Quantitative Trading Module

Mathematical Structure and Limitations of Financial Time Series

Daniel Bloch

VinUnivesity & Paris Sorbonne

21st January 2026

Outline

- **The Fallacy of Mean Squared Error (MSE)**
  - Over-penalisation of Bold, Informative Forecasts
  - Preference for Uninformative 'Unconditional Mean' Models
- **Mean Absolute Error (MAE) and Risk Masking**
  - Inability to Distinguish Error Magnitude and Direction
  - Hiding Concentration of Failures in Market Crashes

Course Reference:
Futuretesting Quantitative Strategies
`http://ssrn.com/abstract=4647103`

Mathematical Properties of Asset Prices: Why Classical Statistics Fail in Markets

Defining Mean Squared Error (MSE)

The most ubiquitous loss function in predictive modelling is the **Mean Squared Error (MSE)**, which measures the average squared difference between the estimated values and the actual outcomes.

- **Mathematical Definition:** For a sequence of forecasts $\hat{y}_t$ and actual realizations y_t over n periods:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

- **Statistical Interpretation:** MSE corresponds to the variance of the forecast error. Minimising MSE is equivalent to finding the **conditional expectation** $E[y|\mathscr{F}_{t-1}]$ under Gaussian assumptions.

Defining Mean Squared Error (MSE)

- **Symmetry:** It treats over-predictions and under-predictions of the same magnitude identically, penalising errors quadratically (i.e., larger errors are penalised disproportionately more than smaller ones).

While MSE is the 'gold standard' in classical statistics, it faces significant challenges when applied to the asymmetric and heavy-tailed world of finance.

Why MSE is ill-suited for financial time series

Mean Squared Error (MSE) is mathematically convenient but economically flawed because it prioritises statistical precision over financial utility.

Key limitations of MSE

- Over-penalises bold, information-rich forecasts,
- Rewards uninformative models that predict the unconditional mean,
- Ignores directional correctness,
- Produces misleading rankings across regimes and structural breaks,
- Conflicts with risk-sensitive financial objectives.

MSE is a purely statistical scoring rule. Financial modelling, by contrast, requires evaluation criteria that reflect economic value, risk asymmetry, and state-dependent relevance.

The MSE

Definition

Let $\widehat{X}_t$ denote the model prediction at time t, and let X_t be the realised value (e.g., an asset return). The MSE is defined as:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(X_t - \widehat{X}_t)^2.$$

MSE penalises forecast errors quadratically, assigning disproportionate weight to large deviations between predicted and observed values.

While this sensitivity is often presented as a strength in classical forecasting settings, in financial markets it becomes a liability, as large deviations are precisely where economic relevance, regime shifts, and tail risk concentrate.

Non-stationarity and temporal regime dependence

MSE implicitly assumes a stationary error process, i.e. that the second moment of forecast errors is time-invariant:

$$E\left[\left(X_t - \widehat{X}_t\right)^2\right] = \sigma^2, \qquad \forall t.$$

This assumption is routinely violated in financial markets:

- Return distributions are heteroskedastic and regime-dependent,
- Large forecast errors cluster during crises and around discrete events (e.g., earnings or policy announcements),
- Predictability appears intermittently rather than uniformly over time.

As a consequence, MSE dilutes episodic predictability and systematically favours low-variance, 'safe' forecasts, often produced by models that ignore the structural dynamics of financial time series.

Disincentivises rare but profitable predictions

MSE penalises forecast errors quadratically. Consider a model that predicts a large move:

- Actual return: $+6\%$
- Forecast: $+4\%$

The resulting squared error is 4%, despite the forecast correctly identifying a rare and economically valuable event.

By contrast, a naive model that predicts $\widehat{X}_t = 0$ every day may achieve a lower MSE, despite being entirely uninformative.

Let $J_t = \mathbb{I}(|X_t| > \delta)$ denote an indicator of large moves. Then:

- $\widehat{X}_t = 0$ yields low expected MSE,
- $\widehat{X}_t = E[X_t \mid J_t = 1]$ may incur higher MSE, even if it captures the only value-relevant states.

This creates an economically perverse incentive: models that identify rare, profitable regimes are penalised, while models that avoid commitment are rewarded.

MSE is symmetric and ignores directionality

MSE penalises forecast errors symmetrically:

$$\left(X_t - \widehat{X}_t\right)^2 = \left(\widehat{X}_t - X_t\right)^2.$$

As a result, a model that predicts $+2\%$ when the market realises -2% is penalised identically to a model that predicts -2% when the market realises $+2\%$.

