Session-Conditioned Gold Regime Forecasting

A Quantitative Framework for Volatility, Breakout Quality, and Near-Term Environment Persistence

Limitless PH Quant Research

15 September 2026

Abstract

This paper presents a public, implementation-neutral framework for estimating whether a recently observed gold-market environment is likely to persist over the next several trading sessions. The objective is not to forecast the exact future path or direction of XAUUSD. Instead, the model targets quantities that are more directly relevant to systematic breakout execution: realized volatility, jump risk, liquidity conditions, directional efficiency, range acceptance, false-break risk, and latent regime persistence.

The framework combines four layers. First, a fixed 08:00-10:00 America/New_York observation window is converted into a compact intraday state vector. Second, heterogeneous autoregressive realized-volatility models (HAR-RV and extensions) estimate the persistence of volatility across daily, weekly, and monthly components. Third, a Hidden Markov or Markov-switching layer estimates the probability and expected duration of latent market regimes. Fourth, slower macro and derivatives variables - including real yields, the U.S. dollar, implied volatility, options term structure, and scheduled-event risk - condition the longer-horizon prior.

A historical-analogue and Bayesian layer converts the current state into directly interpretable recurrence probabilities, while strict walk-forward validation, probability calibration, and execution-level out-of-sample testing are used to determine whether the framework adds value beyond simple rolling-volatility and regime-persistence baselines.

The core hypothesis is intentionally limited: a two-hour New York observation may contain useful information about the *type and intensity* of near-term movement, but its direct information content should decay with forecast horizon. The resulting system is therefore designed as an environment-classification and risk-conditioning framework, not a directional price oracle.

Research status: Methodology paper. No proprietary strategy rules, parameter values, position-sizing rules, execution thresholds, code, account data, internal model names, or unpublished performance statistics are disclosed.

Keywords: gold, XAUUSD, realized volatility, HAR-RV, regime switching, hidden Markov model, market microstructure, breakout quality, implied volatility, Bayesian forecasting, walk-forward validation.

1. Public Disclosure Boundary

This is the **public research edition** of a broader internal quantitative research program. The paper intentionally discloses only the statistical framework necessary to evaluate the research hypothesis.

The following are **not** disclosed:

- proprietary entry and exit logic;
- execution times or broker-specific activation rules;
- position-sizing formulas or risk percentages;
- stop-loss, take-profit, basket, or scaling rules;
- internal strategy or model identifiers;
- source code or production-system architecture;
- proprietary feature thresholds;
- account-level or client-level data;
- unpublished live, forward, or backtest performance statistics;
- private research branches and internal implementation decisions.

The 08:00-10:00 America/New_York period used in this paper is a **research observation window**. It should not be interpreted as disclosure of any proprietary execution window.

2. Research Question

The practical question is:

> Can the current gold-market state be translated into a calibrated probability that the next several sessions will remain favorable for a generic short-horizon breakout process?

A realistic recent backtest can show that a strategy is compatible with the market conditions that actually occurred. It cannot, by itself, estimate the probability that those conditions will recur.

The distinction is:

$$\text{execution realism} \neq \text{statistical recurrence probability}\,.$$

The formal target is therefore an environment probability rather than a price target:

$$P\left(Y_{t,h}=1 \mid I_t\right), h \in \{1, 5, 10, 22\},$$

where $Y_{t,h}$ is a favorable future environment over horizon h, and I_t is the information set actually available at time t.

The principal research horizon is five trading sessions. Longer horizons are included to test how quickly the information in a narrow intraday window decays.

3. Why a New York Intraday State May Matter

The 08:00-10:00 New York window is economically meaningful because it overlaps a period of substantial U.S. market participation and includes many 08:30 ET U.S. macroeconomic releases. Elder, Miao, and Ramchander (2012) document swift and significant responses in metal futures to U.S. macroeconomic news, with 08:30 announcements among the most influential for returns, volatility, and volume.

