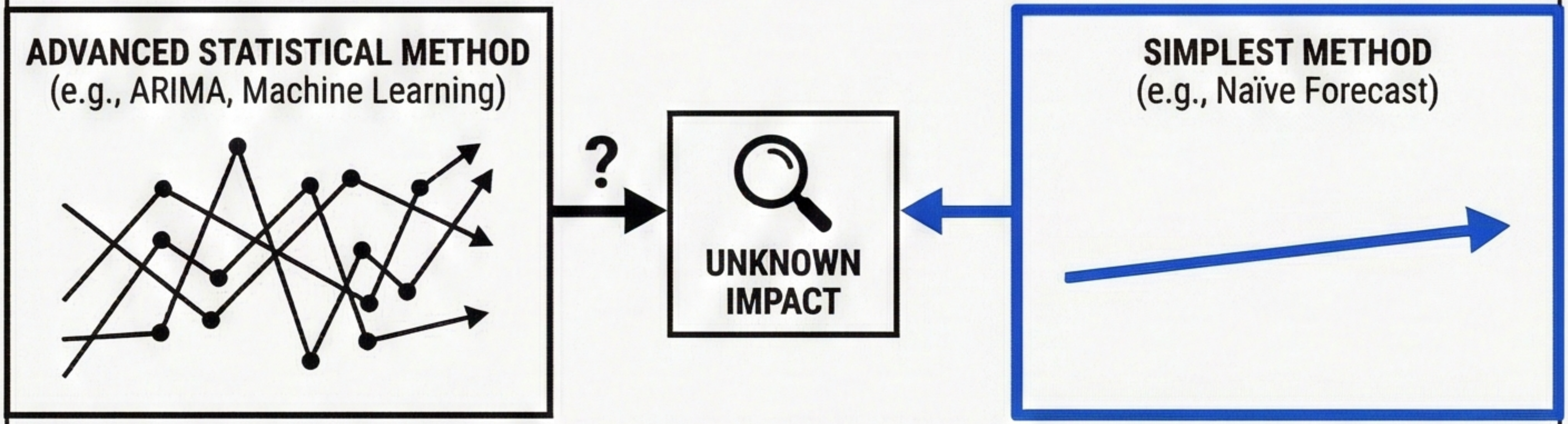
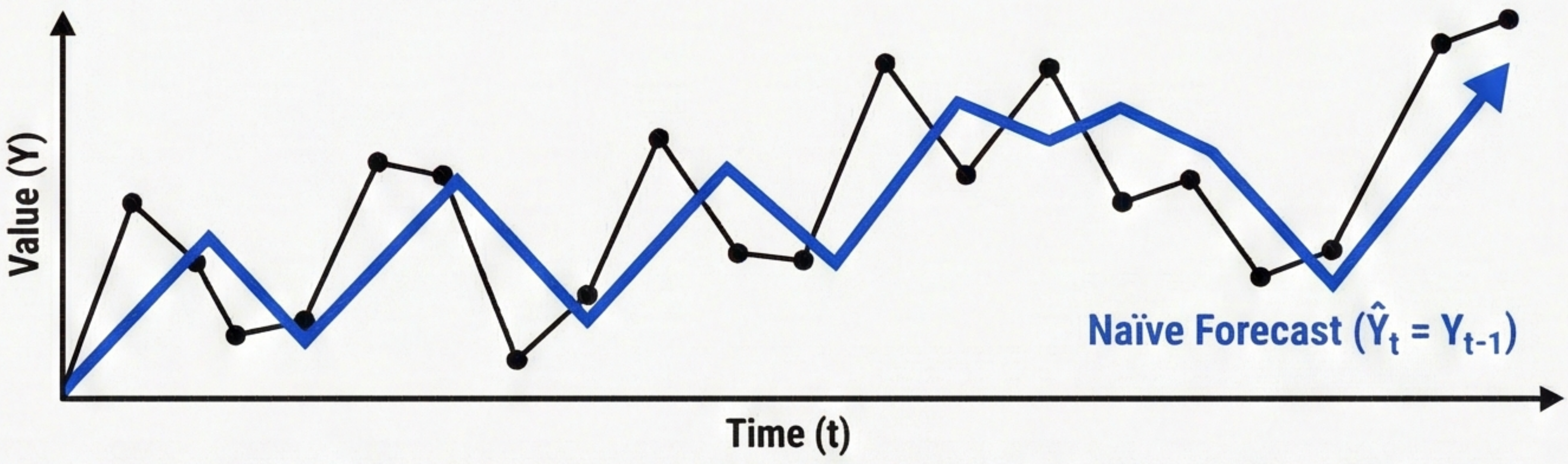
MEAN ABSOLUTE SCALED ERROR (MASE)