In trading and portfolio construction, however:

- The *sign* of the forecast determines profit and loss,
- Positive alpha arises from correct directional alignment.

MSE fails to distinguish between profitable and disastrous errors whenever their magnitudes coincide, rendering it blind to the most basic notion of trading success.

MSE rewards inaction in unpredictable environments

In volatile and noisy markets, a model that always predicts the unconditional mean,

$$\widehat{X}_t = 0,$$

will often achieve a lower MSE than a model that attempts to forecast meaningful movements.

This occurs because:

- Most observations are clustered near the mean,
- Incorrect directional forecasts generate large residuals that are squared,
- Repeatedly predicting zero minimises average squared error over time.

Such a model may achieve excellent MSE performance while delivering zero economic value and zero tradable signal.
MSE therefore systematically rewards conservative, uninformative forecasts and discourages models that attempt to capture economically relevant structure.

Poor robustness under structural breaks

Assume returns follow a process with a structural break at time τ:

$$X_t \sim \begin{cases} N(\mu_1, \sigma_1^2), & t \leq \tau, \\ N(\mu_2, \sigma_2^2), & t > \tau. \end{cases}$$

When evaluated over the full sample, MSE:

- Smooths across the break,
- Treats pre- and post-break errors as arising from a single regime,
- Ignores state-contingent predictability.

Consequently, a model that adapts well after τ but performs poorly before it may be ranked inferior to a model that is uniformly mediocre.

MSE thus obscures regime-dependent performance precisely when adaptation matters most.

Risk asymmetry is ignored

In financial decision-making, upside and downside errors are fundamentally asymmetric:

- Missing a $+5\%$ gain represents an opportunity cost,
- Missing a -5% loss represents a realised and potentially catastrophic loss.

MSE, however, weights both outcomes identically, based solely on squared magnitude.

This symmetry is incompatible with risk-sensitive objectives such as:

- Maximising Sharpe or Sortino ratios,
- Controlling drawdowns,
- Managing tail risk.

As a result, optimising for MSE can lead to models that appear statistically sound but perform poorly under economically meaningful risk criteria.

The Student-t Distribution and Tail Thickness

To model the 'fat tails' observed in financial markets, we move beyond the Gaussian distribution to the **Student-t distribution**, which is parameterised by the degrees of freedom ν.

- **Degrees of Freedom (ν):** This parameter controls the thickness of the tails.
  - As $\nu \to \infty$, the distribution converges to a Normal distribution.
  - For small ν, the probability of extreme outcomes is significantly higher than in the Gaussian case.
- **Moment Existence:** A critical property of the Student-t is that the p-th moment exists only if $p < \nu$.
  - If $\nu \leq 4$, the **kurtosis** is undefined or infinite.
  - If $\nu \leq 2$, the **variance** is infinite.
- **Financial Realism:** Empirical studies of daily returns often find ν between 3 and 5, suggesting that the second and fourth moments are either unstable or non-existent.

This has profound implications for error metrics like MSE.

Failure under fat tails and heteroskedasticity

Financial returns frequently exhibit heavy-tailed behaviour. For example, let

$$X_t \sim \text{Student-}t(\nu), \qquad \nu < 5.$$

In this setting:

- Higher moments may not exist or may be extremely unstable,
- Rare but severe shocks dominate squared-error losses,
- MSE becomes highly sensitive to a small number of extreme observations.

Models that attempt to capture large deviations are disproportionately penalised, while conservative models that suppress signal are overrewarded.

Consequently, under fat tails and time-varying volatility, MSE is neither statistically robust nor economically aligned.

Defining Mean Absolute Error (MAE)

As an alternative to MSE, the **Mean Absolute Error (MAE)** provides a measure of forecast accuracy that is less sensitive to extreme outliers.

- **Mathematical Definition:** For forecasts $\hat{y}_t$ and actuals y_t, the MAE is the average of the absolute residuals:

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t|$$

- **Linear Penalisation:** Unlike MSE, which squares the errors, MAE penalises all errors **linearly**. An error of 10 is exactly 10 times worse than an error of 1, regardless of the distribution.

Defining Mean Absolute Error (MAE)

- **Statistical Interpretation:** Minimising MAE is equivalent to finding the **conditional median** of the distribution. This makes it more 'robust' in the presence of the heavy tails (Student-t or Pareto) discussed previously.