Intraday volatility is also strongly periodic. Andersen and Bollerslev (1997) show that ignoring intraday periodicity can materially distort high-frequency volatility analysis. This supports treating a fixed New York session slice as a distinct state observation rather than pooling all intraday observations indiscriminately.

The working information-content hypothesis is:

$$I\left(X_t^{AM}; \sigma_{t+1:t+h}, Q_{t+1:t+h}\right) > I\left(X_t^{AM}; sign\left(R_{t+1:t+h}\right)\right),$$

where X_t^{AM} is the session state vector, σ is future volatility, Q is future breakout quality, and $I(\cdot;\cdot)$ denotes information content.

In plain language: the session is expected to contain more information about *how the market may move* than *which direction it must move*.

4. Data Architecture

4.1 Market data

Two complementary data sources are appropriate.

Centralized gold futures data are preferred when the research requires actual exchange volume, order-flow imbalance, order-book depth, cancellations, or market-impact estimates.

Broker XAUUSD tick data are preferred when validating the economic usefulness of the final environment model under the actual pricing, spread, and execution conditions faced by a downstream trading system.

This creates a clean research hierarchy:

$$\text{market-information state} \rightarrow \text{environment probability} \rightarrow \text{execution-level validation}.$$

No result should be assumed transferable from centralized futures data to a broker execution environment without explicit testing.

4.2 Macro and derivatives data

Candidate slow or forward-looking variables include:

- U.S. real yields;
- a broad U.S. dollar index;
- gold implied volatility;
- short- and medium-dated ATM implied volatility;
- put/call volatility skew;
- broader rates or equity volatility measures;
- scheduled CPI, employment, and central-bank event indicators;
- positioning and flow measures where historically timestamped data are available.

These variables are not intended to replace the intraday state. They condition the probability that an observed short-term regime will persist.

4.3 Timestamp discipline

Every observation must be stored according to when it was **actually knowable**. The admissible information set is:

$$I_t = \{ X_s : \tau_s^{available} \le t \}.$$

Historical revisions, delayed publications, or restated macro data must not be allowed to leak future information into the model.

5. Intraday Feature Engineering

Let $p_{t,i}$ denote the observed gold price for intraday observation i on trading day t, and define log returns as

$$r_{t,i} = \ln p_{t,i} - \ln p_{t,i-1}.$$

The default research sampling frequency can begin at one minute, with sensitivity tests at finer or coarser frequencies to control for microstructure noise.

5.1 Realized variance

For the fixed session containing M observations:

$$RV_t^{AM} = \sum_{i=1}^{M} r_{t,i}^2.$$

The corresponding full-day measure is

$$RV_t^D = \sum_{i=1}^{N_t} r_{t,i}^2.$$

Realized volatility is the square root of realized variance:

$$\sigma_t^{RV} = \sqrt{RV_t}.$$

5.2 Bipower variation and jump proxy

Following Barndorff-Nielsen and Shephard (2004), bipower variation can be approximated by

$$BV_t = \mu_1^{-2} \sum_{i=2}^{M} |r_{t,i}| |r_{t,i-1}|, \mu_1 = \sqrt{\frac{2}{\pi}}.$$

A simple non-negative jump proxy is

$$J_t = max(RV_t - BV_t, 0).$$

The point is not merely to label a day "volatile." Two sessions can have similar ranges while one is smooth and persistent and the other is dominated by discontinuous event risk.

5.3 Garman-Klass variance

A secondary range-based estimator is

$$\sigma_{GK,t}^2 = \frac{1}{2}\left[\ln\left(\frac{H_t}{L_t}\right)\right]^2 - (2\ln 2 - 1)\left[\ln\left(\frac{C_t}{O_t}\right)\right]^2.$$

This is treated as a challenger feature, not assumed to be superior to realized variance.