PROBLEM: DEMONSTRATING FORECAST VALUE



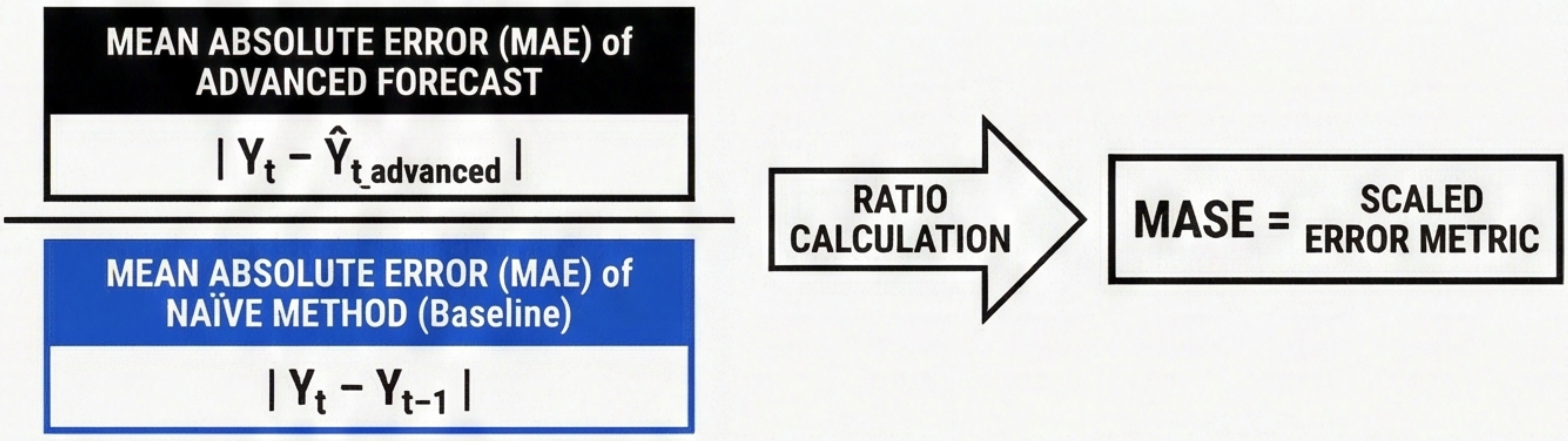
Data scientists struggle to show the true usefulness of complex models without a clear, simple benchmark.

THE BENCHMARK: NAÏVE METHOD (THE 'SCALING FACTOR')



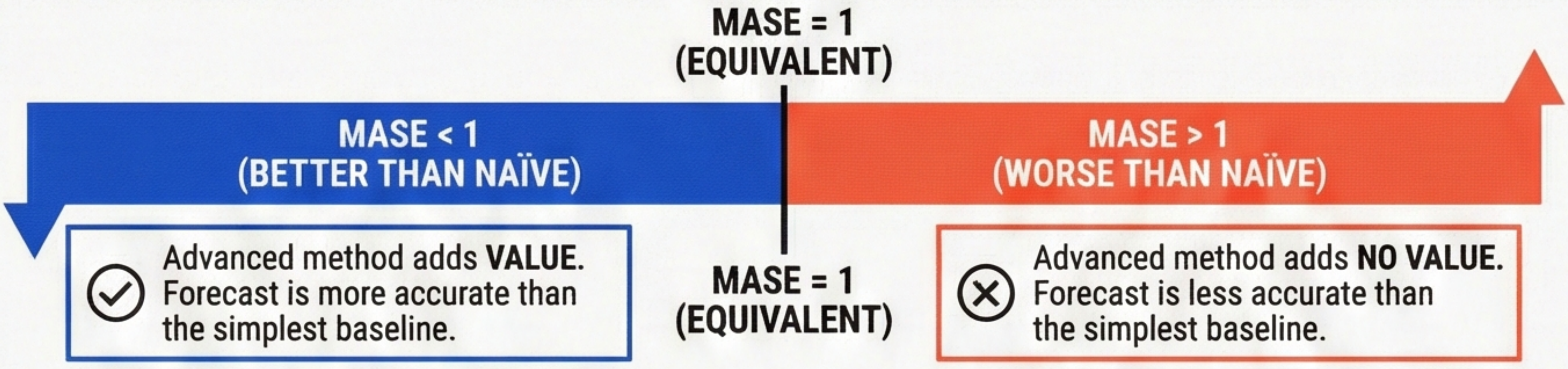
The "simplest method" serves as the reference point. MASE scales the forecast's error relative to the error of this naive baseline.