While MAE solves the problem of outlier sensitivity, it still fails to address the unique requirements of financial decision-making.

Why MAE is ill-suited for financial time series

Although more robust to outliers than MSE, the Mean Absolute Error (MAE) remains ill-suited for evaluating financial forecasting models.

Its key limitations are:

Core shortcomings

- Lack of directional awareness,
- Indifference to timing and path dependence,
- Failure to reward economically valuable forecasts,
- Equal weighting across time, ignoring volatility regimes,
- Insensitivity to structural breaks and regime shifts.

A low MAE reflects statistical smoothness, not economic usefulness. In volatile, jump-driven markets, it provides little assurance of robustness or tradability.

The MAE

Definition

Let $\widehat{X}_t$ denote the model's prediction at time t, and X_t the realised return. The Mean Absolute Error (MAE) is defined as:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|X_t - \widehat{X}_t\right|.$$

Unlike MSE, MAE penalises errors linearly and is often preferred in the presence of outliers.

However, despite this robustness, MAE remains misaligned with the statistical structure of financial time series and with the economic objectives of trading and risk management.

Non-stationarity and regime dependence

Like MSE, MAE implicitly assumes stationarity in the joint distribution of predictions and realised values:

$$E\left[\,|X_t - \widehat{X}_t|\,\right] = \text{constant over time.}$$

In real financial markets, this assumption is violated:

- Volatility evolves over time,
- Return distributions are regime-switching,
- Predictability is episodic and concentrated in high-impact periods.

A model that performs well in calm regimes but fails during transitions or crises may still achieve a low MAE, despite being economically fragile or outright dangerous.

Insensitive to directionality

MAE measures only the magnitude of forecast errors, not their direction. For example, predicting a $+3\%$ return when the realised outcome is -3% produces the same absolute error as predicting -3% when the outcome is $+3\%$.

In trading applications, however:

- Directional mistakes lead directly to losses,
- Correct directional calls can remain profitable even with poor magnitude accuracy.

By ignoring sign information, MAE fails to distinguish between forecasts that are economically useful and those that are actively harmful in directional trading systems.

Penalises useful extremes

Consider a sharp return spike triggered by an unexpected event, such as a central bank announcement. A model that correctly anticipates this jump will necessarily produce a forecast far from the long-run mean, resulting in a large absolute error—even when the sign is correct and the trade is profitable.

By contrast, a naive model that always predicts

$$\widehat{X}_t = 0$$

will typically achieve a lower MAE simply by avoiding large forecasts.

This behaviour:

- Penalises bold but informative predictions,
- Rewards passive forecasts that miss economically meaningful events.

As a result, MAE systematically favours safety over insight, even when insight is precisely what generates trading value.

Poor alignment with economic or trading value

In practical trading and risk management, performance is evaluated through:

- Hit rates on large, infrequent moves,
- Risk-adjusted returns,
- Drawdown behaviour,
- Correct positioning during volatility spikes and regime shifts.

MAE, as a symmetric average loss, does not:

- Distinguish between low-error forecasts in economically irrelevant periods and large errors during critical episodes,
- Reward successful anticipation of rare but high-impact events.

By treating every time point as equally important, MAE conflicts with the way financial returns are generated and realised.

Masking path dependence and temporal clustering

Suppose a model performs poorly during a short but severe market crash, generating large absolute errors over a 10-day window, yet performs well during 200 days of calm conditions.

When averaged over the full sample:

- The overall MAE may appear acceptably low,
- The concentrated nature of the failure is obscured.

In practice, such behaviour would be catastrophic, as losses tend to cluster precisely when capital is most vulnerable.

By averaging errors uniformly over time, MAE masks path dependence and understates the economic impact of clustered risk events.

Aggregation over regimes gives misleading evaluation

Consider a regime-switching environment:

$$X_t \sim \begin{cases} X_{t,1}, & \text{Regime 1}, \\ X_{t,2}, & \text{Regime 2}. \end{cases}$$

A forecasting model may perform well in one regime and poorly in another.

MAE aggregates errors across regimes into a single average, thereby:

- Concealing regime-specific strengths and weaknesses,
- Obscuring fragility in adverse states,
- Producing misleading rankings of model quality.

This is particularly dangerous in finance, where downside risk is asymmetric and concentrated in specific regimes rather than evenly distributed over time.

The end

Thank You !