5.4 Range normalization

The observed session range is

$$Range_t^{AM} = H_t^{AM} - L_t^{AM}.$$

A generic normalized range is

$$NR_t = \frac{Range_t^{AM}}{S_t},$$

where S_t is a trailing volatility scale such as rolling ATR or realized volatility. The exact production scale and thresholds are deliberately outside this public paper.

5.5 Directional efficiency

Directional efficiency distinguishes clean displacement from noisy two-sided movement:

$$DE_t = \frac{|C_t - O_t|}{\sum_{i=1}^{M} |p_{t,i} - p_{t,i-1}|}.$$

Values closer to one indicate a more one-directional path. Values closer to zero indicate greater back-and-forth movement.

5.6 Close-location value

A simple acceptance measure is

$$CLV_t = \frac{2C_t - H_t - L_t}{H_t - L_t}.$$

When the close lies inside the observed range, CLV_t is approximately bounded between -1 and +1. It helps distinguish accepted displacement from an excursion that is largely rejected before the end of the observation window.

5.7 Relative volume and liquidity proxies

Where reliable volume is available:

$$RelVol_t = \frac{V_t^{AM}}{\mathrm{median}\left(V_{t-1:t-L}^{AM}\right)},$$

for a trailing window L fixed before testing.

A simple Amihud-style state variable is

$$ILLIQ_t = \frac{|R_t^{AM}|}{V_t^{AM}}.$$

If full order-book data are available, order-flow imbalance, depth, cancellations, and price-impact measures can be used. If not, broker-level proxies should be labeled clearly as proxies rather than treated as equivalent to centralized depth.

5.8 Session share of daily volatility

A particularly useful state variable is

$$s_t = \frac{RV_t^{AM}}{RV_t^{D}}.$$

This measures whether the chosen session is contributing an unusually large or small fraction of the day's total realized variation.

6. Dual HAR-RV Volatility Layer

Corsi's HAR-RV framework is an appropriate statistical anchor because it captures volatility components operating over heterogeneous horizons while remaining parsimonious and interpretable (Corsi, 2009).

6.1 Full-day HAR-RV

Define daily, weekly, and monthly components:

$$RV_t^{(d)} = RV_t^D,$$

$$RV_t^{(w)} = \frac{1}{5}\sum_{j=0}^{4} RV_{t-j}^D,$$

$$RV_t^{(m)} = \frac{1}{22}\sum_{j=0}^{21} RV_{t-j}^D.$$

A log-HAR specification is

$$\log RV_{t+1}^D = \beta_0 + \beta_d \log RV_t^{(d)} + \beta_w \log RV_t^{(w)} + \beta_m \log RV_t^{(m)} + \varepsilon_{t+1}.$$

6.2 Model A: full-day HAR multiplied by session share

Vanilla HAR-RV predicts full-day volatility. To adapt it to a fixed intraday observation window, estimate a session-share process separately:

$$\widehat{RV}_{t+1}^{AM,A} = \widehat{RV}_{t+1}^D \times \hat{s}_{t+1}.$$

The session-share estimate $\hat{s}_{t+1}$ must be generated using past data only and may be conditioned on documented categories such as event versus non-event sessions.

This approach allows the cleaner full-day persistence estimate to do most of the statistical work while the intraday component allocates the expected volatility across the day.

6.3 Model B: session-siloed HAR

An alternative is to model the fixed session as its own daily time series:

$$\log RV_{t+1}^{AM} = \alpha_0 + \alpha_d \log RV_t^{AM} + \alpha_w \log \overline{RV}_{t,5}^{AM} + \alpha_m \log \overline{RV}_{t,22}^{AM} + u_{t+1}.$$

The session-only model is more targeted but also noisier because each day contributes far fewer observations than a full trading day.

6.4 HAR disagreement as a diagnostic

The two forecasts should not automatically be averaged. Their disagreement can itself be a regime feature:

$$D_t^{HAR} = \log\left(\frac{\widehat{RV}_{t+1}^{AM,B}}{\widehat{RV}_{t+1}^{AM,A}}\right).$$

A persistent divergence may indicate that the fixed New York session is taking a structurally different share of total daily volatility than its historical norm.