THE MASE FORMULA: DECONSTRUCTED



MASE directly compares the average error of the advanced model to the average error of the simplest (naïve) method.

INTERPRETATION: VALUE & IMPACT



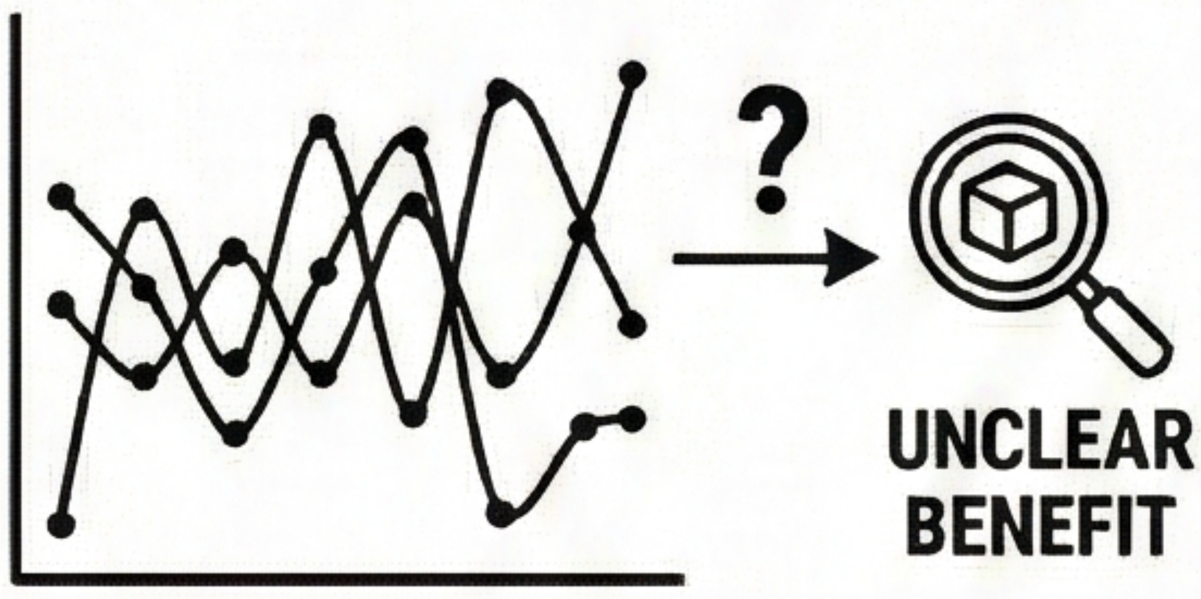
MASE quantifies the precise improvement (or degradation) of a forecast relative to the naive alternative, providing a measure of practical usefulness.

Zbigniew Galar

SEASONAL MEAN ABSOLUTE SCALED ERROR (MASE)

PROBLEM: ASSESSING FORECAST VALUE IN SEASONAL DATA

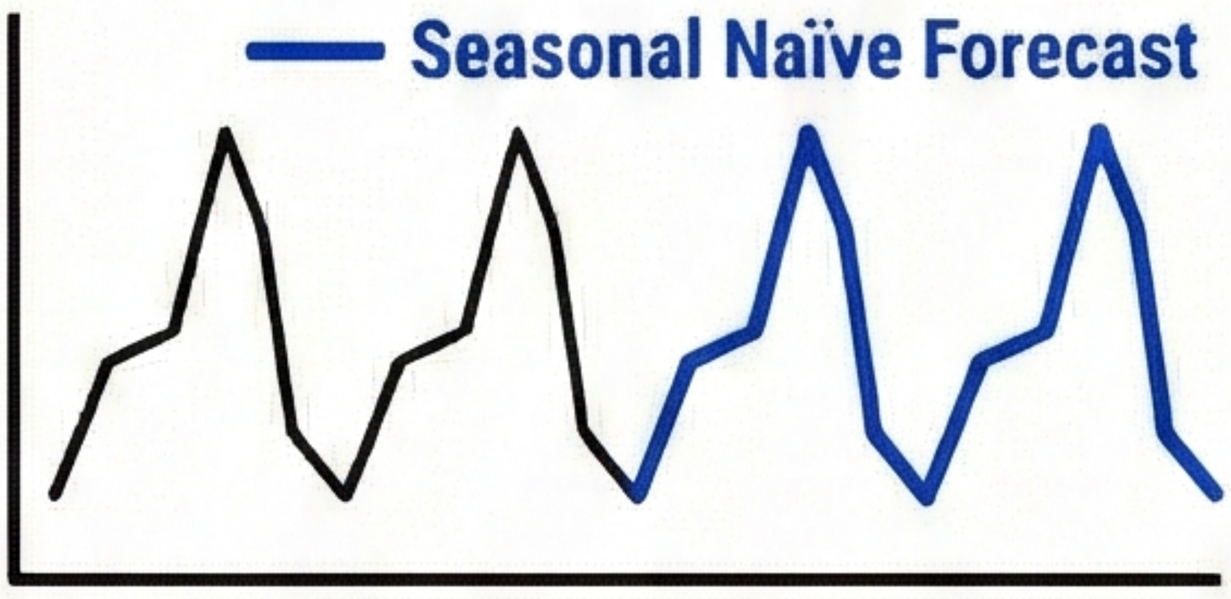
COMPLEX MODELS
(e.g., SARIMA, Deep Learning)



