks. The prediction-band endpoints are generated from the validation quantiles defined in Equation (17). Calibration multipliers are estimated only from 2017 to 2019 and remain fixed throughout 2020–2024. Mean tail-breach magnitude is the conditional average distance between a realized signed return and the nearest breached endpoint, expressed in percentage points. Lower mean width, interval score, and tail-breach magnitude are preferred, while coverage should be assessed relative to the nominal level. These bands describe daily reference-rate uncertainty and should not be interpreted as portfolio VaR or Expected Shortfall.

Table 10 indicates that the model-universe test rejects the null that none of the candidate multifractal specifications improves on market-only Extra Trees. This conclusion is unchanged under alternative block lengths. The evidence therefore survives an adjustment for searching across the prespecified feature universe. Panel B provides a tail-oriented complement to the point-forecast results. At the 95% level, MFRA-ET attains coverage of 93.760%, compared with 94.000% for market-only Extra Trees. Its mean width declines from 2.187 to 2.131, while the interval score falls from 2.938 to 2.874. The corresponding interval-score reduction is 2.160%. Mean tail-breach magnitude also decreases from 0.312 to 0.298 percentage points. At the 99% level, MFRA-ET produces coverage of 98.240%, a mean width of 3.043, and an interval score of 3.789. The corresponding market-only values are 98.400%, 3.127, and 4.019. Thus, MFRA-ET lowers the 99% interval score by 5.730%. Mean tail-breach magnitude falls more substantially, from 0.279 to 0.212 percentage points. The improvement is therefore stronger in the extreme tail than in the central risk region. Nevertheless, coverage remains below the nominal level for both models and is slightly lower for MFRA-ET. The evidence supports sharper daily prediction bands and less severe tail breaches. It does not establish complete interval calibration or a stand-alone VaR system.

4.6. Explanatory-Power and Temporal-Alignment Diagnostics

The level of explanatory power must be stated plainly. Using the unrounded locked-sample forecasts, MFRA-ET produces an R^2 of 0.045 and a forecast–realization correlation

of 0.218. It therefore leaves 95.52% of daily absolute-return variation unexplained. The two measures are internally consistent: the squared correlation is 0.0476, close to the loss-based R^2. Both indicate a small common component rather than broad explanatory power. The matched market-only Extra Trees model produces an R^2 of 0.013 and a correlation of 0.157. MFRA-ET therefore increases the two measures by 0.032 and 0.061, respectively. Table 11 evaluates the uncertainty surrounding both the absolute statistics and these paired increments.

Table 11. Explanatory-power and temporal-alignment diagnostics.

Panel A. Locked-Sample Estimates and Paired Circular Moving-Block 95% Confidence Intervals				
Statistic	**Point Estimate**	**5-Day Block**	**20-Day Block**	**60-Day Block**
MFRA-ET R^2	0.045	[0.00935, 0.076]	[0.001, 0.079]	[−0.006, 0.083]
MFRA-ET correlation	0.21825	[0.149, 0.283]	[0.127, 0.290]	[0.108, 0.295]
ΔR^2 relative to market-only ET	0.035	[0.012, 0.052]	[0.011, 0.051]	[0.011, 0.050]
Δ correlation relative to market-only ET	0.061	[0.019, 0.103]	[0.018, 0.101]	[0.018, 0.098]

Panel B. Temporal-Alignment and Regression Diagnostics			
Diagnostic	**Estimate**	**Null or Placebo Reference**	***p*-Value**
Correctly aligned R^2	0.045	Shift-placebo 95% range: [−0.121, −0.016]	<0.001
Correctly aligned correlation	0.218	Shift-placebo 95% range: [−0.107, 0.099]	<0.001
Mincer–Zarnowitz slope, b	0.859 (0.146)	H_0: b = 0	<0.001
Encompassing increment, c	0.999 (0.244)	H_0: c = 0	<0.001

Panel C. Pooled Annual Rolling-Origin Forecasts		
Model	**R^2**	**Correlation**
Market-only Extra Trees	0.015	0.163
MFRA-ET	0.044	0.215
MFRA-ET minus market-only ET	0.029	0.052

Note: Panel A reports percentile intervals from 10,000 paired circular moving-block bootstrap replications. The 20-day block is the primary specification. Panel B uses 1130 admissible circular shifts that are at least 60 trading days from the correct alignment. Its exact empirical *p*-value is 0.000884 for both alignment statistics. Regression standard errors, shown in parentheses, are heteroskedasticity- and autocorrelation-consistent with five lags. Panel C pools forecasts from the five annual rolling-origin evaluations. All calculations use the same 1249 test observations and fixed forecast-generating rules.

Under the primary 20-day block, the 95% interval for the MFRA-ET R^2 is [0.001, 0.079]. This interval excludes zero only narrowly. Under the more conservative 60-day block, it becomes [−0.006, 0.083]. Thus, the absolute R^2 claim is sensitive to the assumed dependence horizon. The paired increments are more stable. The confidence intervals for ΔR^2 and Δ correlation remain strictly positive under all three block lengths. This distinction is central. The data support incremental explanatory content relative to a matched benchmark, but they do not support high stand-alone explanatory power. The temporal-shift results provide a separate check. The correctly aligned R^2 and correlation lie above the entire central 95% range of their shifted counterparts. Both empirical *p*-values are below 0.001. The Mincer–Zarnowitz slope is 0.859, with a HAC standard error of 0.146. The null of zero alignment is rejected. In the encompassing regression, the coefficient on the MFRA-ET increment is 0.999, with a HAC standard error of 0.244. This coefficient is also significant at $p < 0.001$. Hence, the increment is not absorbed by the market-only forecast. The pooled annual rolling-origin results tell the same relative story, but individual years remain heterogeneous. MFRA-ET produces negative annual R^2 values in 2021 (−0.028), 2023 (−0.007), and 2024 (−0.046). Its annual loss performance is more stable than its annual R^2. This evidence rules out a claim of strong or uniform explanatory power. Taken together, the diagnostics identify a weak, incremental, and time-localized signal. They reject a

characterization of the locked-sample result as mere random alignment, but they do not make the signal large.

