$\boxed{\text{Milestone 1}}$

State-Space Models for Multi-Horizon Financial Forecasting

Group 3 - DS and AI Lab

October 3, 2025

1 Problem Definition

High-level goal. Develop a *state-space* Echo State Network (ESN) for multi-horizon forecasting of financial time series (equity/index returns or prices), with rigorous evaluation across point, probabilistic, and decision-aware metrics under leakage-safe backtesting.

Observed data and targets. Let $s_t \in \mathbb{R}^{d_{\text{obs}}}$ be market features at (discrete) time t (e.g., log-returns, volumes, realized volatility, calendar flags), and $u_t \in \mathbb{R}^{d_{\text{exo}}}$ exogenous covariates (e.g., macro factors). Define targets y_{t+h} for $h \in \mathcal{H}$ (e.g., $\{1, 5, 20\}$ steps ahead).

State-space ESN (latent dynamics and observation). We adopt a fixed, contractive reservoir as latent transition; the readout constitutes the observation equation:

$$x_{t+1} = (1-\alpha)\, x_t + \alpha\, \phi\big(W x_t + U u_t + b\big) + \xi_t, \quad x_t \in \mathbb{R}^N, \quad \rho(W) < 1, \tag{1}$$

$$\hat{y}_{t+h} = C_h x_t + D_h u_t + \varepsilon_{t,h}, \qquad h \in \mathcal{H}. \tag{2}$$

Here ϕ is 1-Lipschitz (e.g., $\tanh$), $\alpha \in (0,1]$ is the leak rate, $\rho(W)$ is the spectral radius enforcing the echo-state property [1, 2, 3]. Training is *readout-only*: (C_h, D_h) are learned via ridge or quantile regression while (W, U, b, α, ϕ) are fixed by design.

Learning objective. For point forecasts we minimize regularized MSE; for probabilistic forecasts we add pinball loss for quantile levels $\mathcal{A} \subset (0,1)$:

$$\begin{aligned} \min_{\{C_h, D_h\}} \quad & \sum_t \sum_{h\in\mathcal{H}} \Big(\|\hat{y}_{t+h} - y_{t+h}\|_2^2 + \lambda \|C_h\|_F^2 + \lambda \|D_h\|_F^2 \Big) \\ & + \sum_{\alpha\in\mathcal{A}} \sum_t \sum_{h\in\mathcal{H}} \underbrace{\big(\alpha - \mathbf{1}\{y_{t+h} < \hat{q}_{\alpha,t,h}\}\big)\,(y_{t+h} - \hat{q}_{\alpha,t,h})}_{\text{pinball loss}}. \end{aligned} \tag{3}$$

Task protocol and constraints. All preprocessing (scalers, feature construction) is fit *only* on training windows; we use strict walk-forward (rolling or expanding) splits to avoid look-ahead bias [6, 7]. Transaction costs and slippage are incorporated into decision metrics.

2 Objectives (Measurable)

1. **Forecast accuracy.** Achieve statistically significant improvements in RMSE/MAE/sMAPE and directional accuracy vs. strong baselines.
2. **Calibration.** Deliver calibrated predictive intervals: nominal coverage within ±2% for 80/95% intervals; low CRPS [5].
3. **Decision-aware utility.** Positive net average P&L and Sharpe with realistic fees; controlled drawdowns; low turnover for robustness.

4. **Ablation discipline.** Quantify effects of $\rho(W)$, sparsity, reservoir size N, leak α, and exogenous inputs u_t.
5. **Reproducibility.** Deterministic seeds, config-logged experiments, published code and splits.

3 Evaluation Setup

3.1 Datasets and Data Sources (Open)

We will assemble series from openly accessible sources:

- **Yahoo Finance** (end-of-day prices, adjusted close, volumes) [12].
- **Stooq** (free historical equities/FX/indices; CSV) [13].
- **Alpha Vantage** (free API tier; equities/FX/crypto) [14].
- **FRED** (macro factors for u_t: rates, CPI, industrial production, VIX proxy when available) [15].

Primary targets use log-returns $r_t = \log P_t - \log P_{t-1}$; multi-horizon returns $r_{t\to t+h} = \sum_{k=1}^{h} r_{t+k}$.

3.2 Metrics

Point: RMSE, MAE;

Directional: $\Pr[\mathrm{sign}(\hat{r}_{t\to t+h}) = \mathrm{sign}(r_{t\to t+h})]$;

Probabilistic: interval coverage & width [5];

4 Literature Review

4.1 Echo State Networks and Reservoir Computing

ESNs implement recurrent dynamics via a fixed, randomly connected reservoir with trainable linear readouts; under contractivity (spectral radius < 1) they exhibit the echo-state property [1]. Leaky-integrator ESNs improve memory-depth control and stability [2]. The broader reservoir computing (RC) paradigm covers ESNs and physical reservoirs; surveys highlight computational efficiency and robustness to noise [3].