Advanced methods struggle to prove usefulness without a clear, simple reference point.

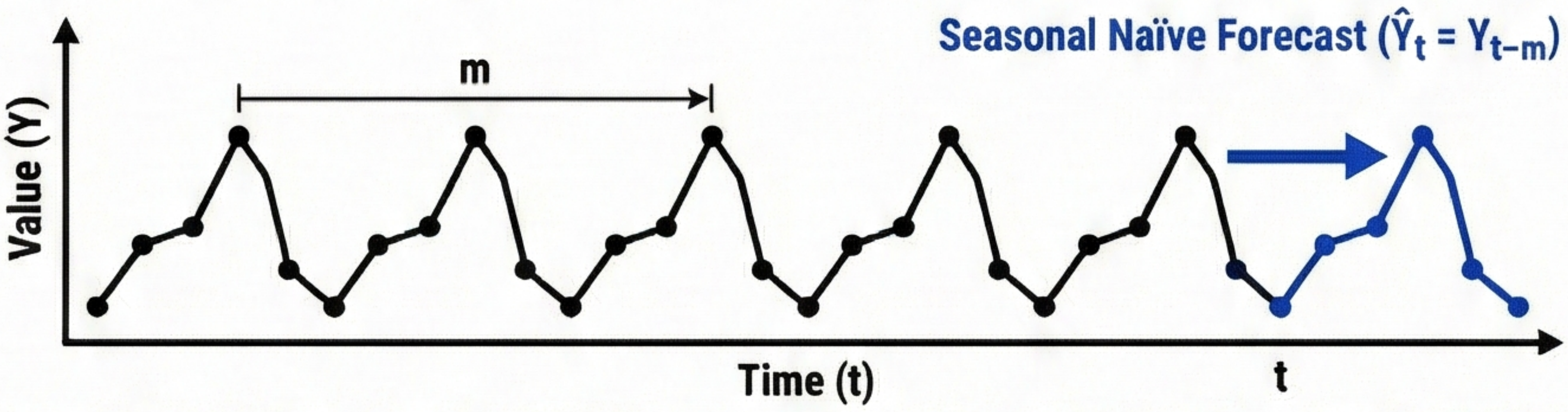
?

SIMPLE BENCHMARK
(Seasonal Naive)



The 'simplest method' provides a necessary baseline for comparison.

THE BENCHMARK: SEASONAL NAÏVE METHOD (THE 'SCALING FACTOR')



The Seasonal Naïve forecast for a given period is equal to the value from the same period in the previous cycle (m).

THE MASE FORMULA: DECONSTRUCTED

MEAN ABSOLUTE ERROR (MAE) of ADVANCED FORECAST

$$| Y_t - \hat{Y}_{t\_advanced} |$$

divided by

MEAN ABSOLUTE ERROR (MAE) of SEASONAL NAÏVE METHOD (Baseline)

$$| Y_t - Y_{t-m} |$$

RATIO CALCULATION →

$$\text{MASE} = \frac{\text{SCALED}}{\text{ERROR METRIC}}$$

MASE directly compares the average error of the advanced model to the average error of the simplest (seasonal naïve) method.

INTERPRETATION: VALUE & IMPACT

MASE = 1
(EQUIVALENT)

MASE < 1 (BETTER THAN NAÏVE)	MASE > 1 (WORSE THAN NAÏVE)
✓ Advanced method adds **VALUE.** Forecast is more accurate than the seasonal naive baseline.	✗ Advanced method adds **NO VALUE.** Forecast is less accurate than the seasonal naïve baseline.