6.5 Jump-augmented HAR

A simple extension is

$$\log RV_{t+1} = \beta_0 + \beta_d \log RV_t + \beta_w \log RV_t^{(w)} + \beta_m \log RV_t^{(m)} + \beta_J J_t + \varepsilon_{t+1}.$$

This lets event-driven jump activity contribute separately from smoother continuous variation.

6.6 HAR-X

Exogenous variables can be incorporated after the plain model is established:

$$\log RV_{t+1} = \beta_0 + \beta_d \log RV_t + \beta_w \log RV_t^{(w)} + \beta_m \log RV_t^{(m)} + \gamma' Z_t + \varepsilon_{t+1},$$

where Z_t can include implied volatility, dollar conditions, real yields, event indicators, or other variables that demonstrate incremental out-of-sample value.

Gold-specific evidence supports this direction. Wei et al. (2020) find that CBOE gold and silver implied-volatility information improves realized-volatility forecasting in HAR-type models, while Gkillas, Gupta, and Pierdzioch (2020) find that geopolitical-risk variables can add information at longer volatility-forecast horizons under a quantile-HAR framework.

7. Dynamic Macro Conditioning

A macro variable should not receive a permanent fixed coefficient merely because it was influential during one market regime.

For a generic real-yield factor x_t, estimate a rolling relationship:

$$r_t^{Gold} = \alpha_t + \beta_t \Delta x_t + \epsilon_t.$$

The model should retain at least:

- $\hat{\beta}_t$, the current sensitivity estimate;
- R_t^2, the explanatory strength;
- the stability of the sign and magnitude of $\hat{\beta}_t$;
- residual volatility.

A generic confidence-weighted contribution can be written as

$$MacroContribution_t = w_t \hat{\beta}_t \Delta x_t,$$

where w_t is an increasing function of the recent explanatory strength and parameter stability.

The principle is straightforward: when the relationship weakens, the model should become less confident rather than continue to apply a stale coefficient.

8. Options and Forward-Uncertainty Layer

Realized-volatility models are backward-looking. Options provide a market-priced forward view of uncertainty.

A basic term-structure measure is

$$Term_t = IV_t^{short} - IV_t^{medium},$$

or equivalently

$$TermRatio_t = \frac{IV_t^{short}}{IV_t^{medium}}.$$

A matched-delta skew measure can be represented as

$$Skew_t = IV_t^{put,\Delta} - IV_t^{call,\Delta}.$$

The derivatives layer can therefore contain

$$O_t = \left[IV_t^{short}, IV_t^{medium}, Term_t, Skew_t \right).$$

The literature supports treating implied volatility as potentially useful information rather than merely a restatement of realized volatility. Wei et al. (2020), for example, report improved gold-volatility forecasts when CBOE implied-volatility indices are incorporated into HAR-type models.

9. Latent Regime Model

Observed features can be interpreted as noisy emissions from an unobserved market state S_t.

A simple three-state model might be summarized economically as:

1. low-volatility or compression state;
2. directional expansion state;
3. stress or jump-dominated state.

These labels are descriptive. They must be assigned from feature characteristics, not retroactively from trading profitability.

9.1 Hidden Markov structure

Let

$$S_t \in \{1, \dots, K\}.$$

The transition matrix contains

$$P_{ij} = P(S_{t+1} = j \mid S_t = i).$$

If p_t is the current vector of state probabilities, then

$$p_{t+h} = p_t P^h.$$

Expected regime duration for state i is

$$E[D_i] = \frac{1}{1 - P_{ii}}.$$

Hamilton's regime-switching framework provides the foundational logic for this class of models (Hamilton, 1989).