4.7. Return-Direction Heterogeneity and Alternative-Target Evidence

Absolute return removes the sign by construction. This transformation is appropriate for a magnitude forecast, but an aggregate loss comparison may conceal directional heterogeneity. Panel A of Table 12 therefore evaluates the fixed locked-sample forecasts separately on won-depreciation and won-appreciation days.

Table 12. Directional and alternative-target robustness.

Panel A. Forecast Performance by Realized Return Direction

Realized Direction	N	Market-Only ET RMSE	MFRA-ET RMSE	RMSE Gain (%)	20-Day Block 95% CI	Market-Only ET MAE	MFRA-ET MAE	MAE Gain (%)	20-Day Block 95% CI
Won depreciation, $r_{t+1} > 0$	658	0.336	0.334	0.553	[−1.072, 2.235]	0.261	0.258	1.283	[−0.580, 3.200]
Won appreciation, $r_{t+1} < 0$	578	0.410	0.400	2.416	[1.053, 3.564]	0.293	0.284	3.164	[1.166, 5.161]

Panel B. Squared-Return Target, $\beta_{t+1}=r^2_{t+1}$

Model	RMSE	MAE
Market-only Extra Trees	0.660	0.342
MFRA-ET	0.652	0.329
Relative gain (%)	1.208	3.894
20-day block 95% CI	[0.046, 2.666]	[1.294, 6.535]

Note: Panel A evaluates the original absolute-return forecasts. Positive USD/KRW returns denote won depreciation, while negative returns denote won appreciation. Thirteen zero-return observations are excluded. Panel B refits the same fixed specifications with squared daily return as the target. Confidence intervals use 10,000 paired circular moving-block bootstrap replications. The 20-day block is primary. With 5-day and 60-day blocks, the depreciation-day intervals continue to include zero, whereas the appreciation-day intervals remain positive. For the squared-return target, the 60-day RMSE interval includes zero, but the corresponding MAE interval remains positive. Squared daily return is not interpreted as high-frequency realized volatility.

Panel A reveals that the aggregate gain is not directionally uniform. On won-depreciation days, MFRA-ET lowers RMSE by 0.553% and MAE by 1.283%. Both confidence intervals include zero. These observations do not establish a reliable improvement. On won-appreciation days, RMSE and MAE decline by 2.416% and 3.164%. Their confidence intervals remain positive. The primary forecasting gain is therefore concentrated more strongly in appreciation episodes. This asymmetry does not turn MFRA-ET into a directional model. The forecast remains nonnegative and contains no sign prediction. The split only identifies where its magnitude forecasts are more accurate. The result also prevents the absolute-return target from concealing a direction-dependent performance pattern. Panel B provides a separate target-form test. With r^2_{t+1} as the dependent variable, MFRA-ET lowers RMSE by 1.208% and MAE by 3.894%. Both 20-day block intervals are positive. The incremental ranking is therefore not confined to the absolute-value transformation. The evidence remains qualified because squared daily return is itself noisy and the longer-block RMSE interval includes zero. This exercise supports target-form robustness within daily data. It does not replicate a high-frequency realized-volatility experiment.

4.8. Cross-Currency Transferability

The preceding results establish temporal evidence within USD/KRW. They do not establish that the frozen forecasting design transfers to other currencies. Table 13 therefore applies the complete USD/KRW protocol to eight additional H.10 exchange rates without pair-specific retuning. Positive gains denote lower loss than the corresponding market-only Extra Trees model.

Table 13. No-retuning cross-currency transferability test, 2020–2024.

Currency Pair	Global ET + MFDFA RMSE Gain	Global ET + MFDFA MAE Gain	MFRA-ET RMSE Gain	MFRA-ET MAE Gain
USD/KRW	0.810%	1.176%	1.627%	2.245%
USD/JPY	0.390%	0.141%	−0.430%	−0.619%
USD/CNY	−0.335%	1.148%	−1.138%	−0.056%
USD/INR	−0.519%	−2.768%	−0.361%	−2.150%
USD/SGD	0.098%	−0.358%	0.378%	−0.178%
USD/THB	0.348%	0.118%	1.552%	2.209%
USD/MYR	−0.158%	−0.106%	−0.692%	0.529%
EUR/USD	−0.617%	−2.461%	−1.930%	−4.763%
GBP/USD	1.390%	1.619%	0.139%	0.251%

Note: Global ET + MFDFA denotes Extra Trees containing all market and MFDFA variables without routing. MFRA-ET uses pair-specific $\Delta\alpha$ terciles estimated exclusively from the pre-2020 development sample. All other model choices are frozen at their USD/KRW values. Each test contains 1249 forecasts. Positive gains favor the candidate model. Pairwise inference uses five-lag Newey–West Diebold–Mariano tests. Holm adjustment is applied across the eight external currencies within each model-loss family.