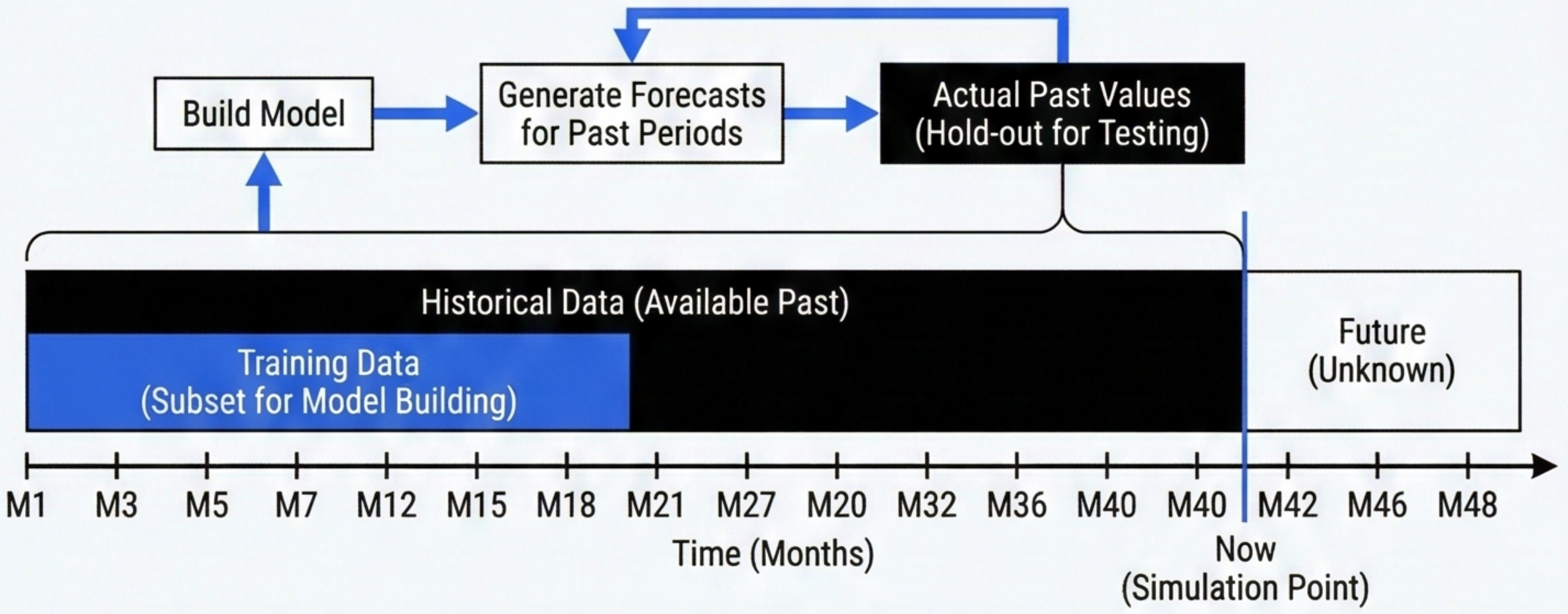
MASE quantifies the precise improvement (or degradation) of a forecast relative to the seasonal naïve alternative, providing a robust measure of usefulness.

Zbigniew Galar

BACKTESTING IN MONTHLY TIME SERIES FORECASTING

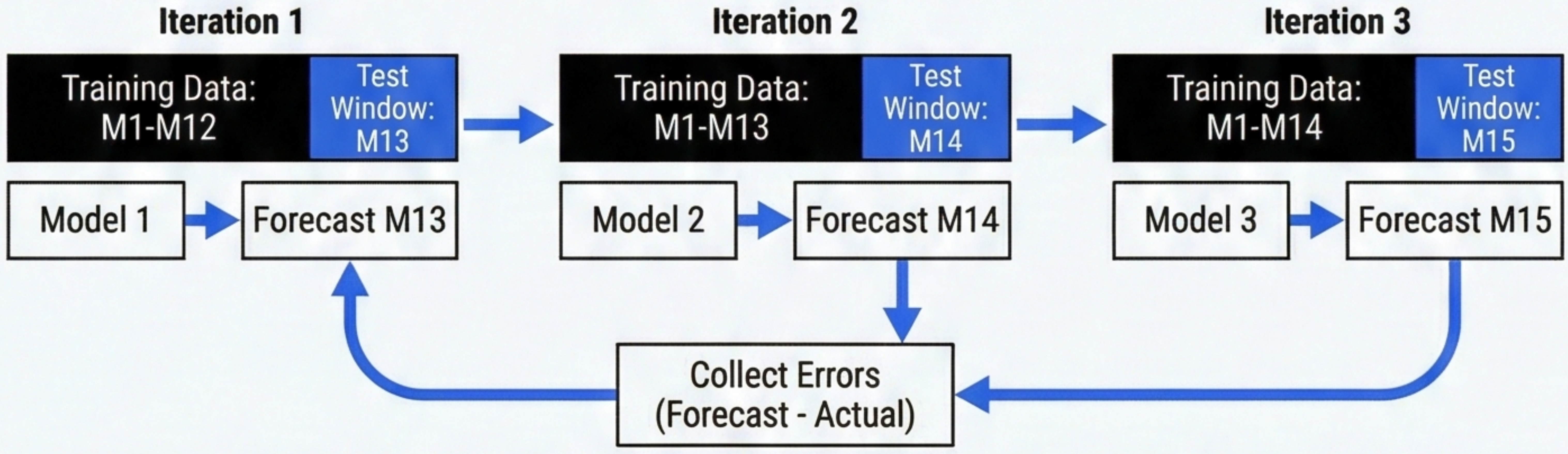
A Retrospective Evaluation Method to Assess Forecast Model Performance on Historical Data