9.2 Time-varying transition probabilities

Transition probabilities can be conditioned on slower variables:

$$P(S_{t+1} = j \mid S_t = i, Z_t) = \frac{\exp(a_{ij} + b_{ij}' Z_t)}{\sum_k \exp(a_{ik} + b_{ik}' Z_t)}.$$

This lets the macro environment alter the probability that a current intraday regime will persist without forcing macro variables to dictate the direction of gold.

Gold-specific work using Markov regime-switching GARCH-MIDAS models finds evidence that policy-uncertainty variables contain volatility information and that regime-switching specifications can improve statistical forecasting performance (Ma et al., 2021).

10. Historical-Analogue Layer

A transparent analogue model is useful because it answers a simple question:

> When the market previously looked similar to today, what happened next?

Let the standardized state vector be

$$z_t = \left[\frac{x_{t1} - \mu_1}{\sigma_1}, \dots, \frac{x_{tp} - \mu_p}{\sigma_p} \right).$$

For Euclidean distance:

$$d(t, s) = \sqrt{\sum_{j=1}^{p} (z_{tj} - z_{sj})^2}.$$

A covariance-aware alternative is Mahalanobis distance:

$$d_M(t,s)=\sqrt{(x_t-x_s)'\Sigma^{-1}(x_t-x_s)}.$$

For the k nearest historical states $N_k(t)$, the raw recurrence estimate is

$$\hat{p}_{t,h}^{NN}=\frac{1}{k}\sum_{s\in N_k(t)}Y_{s,h}.$$

A distance-weighted version is

$$\hat{p}_{t,h}^{WNN}=\frac{\sum_{s\in N_k(t)}w_{ts}Y_{s,h}}{\sum_{s\in N_k(t)}w_{ts}}.$$

This layer is deliberately interpretable: the researcher can inspect the number of comparable historical states, their distances, and the distribution of subsequent outcomes.

11. Bayesian Recurrence Estimate

If the analogue set contains x favorable outcomes and $n-x$ unfavorable outcomes, a Beta prior gives

$$p\sim Beta(\alpha,\beta),$$

and the posterior is

$$p\mid x,n\sim Beta(\alpha+x,\beta+n-x).$$

With a neutral $Beta(1,1)$ prior, the posterior mean is

$$E[p\mid x,n]=\frac{x+1}{n+2}.$$

This avoids presenting a raw historical fraction as if it were a perfectly known probability and naturally produces a credible interval that widens when the number of genuinely comparable historical states is small.

12. Environment Targets

The public framework does not disclose proprietary trade rules. Instead, it defines generic research targets applicable to any short-horizon breakout process.

Examples include:

$$Y_{t,h}^{favorable}=1\left(R_{t:t+h}^{strategy}>0\right),$$

$$Y_{t,h}^{drawdown}=1\left(DD_{t:t+h}<c\right),$$

and event-level quantities such as

$$MFE_t, MAE_t, FalseBreak_t, Continuation_t.$$

The threshold c and the exact definitions of a proprietary strategy's favorable environment are intentionally not disclosed.

A broader probabilistic target is

$$P\left(\text{favorable environment over next } h \text{ sessions} \mid I_t\right).$$

13. Probability Fusion

One possible implementation combines multiple calibrated probability components:

$$p_t^{final} = \sum_{m=1}^{M} w_m p_{t,m}, w_m \geq 0, \sum_m w_m = 1.$$

Candidate components are:

- volatility-persistence probability;
- latent-regime persistence probability;
- historical-analogue probability;
- macro-conditioned probability;
- derivatives-conditioned probability.

Weights must be estimated or fixed **before** evaluation on the final holdout. A more sophisticated stacking model is permissible only if it improves calibration out of sample.

14. Why a Recent One- or Two-Week Test Is Not Enough

Suppose a recent test contains only n trading sessions. Even an apparently strong success rate has substantial statistical uncertainty when n is small.