4.2 State-Space Models and Modern Sequence Models

Classical linear Gaussian state-space models (SSMs) and Kalman filtering/smoothing are foundational for time-series forecasting [4]. Recent deep SSMs (e.g., structured state-space sequence models) aim to combine long-range memory with stable training [9]. We adopt a *hybrid* stance: keep the latent transition *fixed* (ESN) to reduce variance/overfitting in noisy financial regimes, while learning only low-variance linear heads.

4.3 Financial Time Series Forecasting: Practices and Pitfalls

Robust evaluation requires leakage-safe splits, appropriate stationarity transforms (returns vs. prices), and correction for multiple trials to avoid backtest overfitting [7]. Proper scoring rules (CRPS) and coverage checks assess probabilistic quality [5]. Surveys of deep learning in finance/time series offer baselines (ARIMA/VAR; MLP/RNN/LSTM/GRU; light Transformers) and caution on non-stationarity [10, 11].

4.4 Baselines and Benchmarks (to be implemented)

- **Classical:** ARIMA/ARIMAX; Linear SSM (Kalman); Ridge/LASSO on lagged features.
- **Neural:** MLP (lags); LSTM/GRU (small capacity); Temporal Convolution; lightweight Transformer.
- **Ablative ESN:** vary $\rho(W) \in \{0.6, 0.8, 0.95\}$, sparsity $\in \{5\%, 10\%, 20\%\}$, $N \in \{200, 500, 1000\}$; leaky vs. non-leaky; with/without u_t.

Benchmarks use identical walk-forward splits and cost assumptions.

5 Gaps and Opportunities

1. **Variance control via fixed dynamics.** Many DL models over-parameterize dynamics for noisy financial series. ESNs fix W, reducing estimation variance while retaining nonlinearity—a principled bias–variance trade-off not fully explored in recent finance benchmarks.
2. **SSM framing of ESNs.** Casting ESNs explicitly as state-space models (1)–(2) clarifies assumptions (contractivity, leak) and supports probabilistic outputs (quantiles), bridging RC and forecasting literature.
3. **Decision-aware, leakage-safe evaluation.** Unified assessment across point, probabilistic, and trading metrics with DM tests and cost-adjusted backtests remains inconsistently applied; our protocol aims to be end-to-end reproducible and statistically sound.
4. **Exogenous signals.** Systematic integration of macro factors (FRED) into u_t and analysis of their incremental value across horizons is under-documented in ESN studies.

Data/Modality Tags: DL · Time-Series · Finance · SSM · Reservoir Computing · Probabilistic ML

References

[1] H. Jaeger, *The "Echo State" Approach to Analysing and Training Recurrent Neural Networks*, GMD Report 148, 2001.

[2] H. Jaeger, M. Lukoševičius, D. Popovici, and U. Siewert, "Optimization and applications of echo state networks with leaky-integrator neurons," *Neural Networks*, 20(3):335–352, 2007.

[3] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, 3(3):127–149, 2009.

[4] J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*, 2nd ed., Oxford Univ. Press, 2012.

[5] T. Gneiting and A. E. Raftery, "Strictly Proper Scoring Rules, Prediction, and Estimation," *JASA*, 102(477):359–378, 2007.

[6] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed., OTexts, 2021. https://otexts.com/fpp3/

[7] D. H. Bailey, J. Borwein, M. López de Prado, and Q. J. Zhu, "Pseudo-mathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance," *Notices of the AMS*, 61(5):458–471, 2014.

[8] F. X. Diebold and R. S. Mariano, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3):253–263, 1995.

[9] A. Gu, T. Goel, A. Gupta, et al., "Efficiently Modeling Long Sequences with Structured State Spaces," *ICLR*, 2022. (S4)

[10] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied Soft Computing*, 90:106181, 2020.

[11] S. Gu, B. T. Kelly, and D. Xiu, "Empirical Asset Pricing via Machine Learning," *The Review of Financial Studies*, 33(5):2223–2273, 2020.

[12] Yahoo Finance Historical Data. https://finance.yahoo.com/

[13] Stooq: Free historical market data. https://stooq.com/

[14] Alpha Vantage API for financial time series. https://www.alphavantage.co/

[15] Federal Reserve Economic Data (FRED), Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/