1. CORE CONCEPT: RETREASPECTIVE SIMULATION



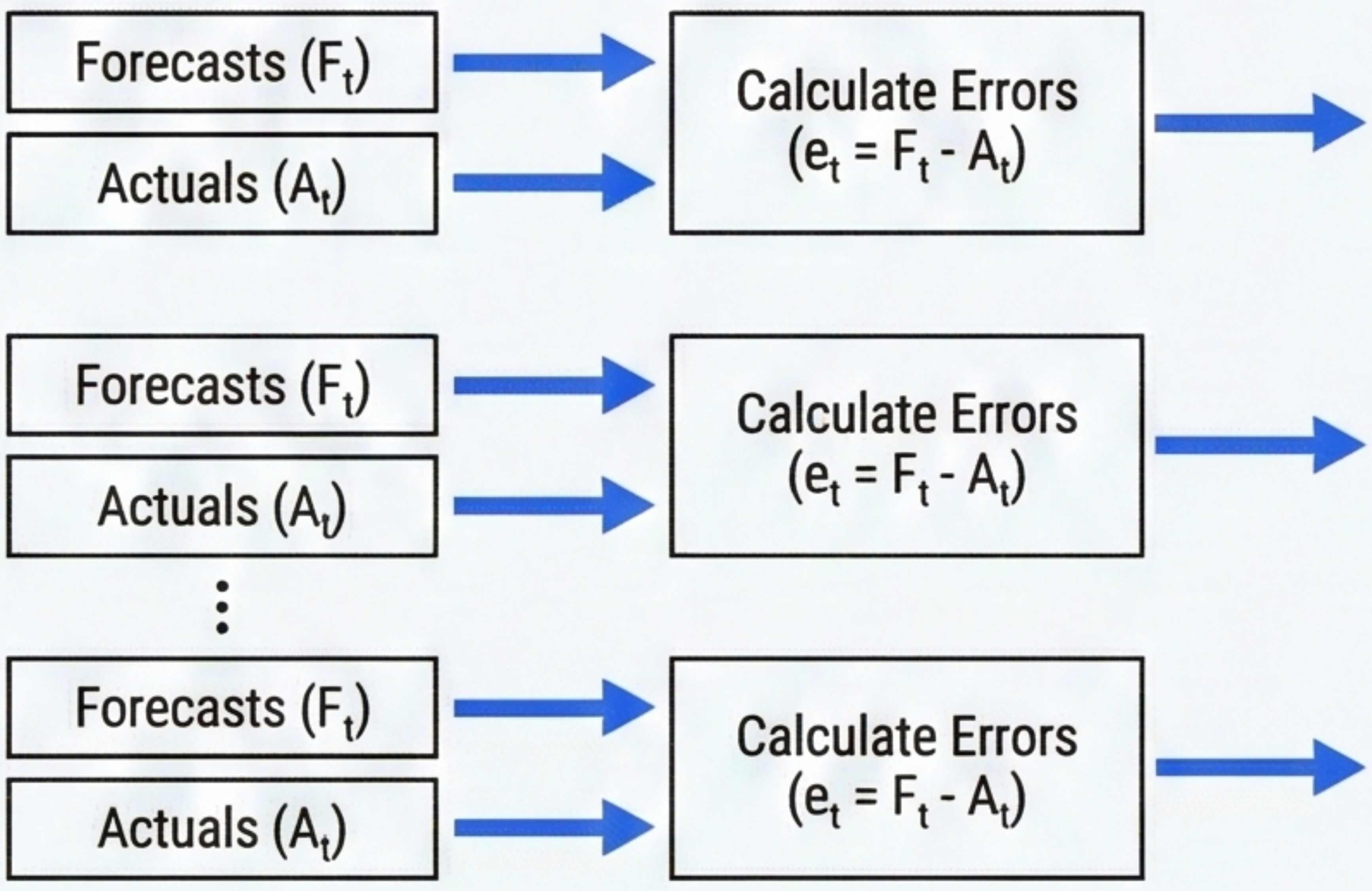
Backtesting simulates past conditions by retraining the model on prior data and evaluating its predictions against known, subsequently observed outcomes.