Table 13 provides no basis for a universal cross-currency claim. Global MFDFA augmentation improves both losses for USD/JPY, USD/THB, and GBP/USD. It produces mixed gains for USD/CNY and USD/SGD, while failing to improve USD/INR, USD/MYR, or EUR/USD consistently. Among the external currencies, the routed MFRA-ET specification improves both losses only for USD/THB and GBP/USD. The GBP/USD gains are close to zero. The inferential evidence is equally restrained. For USD/THB, the routed-model improvement is significant under absolute loss ($p = 0.043$) but not under squared loss ($p = 0.140$). The positive results for the other external currencies are not significant at the 5% level. No positive external result survives Holm adjustment. Across the eight external pairs, the median global-MFDFA gains are −0.030% for RMSE and 0.006% for MAE. The corresponding MFRA-ET medians are −0.395% and −0.117%. This frozen-specification exercise is more demanding than pair-specific model selection. Its negative results do not prove that MFDFA variables cannot forecast another currency after separate tuning. They show that the USD/KRW design does not travel mechanically. External validity is therefore partial and currency-dependent. The strongest transported evidence arises for USD/THB, but one additional positive case cannot support a general conclusion about Asian currencies.

4.9. Discussion

The results carry several broader implications. First, near-random H(2) does not imply that USD/KRW returns lack temporal organization. The declining H(q) function, broad singularity spectrum, and rolling variation in $\Delta\alpha$ reveal heterogeneity across fluctuation sizes. Small changes display different persistence from large movements. This finding supports the multifractal view that one global exponent cannot describe a financial market with heterogeneous horizons (Bolgorian and Gharli [63]). The expanded surrogate evidence sharpens this interpretation without assigning economic meaning to the surrogate *p*-values. Shuffled series retain substantial spectrum width, which indicates that

the heavy-tailed marginal distribution contributes to the observed multifractality. The empirical ΔH and $\Delta\alpha$ nevertheless exceed the shuffled distributions with $p = 0.001$ and $p = 0.002$. The phase-randomized values of $p_{MC} = 0.001$ correspond to zero exceedances among 999 realizations. The IAAFT values of $p_{MC} = 0.005$ likewise correspond to zero exceedances among 199 realizations. By contrast, the shuffled $\Delta\alpha$ comparison has one exceedance and $p_{MC} = 0.002$. The floor values are therefore reported as properties of the finite Monte Carlo design. They are not interpreted as exact population probabilities or evidence of economic importance. Thus, neither the empirical marginal distribution nor its combination with approximate linear spectral dependence reproduces the full observed width. This conclusion concerns statistical scaling structure only. Predictive relevance must be established separately through temporal out-of-sample evidence. Second, unconditional multifractality does not imply strong one-day predictability. Predictive content must also be distinguished from broad explanatory power. The MFRA-ET R^2 of 0.045 and correlation of 0.218 are low. The model leaves about 95.5% of daily variation unexplained. The new diagnostics nevertheless show that its small incremental contribution is not reproduced when forecast dates are shifted. The paired increments in R^2 and correlation remain positive under block lengths of 5, 20, and 60 trading days. The forecast-encompassing regression also retains a significant MFRA-ET increment. The appropriate interpretation is therefore weak and time-localized predictability, rather than either random alignment or high explanatory power. The evidence is also architecture-dependent. MFDFA variables improve Extra Trees and random forest, and the gains remain positive under all 30 examined seeds for both learners. They do not improve linear regression or histogram gradient boosting. This pattern is consistent with information that enters through nonlinear interactions and threshold responses, although it does not prove that every nonlinear architecture can exploit the same features. The defensible conclusion is therefore limited to stable incremental information within the bagged-tree class used here (Wang and Lee [64]; Gao et al. [65]; Chi et al. [66]; Leventides et al. [67]).

Third, the attribution diagnostics narrow the competing explanations without eliminating them by assumption. An equal-sized block of non-MFDFA summaries fails to reproduce the global MFDFA gain. The 2×2 comparison shows that MFDFA inputs and $\Delta\alpha_t$-based routing each lower both locked-sample losses. Jointly delayed MFDFA blocks also perform materially worse than the correctly aligned block. These results argue against a mechanical feature-count explanation and against arbitrary temporal placement. The alternative-gate results are less categorical. Short-horizon volatility and moving-average displacement also benefit from local routing, and their losses are not significantly different from those of the $\Delta\alpha_t$ route at the 5% level. We therefore interpret $\Delta\alpha_t$ as an informative MFDFA-derived state descriptor, not as a uniquely identified causal source of exchange-rate risk. This distinction has been carried through the interpretation of the empirical results. The expanding-window audit also separates threshold drift from changing state occupancy. Although Figure 3 shows pronounced movements in $\Delta\alpha_t$, the annual lower and upper terciles remain within the narrow ranges of 0.241–0.248 and 0.422–0.434. Only 1.761% of test observations are reassigned relative to the fixed pre-2020 cutoffs. At the same time, annual state shares vary substantially. The system therefore captures structural variation mainly through changing occupancy of the predictor-space regions rather than large movements in the boundaries themselves. The positive pooled gains under annual threshold updating reduce concern that stale cutoffs generate the result. They do not make the terciles structural economic boundaries or establish a permanent regime process. Fourth, the revised permutation analysis rejects a strict feature hierarchy. Under 20-day block permutation, ΔH has the largest point importance, while $\Delta\alpha$ ranks second. The uncertainty intervals overlap. $\Delta\alpha$ ranks first in only 12.040% of moving-block replications, although it enters the

top three in 65.520%. Its pairwise differences from ΔH, the 20-day mean absolute return, and upside RV20 are not significant after Holm adjustment. The appropriate interpretation is therefore set-valued rather than ordinal. Cross-scale dispersion and recent volatility belong to an upper importance tier, but the data do not identify a unique leading predictor. This diagnostic concerns reliance within one fitted learner. It is not used to establish causal importance or to select the forecasting model.