For a binary outcome, a Wilson interval is

$$\hat{p}_W = \frac{\hat{p} + z^2/(2n)}{1 + z^2/n},$$

with half-width

$$H = \frac{z}{1 + z^2/n} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}.$$

The deeper issue is dependence. Five-session forward outcomes overlap, volatility clusters, and neighboring observations are not independent. Raw sample size therefore overstates effective information.

A rough effective-sample-size approximation is

$$N_{eff} \approx \frac{N}{1+2\sum_{k=1}^{K} \rho_k},$$

where ρ_k is serial correlation at lag k.

Block bootstrap, HAC/Newey-West errors, and purged walk-forward validation are therefore preferable to naive independent-sample confidence intervals when forward labels overlap.

15. Validation Protocol

The model should be considered useful only if it survives a frozen, chronological validation protocol.

15.1 Freeze before testing

Before fitting the final model, freeze:

- session definition and timezone handling;
- feature equations;
- target equations;
- forecast horizons;
- event-day treatment;
- missing-data rules;
- normalization rules;
- baseline models;
- model-selection metrics;
- final holdout period.

15.2 Walk-forward design

Random train/test splits are inappropriate for this problem. Every prediction at time t must use only information available before t.

A generic expanding design is:

$$T_t = \{1, \ldots, t-1\},$$

fit on T_t, forecast t, move forward, and repeat.

Rolling windows can be tested as a robustness alternative when structural change is a concern.

15.3 Purging and embargo

If the target is the next five sessions, adjacent labels share future returns. Training observations whose label windows overlap the validation period should be purged, with an embargo around the boundary where appropriate.

15.4 Probability calibration

For probability forecasts p_t and binary outcomes y_t, the Brier score is

$$BS = \frac{1}{T}\sum_{t=1}^{T}\left(p_t - y_t\right)^2.$$

Log loss is

$$LL = -\frac{1}{T}\sum_{t=1}^{T}\left[y_t \ln p_t + \left(1 - y_t\right)\ln\left(1 - p_t\right)\right].$$

Reliability plots should verify whether events forecast at approximately 70% actually occur near 70% of the time.

15.5 Volatility forecast loss

A common robust volatility loss is QLIKE:

$$QLIKE_t = \frac{RV_t}{\widehat{RV}_t} - \ln\left(\frac{RV_t}{\widehat{RV}_t}\right) - 1.$$

Mean squared error and mean absolute error should be reported alongside QLIKE.

16. Baselines Before Complexity

Every complex model should be required to beat simple alternatives such as:

1. unconditional historical environment probability;
2. previous-regime persistence;
3. trailing realized volatility;
4. plain HAR-RV;
5. simple historical analogues.

Only after those baselines are established should the research advance to:

- jump-augmented HAR;
- HAR-X;
- low-state HMM;
- time-varying transition models;
- Markov-switching GARCH-MIDAS;
- regularized nonlinear models.

Deep neural networks should be treated as late-stage challengers rather than default solutions. A more complex model is useful only if its out-of-sample improvement is stable, economically meaningful, and not driven by a small subperiod.

17. Falsifiable Hypotheses

The framework is designed to fail cleanly if the data do not support it.

H1. The fixed New York session's realized-volatility and range state improves one- to five-session volatility forecasts beyond a trailing-volatility baseline.

H2. Session-conditioned HAR forecasts outperform a naive rolling-volatility forecast under QLIKE out of sample.

H3. Jump variables add information about future false-break or unstable-expansion environments beyond total realized variance alone.

H4. A low-state HMM improves calibrated next-week regime probabilities beyond the assumption that yesterday's state simply persists.

H5. Historical analogues produce better-calibrated recurrence probabilities than the unconditional base rate.

H6. Macro and derivatives variables add incremental value after the intraday and HAR layers are already known.

H7. The direct value of one narrow intraday observation decays as the forecast horizon lengthens.

H8. A model that improves volatility forecasting but does not improve downstream risk-adjusted strategy outcomes should not be treated as economically useful.

18. Minimal Empirical Test

Before building a complete production model, the smallest useful experiment is:

Step 1. Build one daily state vector from the fixed New York observation window.