2. ROLLING WINDOW APPROACH (STEP-BY-STEP FLOW)



The training window expands or shifts over time, and a forecast is generated for the immediate next period(s) in each step. The process repeats until the end of the available data.

3. EVALUATION METRICS & AGGREGATION



MAE (Mean Absolute Error)

$$MAE = \frac{1}{n}\sum |e_t|$$

Average magnitude of errors without considering direction.

RMSE (Root Mean Squared Error)

$$RMSE = \sqrt{\frac{1}{n} * \sum (e_t)^2}$$

Penalizes larger errors more severely.

MAPE (Mean Absolute Percentage Error)

$$MAPE = \frac{1}{n} * \sum \left|\frac{e_t}{A_t}\right| * 100\%$$

Expresses error as a percentage of actual values.

⚠ **Alert:** Undefined if $A_t = 0$.

4. PURPOSE & LIMITATIONS

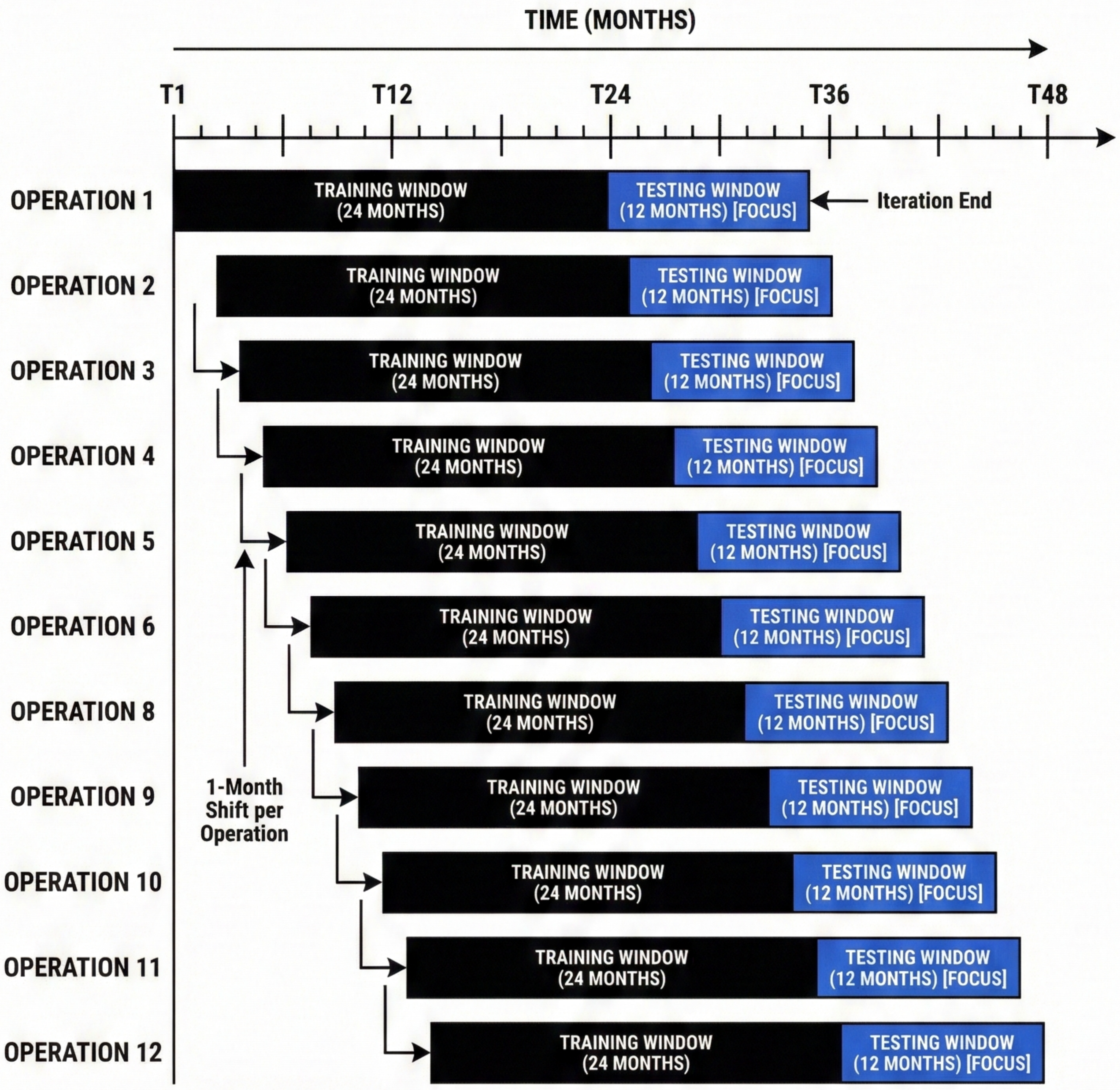
PURPOSE	LIMITATIONS
Model Selection: Compare different forecasting models (e.g., ARIMA, ETS) on the same historical data.	⚠ **Alert: Historical Bias:** Assumes past patterns will continue into the future.
Parameter Tuning: Optimize model hyperparameters to minimize backtest errors.	**Data Leakage Risk:** Ensure future information is strictly excluded from training data at each step.
Performance Estimation: Provide a realistic estimate of forecast accuracy on unseen future data.	**Computational Cost:** Requires retraining the model many times, which can be resource-intensive.