The Korean setting remains important for interpreting the USD/KRW evidence, but it cannot establish a universal moderator. The won is exposed to global dollar cycles, cross-border portfolio flows, geopolitical risk, and domestic monetary conditions. These forces operate at different horizons and may alter the conditional forecasting relationship. Their relevance provides economic context rather than causal identification. The larger 2022–2024 point gains coincide with higher volatility, Federal Reserve tightening, dollar strength, geopolitical uncertainty, and repeated pressure on the won. Direct scaling shows that the RMSE pattern is not explained by a purely multiplicative volatility effect, although the corresponding MAE evidence remains nonsignificant. More importantly, equal weighting across ex ante volatility states reduces the later-minus-earlier RMSE difference to 0.282 percentage points ($p = 0.907$) and the MAE difference to 1.850 percentage points ($p = 0.410$). Conditional loss-differential regressions also produce nonsignificant later-period coefficients under absolute and squared loss ($p = 0.516$ and $p = 0.139$). The later-period pattern is therefore neither a mechanical scale artifact nor robust evidence of a distinct structural regime. It cannot establish that monetary tightening or market stress increases MFDFA predictability. The cross-currency test reinforces this boundary. The frozen specification performs positively for some currencies but negatively for others. It does not support a common exchange-rate effect. Within USD/KRW, the mechanism proposed in Section 2.2 remains conditional. The mean next-day absolute returns across the three $\Delta\alpha$ states are not monotonic. MFRA-ET has lower point losses than global MFDFA-augmented Extra Trees under both loss functions. However, the routing contrasts do not survive family-wise multiplicity adjustment. The Holm-adjusted p-values are 0.099 under squared loss and 0.077 under absolute loss. The corresponding Bonferroni-adjusted values are 0.989 and 0.257. The point estimates are consistent with a possible currency-specific state-dependent response surface, but they do not establish an independent routing effect. Nor do they identify $\Delta\alpha$ as a causal determinant, a sufficient statistic for risk, or a universally transferable predictor.

The target distinction is central to this interpretation. Absolute return measures the magnitude of the change between consecutive daily reference observations. It does not reconstruct the exchange-rate path within the day. Nor does it predict whether the won will appreciate or depreciate. The directional decomposition shows that the improvement is concentrated more strongly on won-appreciation days. The squared-return experiment preserves the direction of the aggregate gain. These findings support a daily conditional-magnitude interpretation, but not an intraday volatility or directional forecasting interpretation. Within the USD/KRW setting examined here, financial institutions could evaluate $\Delta\alpha$ as a supplementary descriptor of the current cross-scale state. Its possible value is limited to complementing the interpretation of conventional market indicators for this currency pair. Spectrum width alone does not reveal the direction or expected size of the next exchange-rate movement. A high-$\Delta\alpha$ observation should therefore prompt conditional reassessment rather than an automatic high-risk classification. The validation-quantile-calibrated bands translate each magnitude forecast into an ex ante threshold for an unusually large daily reference-rate movement. MFRA-ET produces narrower bands, lower interval scores, and smaller conditional breaches than market-only Extra Trees. The difference is most visible at the 99% level, where mean tail-breach magnitude declines from 0.279 to 0.212 percentage

points. For the analyzed USD/KRW pair, these thresholds could be evaluated as an illustrative supplementary input to low-frequency exposure screening or more detailed stress analysis. One possible application for the Bank of Korea would be experimental shadow monitoring of daily USD/KRW conditions, subject to independent validation against simpler benchmarks. A flagged observation could prompt analyst review, but it should not trigger a mechanical intervention or be interpreted as an autonomous policy signal. This pair-specific suggestion should not be extended to other currencies or institutional settings without separate estimation, prospective validation, and an explicit assessment of implementation costs. Shadow deployment would also permit continuous comparison with the market-only Extra Trees benchmark. Such an implementation requires monitoring of the MFDFA feature distribution, state prevalence, threshold stability, and prediction-band coverage. If the incremental accuracy weakens or the state distribution drifts materially, the simpler market-only specification provides a transparent fallback. Their coverage nevertheless remains below nominal. The results should therefore be interpreted as comparative evidence of improved tail monitoring, not as fully calibrated VaR or Expected Shortfall. MFRA-ET also cannot generate a directional position because its forecast contains no sign. It cannot replace intraday volatility systems, liquidity monitoring, exposure limits, or stress tests. The evidence supports modest daily-displacement predictability rather than intraday risk measurement or transaction-cost-adjusted profitability.

Several limitations qualify the findings. Absolute return is an observable daily displacement measure, not an efficient estimator of integrated variance. The H.10 series contains one reference observation per day. It cannot capture temporary intraday jumps, bid–ask dynamics, market microstructure noise, or within-day liquidity stress. A study of integrated variance would therefore require consistently sampled high-frequency quotes or transactions and explicit treatment of the 24 h foreign-exchange trading day. The directional results also show that forecast improvement is not uniform. It is statistically distinguishable on won-appreciation days but not on won-depreciation days. The positive squared-return result reduces concern about dependence on one transformation, but it does not replace a realized-volatility experiment. The no-retuning currency panel broadens the external diagnostic, but it does not establish general validity. It uses one forecast period and a specification selected for USD/KRW. Its mixed results show that transferability is conditional on the currency. Pair-specific tuning could produce different forecast rankings, but such tuning would constitute a separate forecasting experiment. The subperiod windows also remain descriptive. Feature importance is model-specific, rank-uncertain, and noncausal. The dependence-aware analysis does not identify $\Delta\alpha$ as the uniquely most important predictor. Future research should use broader currency panels, repeated forecast vintages, panel-level inference, and currency-specific validation. Within USD/KRW, multifractal variables provide information beyond a single Hurst exponent and conventional market inputs. The external results show that this conclusion should not be extended automatically to other currencies.