Step 2. Measure future realized volatility over one and five sessions.

Step 3. Compare trailing RV, full-day HAR, session-siloed HAR, and full-day HAR multiplied by session share.

Step 4. Bucket strictly out-of-sample volatility forecasts into quantiles and determine whether generic breakout-quality metrics separate meaningfully across the forecast distribution.

Step 5. Build a transparent nearest-neighbor analogue model.

Step 6. Proceed to a latent-regime model only if the simpler tests show genuine out-of-sample information.

This sequencing protects the research from becoming a search for a complicated model that merely explains historical noise.

19. Interpretation and Risk Use

The environment probability is not a trade signal by itself.

A generic downstream policy can be represented as

$$a_t^{*} = \arg\max_{a \in A} E\left[U\left(R_{t+1:t+h}\right) \mid I_t, a\right],$$

where a is a risk action and $U(\cdot)$ is a utility function that penalizes drawdown and tail risk.

The public implication is intentionally limited: probabilistic environment models are more defensible as **risk conditioners** than as mechanisms for increasing leverage. The first economic objective should be identifying historically adverse environments and reducing avoidable exposure, not maximizing risk whenever model confidence is high.

20. Limitations

Several limitations are structural.

First, a two-hour observation cannot encode future geopolitical shocks, policy surprises, or other genuinely new information.

Second, the useful predictive horizon for microstructure variables can be much shorter than the useful horizon for volatility persistence or macro state.

Third, centralized futures data and broker XAUUSD data are not identical markets. Transferability must be tested.

Fourth, economic-calendar and macro variables are highly vulnerable to look-ahead bias if historical revisions or publication timing are mishandled.

Fifth, latent-state models can create visually convincing regimes even when the states have no incremental economic value. State interpretability is not validation.

Sixth, repeated model experimentation creates data-snooping risk. The final holdout must remain untouched until the research specification is frozen.

Finally, this paper presents a framework rather than a result. No claim of live profitability, directional accuracy, or validated trading edge follows from the equations alone.

21. Conclusion

A recent execution-realistic test can demonstrate compatibility with the market environment that actually occurred. It does not measure how likely that environment is to recur.

A defensible recurrence model therefore needs to separate three questions:

1. **What state is the market in now?**

2. **How persistent is that state likely to be?**
3. **Does that state materially change the economics of a downstream breakout process?**

The proposed framework answers the first question with intraday realized-volatility, jump, range, volume, and acceptance features; the second with HAR-RV, latent-regime transitions, options, and macro conditioning; and the third with strictly out-of-sample strategy-level validation.

The important shift is conceptual. The target is not "Where will gold be next month?" It is:

$$P\left(\text{favorable near-term market environment} \mid \text{current observable state}\right).$$

That probability is useful only if it is calibrated, robust across time, and economically meaningful after execution costs. Until those conditions are met, the framework should remain a research hypothesis rather than a claimed edge.

References

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance, 4*(2-3), 115-158. https://doi.org/10.1016/S0927-5398(97)00004-2

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics, 2*(1), 1-37. https://doi.org/10.1093/jjfinec/nbh001

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics, 7*(2), 174-196. https://doi.org/10.1093/jjfinec/nbp001

Elder, J., Miao, H., & Ramchander, S. (2012). Impact of macroeconomic news on metal futures. *Journal of Banking & Finance, 36*(1), 51-65. https://doi.org/10.1016/j.jbankfin.2011.06.007

Fang, L., Chen, B., Yu, H., & Qian, Y. (2018). The importance of global economic policy uncertainty in predicting gold futures market volatility: A GARCH-MIDAS approach. *Journal of Futures Markets, 38*, 413-422. https://doi.org/10.1002/fut.21897

Gkillas, K., Gupta, R., & Pierdzioch, C. (2020). Forecasting realized gold volatility: Is there a role of geopolitical risks? *Finance Research Letters, 35*, 101280. https://doi.org/10.1016/j.frl.2019.08.028