Zbigniew Galar

ADVANCED BACKTESTING: ROLLING WINDOW CONCEPT

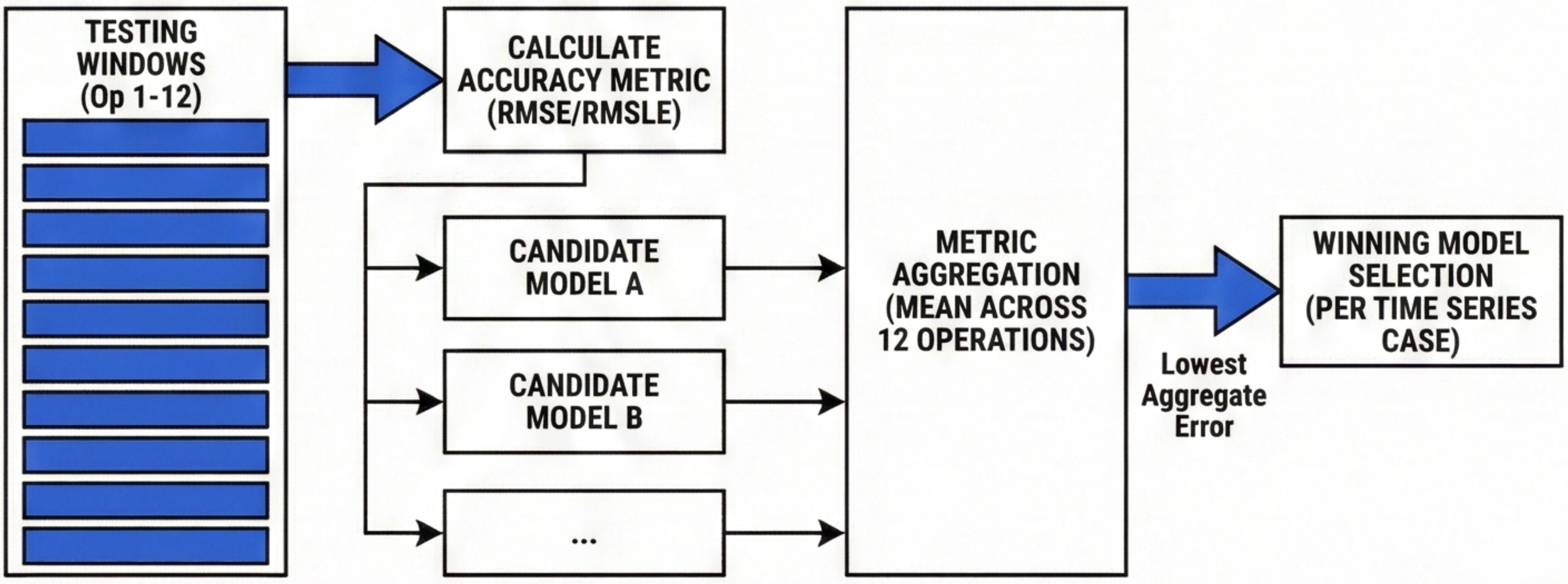
12-Month Testing Window, 24-Month Training Window, 12 Iterations with 1-Month Shift.
Minimizes Seasonality Bias for Model Selection based on Aggregate RMSE/RMSLE.

TIMELINE & WINDOWS STRUCTURE (12 ITERATIONS)



Note: Shifting window across full cycle removes seasonality influence on aggregate performance.

EVALUATION & MODEL SELECTION FLOW



Zbigniew Galar