5. Conclusions

This study examined whether time-varying multifractal states improve machine-learning forecasts of the next observed USD/KRW return magnitude. The target is a daily reference-rate displacement measure. It is not an estimator of intraday realized volatility and contains no directional forecast. The full-sample H(2) estimate is close to 0.5, suggesting weak average second-order persistence. This result does not imply an absence of temporal structure. Both ΔH and $\Delta\alpha$ reveal substantial heterogeneity across fluctuation sizes. Expanded shuffled, phase-randomized, and IAAFT diagnostics show that the observed widths are not reproduced by the examined empirical null distribu-

tions. Zero-exceedance cases are reported at their finite Monte Carlo floors of 0.001 or 0.005 and are not interpreted as zero tail probabilities. These diagnostics establish statistical scaling evidence under the specified nulls, not economic value, predictive usefulness, or causal identification The forecasting result remains positive across alternative windows, detrending orders, scale grids, and symmetric moment ranges, although its magnitude varies. The capacity-matched non-MFDFA block does not reproduce the gain. The 2×2 decomposition produces positive point-loss contributions from both MFDFA inputs and $\Delta\alpha_t$-based routing. This decomposition is descriptive rather than confirmatory. In Table 5, the direct comparisons between MFRA-ET and global MFDFA-augmented Extra Trees do not survive Holm or Bonferroni adjustment. The evidence therefore does not establish a statistically independent routing advantage. On a common sample, the aligned global MFDFA block outperforms 99 delayed blocks, with upper-tail p-values of 0.020 for RMSE and 0.010 for MAE. Some conventional state gates also improve on the global model, and their differences from $\Delta\alpha_t$ routing are not consistently significant. The supported conclusion concerns time-aligned information in the broader MFDFA feature architecture, rather than the superiority of one routing rule. MFDFA variables improve Extra Trees and random forest across all 30 examined seeds, but they do not improve linear regression or histogram gradient boosting. The evidence therefore supports time-aligned, architecture-dependent information in MFDFA-derived variables. It does not establish multifractality as the sole or uniquely causal source of the forecast gain. Dependence-aware grouped permutation importance also provides no evidence of a uniquely dominant input. $\Delta\alpha$ ranks first in 12.040% of moving-block replications and enters the top three in 65.520%. None of its 16 pairwise importance differences is significant after Holm adjustment. The predictive claim therefore rests on matched out-of-sample loss reductions, not on an unstable ordinal importance ranking. MFRA-ET achieves the lowest RMSE and MAE in the locked 2020–2024 test period. Its R^2 of 0.045 and correlation of 0.218 indicate limited total explanatory power. Relative to market-only Extra Trees, however, R^2 increases by 0.032 and correlation increases by 0.0612. Under the primary 20-day block bootstrap, the corresponding 95% intervals are [0.011, 0.051] and [0.018, 0.101]. Both increments remain positive with 5-day and 60-day blocks. Temporal-shift placebos reject accidental calendar-time alignment at $p < 0.001$. The Mincer–Zarnowitz and forecast-encompassing regressions also reject zero incremental information. These results do not imply strong predictability. The 60-day interval for the absolute R^2 includes zero, and annual R^2 is negative in three evaluation years. Annual expanding-window estimation, which recomputes both $\Delta\alpha$ thresholds before each evaluation year, lowers USD/KRW RMSE in all five years and MAE in four. The annual cutoffs remain close to their pre-2020 values, and only 1.761% of test observations change state. The pooled gains remain positive under dependence-robust inference. Higher-order detrending preserves the loss reduction, whereas the 504-day window largely removes it. The raw 2022–2024 point gains exceed those from 2020 to 2021, but their moving-block differences are not significant. Scale-normalized RMSE retains a later-period difference, while normalized MAE does not. Once the preforecast volatility-state distribution is equalized, the RMSE and MAE differences fall to 0.282 and 1.850 percentage points, with $p = 0.907$ and $p = 0.410$. Conditional later-period coefficients are also nonsignificant. The calendar split therefore provides no evidence that tightening or structural change causes stronger MFDFA predictability. The no-retuning currency panel likewise produces mixed results, and no positive external gain survives multiplicity adjustment. The evidence identifies a weak, incremental, and USD/KRW-specific forecasting contribution. It does not establish broad explanatory power, crisis-period causality, universal cross-currency transferability, or economically large trading profits.

For the USD/KRW pair examined here, the results support only an illustrative monitoring implication, subject to an explicit implementation trade-off. In this setting, $\Delta\alpha$ and ΔH could be evaluated as supplementary state descriptors alongside recent volatility and persistence measures. Their inclusion nevertheless requires a daily 252-observation MFDFA update and the maintenance of three local ensembles. This additional burden must be weighed against an improvement of approximately 0.006 percentage points, or 0.600 basis points, in the principal forecast-error measures. The routing increment itself does not survive family-wise multiplicity adjustment. MFRA-ET is therefore most defensible as a shadow or secondary monitoring model within an existing daily analytics system. It should not replace intraday surveillance, liquidity monitoring, stress testing, or a simpler benchmark model. The nonmonotonic state means also confirm that spectrum width alone does not determine next-day return magnitude. Any operational use requires continuing validation of feature stability, state prevalence, forecast accuracy, and prediction-band coverage. These variables do not identify the cause of exchange-rate pressure, prescribe a hedge, or justify an automatic policy response. Several limitations define the scope of the conclusions. Absolute return is transparent, but it is less informative than intraday realized volatility and cannot measure within-day liquidity stress. The external currency panel is a frozen-specification stress test rather than a set of pair-specific forecasting studies. Its mixed results prevent a general conclusion about major or Asian currencies. The gain depends on the MFDFA measurement horizon, although it remains positive across the examined detrending orders, scale grids, and moment ranges. It becomes negligible under the 504-day window. Alternative quantile partitions confirm that the routing result is not unique to terciles, while annual expanding-window estimates show that it is not dependent on stale pre-2020 cutoffs. Adding more states does not improve performance consistently. The thresholds should therefore be treated as parsimonious forecasting devices rather than structural economic boundaries. Their relative stability coexists with large changes in annual state prevalence. A discrete routing rule may also fail to represent gradual transitions between multifractal states. The validation-quantile-calibrated 95% and 99% prediction bands are narrower and produce lower interval scores and smaller tail breaches than the corresponding market-only bands. Their coverage nevertheless remains below nominal. They should be interpreted as daily reference-rate prediction bands rather than fully calibrated portfolio VaR or Expected Shortfall estimates. Most importantly, the model leaves most daily USD/KRW variation unexplained and supplies no directional signal. It should not be used as a stand-alone policy, hedging, or trading rule. Future research should combine broader currency panels, repeated forecast origins, pair-specific validation, panel-level predictive-ability tests, intraday data, and explicit hedging objectives with turnover and transaction costs.