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*(2), 357-384. https://doi.org/10.2307/1912559

Ma, F., Lu, X., Wang, L., & Chevallier, J. (2021). Global economic policy uncertainty and gold futures market volatility: Evidence from Markov regime-switching GARCH-MIDAS models. *Journal of Forecasting, 40*(6), 1070-1085. https://doi.org/10.1002/for.2753

Wei, Y., Liang, C., Li, Y., Zhang, X., & Wei, G. (2020). Can CBOE gold and silver implied volatility help to forecast gold futures volatility in China? Evidence based on HAR and Ridge regression models. *Finance Research Letters, 35*, 101287. https://doi.org/10.1016/j.frl.2019.09.002

World Gold Council. Gold Return Attribution Model (GRAM). Goldhub. Accessed 2026.

Appendix A. Core Equations

A1. Realized variance

$$RV_t = \sum_i r_{t,i}^2 .$$

A2. Bipower variation and jump proxy

$$BV_t = \mu_1^{-2} \sum_i |r_{t,i}| |r_{t,i-1}|, J_t = max(RV_t - BV_t, 0).$$

A3. Full-day HAR-RV

$$\log RV_{t+1} = \beta_0 + \beta_d \log RV_t + \beta_w \log \overline{RV}_{t,5} + \beta_m \log \overline{RV}_{t,22} + \varepsilon_{t+1} .$$

A4. Session share

$$s_t = \frac{RV_t^{AM}}{RV_t^D} .$$

A5. Session forecast from full-day HAR

$$\widehat{RV}_{t+1}^{AM,A} = \widehat{RV}_{t+1}^D \hat{s}_{t+1} .$$

A6. Session-siloed HAR

$$\log RV_{t+1}^{AM} = \alpha_0 + \alpha_d \log RV_t^{AM} + \alpha_w \log \overline{RV}_{t,5}^{AM} + \alpha_m \log \overline{RV}_{t,22}^{AM} + u_{t+1} .$$

A7. HAR divergence

$$D_t^{HAR} = \log \left(\frac{\widehat{RV}_{t+1}^{AM,B}}{\widehat{RV}_{t+1}^{AM,A}} \right) .$$

A8. Rolling macro sensitivity

$$r_t^{Gold} = \alpha_t + \beta_t \Delta x_t + \epsilon_t .$$

A9. Markov transition

$$P_{ij} = P(S_{t+1} = j \mid S_t = i), p_{t+h} = p_t P^h .$$

A10. Expected regime duration

$$E[D_i] = \frac{1}{1 - P_{ii}} .$$

A11. Historical analogue probability

$$\hat{p}_{t,h} = \frac{1}{k} \sum_{s \in N_k(t)} Y_{s,h}.$$

A12. Beta-Binomial posterior

$$p \mid x, n \sim Beta(\alpha + x, \beta + n - x).$$

A13. Brier score

$$BS = \frac{1}{T} \sum_{t} (p_t - y_t)^2.$$

A14. QLIKE

$$QLIKE_t = \frac{RV_t}{\widehat{RV}_t} - \ln\left(\frac{RV_t}{\widehat{RV}_t}\right) - 1.$$

Appendix B. Public Research Checklist

Before any claim of predictive value is published, confirm that:

1. the session clock is timezone-aware and daylight-saving correct;
2. every feature is available at the forecast origin;
3. forward labels are purged or embargoed where they overlap;
4. simple baselines are reported alongside complex models;
5. probabilities are evaluated for calibration, not only classification accuracy;
6. volatility forecasts are evaluated under QLIKE and standard loss functions;
7. the final holdout has not influenced feature or threshold selection;
8. execution-level usefulness is tested independently of statistical fit;
9. no result depends on one narrow historical subperiod;
10. no proprietary strategy logic is required to reproduce the public statistical methodology.