Author Contributions: Conceptualization, D.W. and Y.H.; methodology, D.W. and Y.H.; software, D.W.; validation, D.W.; formal analysis, D.W.; investigation, D.W.; resources, D.W.; data curation, D.W.; writing—original draft preparation, D.W.; writing—review and editing, Y.H.; funding acquisition, D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the 2025 Annual Scientific Research Project of Kashgar University, grant number GCC2025SK-012; the 2026 Annual School-Level Project of Kashgar University, grant number 022026165; and the 2026 Annual Education and Teaching Reform Project of Kashgar University, grant number KD2026J012.

Data Availability Statement: All exchange-rate series are publicly available from the Federal Reserve H.10 release and the Federal Reserve Bank of St. Louis. The identifiers are DEXKOUS, DEXJPUS, DEXCHUS, DEXINUS, DEXSIUS, DEXTHUS, DEXMAUS, DEXUSEU, and DEXUSUK.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Grobys, K. A Multifractal Model of Asset (in) Variances. *J. Int. Financ. Mark. Inst. Money* **2023**, *85*, 101767. [CrossRef]
2. Caporin, M.; Rodríguez-Caballero, C.V.; Ruiz, E. The Factor Structure of Exchange Rates Volatility: Global and Intermittent Factors. *Empir. Econ.* **2024**, *67*, 31–45. [CrossRef]
3. Yu, X.; Li, Y.; Wang, X. RMB Exchange Rate Forecasting Using Machine Learning Methods: Can Multimodel Select Powerful Predictors? *J. Forecast.* **2024**, *43*, 644–660. [CrossRef]
4. Arslan, M.; Hunjra, A.I.; Ahmed, W.S.; Ben Zaied, Y. Forecasting Multi-Frequency Intraday Exchange Rates Using Deep Learning Models. *J. Forecast.* **2024**, *43*, 1338–1355. [CrossRef]
5. Lee, S.S.; Wang, M. Jumps and Post-FOMC Announcement Returns in Currency Markets. *Rev. Asset Pricing Stud.* **2025**, *15*, 247–287. [CrossRef]
6. Grobys, K. On Discontinuity and Persistence in the Returns on Financial Assets: New Simulation-Based Evidence Derived From a Multifractal Model of Asset Returns. *Comput. Econ.* **2026**, 1–48. [CrossRef]
7. Fathi, M.; Grobys, K.; Kolari, J.W. On the Realized Risk of Foreign Exchange Rates: A Fractal Perspective. *J. Risk Financ. Manag.* **2024**, *17*, 79. [CrossRef]
8. Ünal, B. Time-Varying Fractal Analysis of Exchange Rates. *Chaos Theory Appl.* **2023**, *5*, 242–255. [CrossRef]
9. Sandubete, J.E.; Beleña, L.; García-Villalobos, J.C. Testing the Efficient Market Hypothesis and the Model-Data Paradox of Chaos on Top Currencies from the Foreign Exchange Market (FOREX). *Mathematics* **2023**, *11*, 286. [CrossRef]
10. Saâdaoui, F. Segmented Multifractal Detrended Fluctuation Analysis in Key Economic Sectors. *Fluct. Noise Lett.* **2024**, *23*, 2440026. [CrossRef]
11. Ali, H.; Aftab, M.; Aslam, F.; Ferreira, P. Inner multifractal dynamics in the jumps of cryptocurrency and forex markets. *Fractal Fract.* **2024**, *8*, 571. [CrossRef]
12. Lee, M.; Cho, Y.; Ock, S.E.; Song, J.W. Analyzing Asymmetric Volatility and Multifractal Behavior in Cryptocurrencies Using Capital Asset Pricing Model Filter. *Fractal Fract.* **2023**, *7*, 85. [CrossRef]
13. Ke, J.; Duan, Y.; Xu, C.; Zhang, Y. Cross-Correlation Multifractal Analysis of Technological Innovation, Financial Market and Real Economy Indices. *Fractal Fract.* **2023**, *7*, 267. [CrossRef]
14. Mariani, M.C.; Kubin, W.; Asante, P.K.; Tweneboah, O.K. Volatility Analysis of Financial Time Series Using the Multifractal Conditional Diffusion Entropy Method. *Fractal Fract.* **2024**, *8*, 274. [CrossRef]
15. Yang, L.; Wang, H.; Gu, C.; Yang, H. Weighted signed networks reveal interactions between US foreign exchange rates. *Entropy* **2024**, *26*, 161. [CrossRef] [PubMed]
16. Morales-Zumaquero, A.; Sosvilla-Rivero, S. An Empirical Evaluation of Transitory and Permanent Components of Exchange Rate Volatility. *Appl. Econ.* **2026**, *58*, 2829–2844. [CrossRef]
17. Peng, L.; Pan, Z.; Liang, C.; Umar, M. Exchange rate volatility predictability: A new insight from climate policy uncertainty. *Econ. Anal. Policy* **2023**, *80*, 688–700. [CrossRef]
18. Aftab, M.; Mehmood, A. Does central bank transparency deter the exchange rate volatility? New evidence from Asian emerging markets. *J. Cent. Bank. Theory Pract.* **2023**, *12*, 133–163. [CrossRef]
19. Wang, X.; Wang, Z. Analysis of the Response of Exchange Rates to Specific FOMC Announcements Using High-Frequency Data. *Borsa Istanb. Rev.* **2022**, *22*, 133–144. [CrossRef]
20. Uzun, S.; Sensoy, A.; Nguyen, D.K. Jump Forecasting in Foreign Exchange Markets: A High-frequency Analysis. *J. Forecast.* **2023**, *42*, 578–624. [CrossRef]
21. Sattarhoff, C.; Lux, T. Forecasting the Variability of Stock Index Returns with the Multifractal Random Walk Model for Realized Volatilities. *Int. J. Forecast.* **2023**, *39*, 1678–1697. [CrossRef]
22. Ahmed, H.; Aslam, F.; Ferreira, P. Navigating Choppy Waters: Interplay between Financial Stress and Commodity Market Indices. *Fractal Fract.* **2024**, *8*, 96. [CrossRef]
23. Zhang, W.; Huang, Z. Nonlinear Dynamics of RMB Exchange Rate Volatility: A Multifractal Perspective Within the G-Expectation Framework. *Fractal Fract.* **2025**, *9*, 746. [CrossRef]
24. Alexandridis, A.K.; Panopoulou, E.; Souropanis, I. Forecasting Exchange Rate Volatility: An Amalgamation Approach. *J. Int. Financ. Mark. Inst. Money* **2024**, *97*, 102067. [CrossRef]
25. Tang, X.; Xie, Y. Exchange Rate Forecasting: A Deep Learning Framework Combining Adaptive Signal Decomposition and Dynamic Weight Optimization. *Int. J. Financ. Stud.* **2025**, *13*, 151. [CrossRef]
26. Ding, Z.; Granger, C.W.J.; Engle, R.F. A Long Memory Property of Stock Market Returns and a New Model. *J. Empir. Financ.* **1993**, *1*, 83–106. [CrossRef]
27. Granger, C.W.J.; Ding, Z. Some Properties of Absolute Return: An Alternative Measure of Risk. *Ann. Econ. Stat.* **1995**, *40*, 67–91. [CrossRef]
28. Forsberg, L.; Ghysels, E. Why Do Absolute Returns Predict Volatility So Well? *J. Financ. Econom.* **2007**, *5*, 31–67. [CrossRef]

29. Andersen, T.G.; Bollerslev, T.; Diebold, F.X.; Labys, P. Modeling and Forecasting Realized Volatility. *Econometrica* **2003**, *71*, 579–625. [CrossRef]
30. Christensen, K.; Siggaard, M.; Veliyev, B. A Machine Learning Approach to Volatility Forecasting. *J. Financ. Econom.* **2023**, *21*, 1680–1727. [CrossRef]
31. Hewamalage, H.; Ackermann, K.; Bergmeir, C. Forecast Evaluation for Data Scientists: Common Pitfalls and Best Practices. *Data Min. Knowl. Discov.* **2023**, *37*, 788–832. [CrossRef] [PubMed]
32. Pabuccu, H.; Barbu, A. Feature Selection with Annealing for Forecasting Financial Time Series. *Financ. Innov.* **2024**, *10*, 87. [CrossRef]
33. Pirayesh Neghab, D.; Cevik, M.; Wahab, M.I.M.; Basar, A. Explaining Exchange Rate Forecasts with Macroeconomic Fundamentals Using Interpretive Machine Learning. *Comput. Econ.* **2025**, *65*, 1857–1899. [CrossRef]
34. Mao, Y.; Chen, Z.; Liu, S.; Li, Y. Unveiling the Potential: Exploring the Predictability of Complex Exchange Rate Trends. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108112. [CrossRef]
35. Barthélémy, S.; Gautier, V.; Rondeau, F. Early Warning System for Currency Crises Using Long Short-term Memory and Gated Recurrent Unit Neural Networks. *J. Forecast.* **2024**, *43*, 1235–1262. [CrossRef]
36. Wang, J.; Dong, Y.; Liu, J. A Novel Multifactor Clustering Integration Paradigm Based on Two-Stage Feature Engineering and Improved Bidirectional Deep Neural Networks for Exchange Rate Forecasting. *Digit. Signal Process.* **2023**, *143*, 104258. [CrossRef]
37. Zhou, Y.; Zhu, X. Forecasting USD/RMB Exchange Rate Using the ICEEMDAN–CNN–LSTM Model. *J. Forecast.* **2025**, *44*, 200–215. [CrossRef]
38. Date, P.; Maunthrooa, J. Modelling and Forecasting of Exchange Rate Pairs Using the Kalman Filter. *J. Forecast.* **2025**, *44*, 606–622. [CrossRef]
39. Souto, H.G.; Heuvel, S.K.; Neto, F.L. Time-Mixing and Feature-Mixing Modelling for Realized Volatility Forecast: Evidence from TSMixer Model. *J. Financ. Data Sci.* **2024**, *10*, 100143. [CrossRef]
40. Nguyen, H.T.; Nguyen, H.; Tran, M.-N. Deep Learning Enhanced Volatility Modeling with Covariates. *Financ. Res. Lett.* **2024**, *69*, 106145. [CrossRef]
41. Mishra, A.K.; Renganathan, J.; Gupta, A. Volatility Forecasting and Assessing Risk of Financial Markets Using Multi-Transformer Neural Network Based Architecture. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108223. [CrossRef]
42. Jazayeri, K. Predicting Multi-Horizon Currency Exchange Rates Using a Stacked Ensemble of Random Forest and SVR. *IDT* **2024**, *18*, 297–325. [CrossRef]
43. Zhu, Q.; Diao, X. From Stochastic to Rough Volatility: A New Deep Learning Perspective on Hedging. *Fractal Fract.* **2023**, *7*, 225. [CrossRef]
44. Chen, F.; Sha, Y.; Ji, H.; Peng, K.; Liang, X. Integrating Multifractal Features into Machine Learning for Improved Prediction. *Fractal Fract.* **2025**, *9*, 205. [CrossRef]
45. Lee, S.; Cho, P. Graph-Based Stock Volatility Forecasting with Effective Transfer Entropy and Hurst-Based Regime Adaptation. *Fractal Fract.* **2025**, *9*, 339. [CrossRef]
46. Cho, P.; Lee, M. Dynamic Black–Litterman Portfolios Incorporating Asymmetric Fractal Uncertainty. *Fractal Fract.* **2024**, *8*, 642. [CrossRef]
47. Ji, J.; Wang, S.; Wei, Y. Multi-Scale Explainable AI for RMB Exchange Rate Drivers. *Forecasting* **2026**, *8*, 7. [CrossRef]
48. Zhu, Q.; Li, Y. Fractal Dimension and Chaotic Dynamics of Multiscale Network Factors in Asset Pricing: A Wavelet Packet Decomposition Approach Based on the Fractal Market Hypothesis. *Fractal Fract.* **2026**, *10*, 196. [CrossRef]
49. Liu, R.; Segnon, M.; Cepni, O.; Gupta, R. Forecasting Volatility of Commodity, Currency, and Stock Markets: Evidence From Markov-Switching Multifractal Models. *J. Forecast.* **2026**, *45*, 2905–2941. [CrossRef]
50. Müller, U.A.; Dacorogna, M.M.; Davé, R.D.; Olsen, R.B.; Pictet, O.V.; von Weizsäcker, J.E. Volatilities of Different Time Resolutions: Analyzing the Dynamics of Market Components. *J. Empir. Financ.* **1997**, *4*, 213–239. [CrossRef]
51. Calvet, L.; Fisher, A. Multifractality in Asset Returns: Theory and Evidence. *Rev. Econ. Stat.* **2002**, *84*, 381–406. [CrossRef]
52. Calvet, L.E.; Fisher, A.J. How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *J. Financ. Econom.* **2004**, *2*, 49–83. [CrossRef]
53. Lux, T. The Markov-Switching Multifractal Model of Asset Returns: GMM Estimation and Linear Forecasting of Volatility. *J. Bus. Econ. Stat.* **2008**, *26*, 194–210. [CrossRef]
54. Zeng, Y.; Tang, B.; Zhou, Z.; Lee, R.S.T. DFC-LSTM: A Novel LSTM Architecture Integrating Dynamic Fractal Gating and Chaotic Activation for Value-at-Risk Forecasting. *Mathematics* **2025**, *13*, 3587. [CrossRef]
55. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Havlin, S.; Bunde, A.; Stanley, H.E. Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series. *Physica A Stat. Mech. Its Appl.* **2002**, *316*, 87–114. [CrossRef]
56. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely Randomized Trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]
57. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
58. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]

59. Diebold, F.X.; Mariano, R.S. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **2002**, *20*, 134–144. [CrossRef]
60. Hansen, P.R. A Test for Superior Predictive Ability. *J. Bus. Econ. Stat.* **2005**, *23*, 365–380. [CrossRef]
61. Gneiting, T.; Raftery, A.E. Strictly Proper Scoring Rules, Prediction, and Estimation. *J. Am. Stat. Assoc.* **2007**, *102*, 359–378. [CrossRef]
62. Mincer, J.; Zarnowitz, V. The Evaluation of Economic Forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*; Mincer, J., Ed.; National Bureau of Economic Research: Cambridge, MA, USA, 1969; pp. 3–46.
63. Bolgorian, M.; Gharli, Z. A Multifractal Detrended Fluctuation Analysis of Gold Price Fluctuations. *Acta Physica Pol. B* **2011**, *42*, 159–168.
64. Wang, L.; Lee, R.S. Stock Index Return Volatility Forecast via Excitatory and Inhibitory Neuronal Synapse Unit with Modified MF-ADCCA. *Fractal Fract.* **2023**, *7*, 292. [CrossRef]
65. Gao, J.; Wang, J.; Wei, D.; Zeng, B. An Innovative Decision-Making System Integrating Multifractal Analysis and Volatility Forecasting. *Ann. Oper. Res.* **2026**, *357*, 45–87. [CrossRef]
66. Chi, Z.; Gao, J.; Wang, C. Global Stock Market Volatility Forecasting Incorporating Dynamic Graphs and All Trading Days. *J. Forecast.* **2026**, *45*, 1311–1324. [CrossRef]
67. Leventides, J.; Melas, E.; Poulios, C.; Livada, M.; Poulios, N.C.; Boufounou, P. Application of the Fractal Brownian Motion to the Athens Stock Exchange. *Fractal Fract.* **2024**, *8*, 454. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.