computer sciences and mathematics forum

MDPI

Article

Refinements of Reservoir Neural Networks for Chaotic Time-Series Forecasting †

Roman Vysotskiy [1,*], Shahd Hammoud [2] and Anton Kovantsev [1]

[1] AI Institute (Institute for Artificial Intelligence), ITMO University, Kronverksky Pr. 49, St. Petersburg 197101, Russia; ankovantcev@itmo.ru
[2] Higher Institute for Applied Sciences and Technology, Damascus P.O. Box 31983, Syrian Arab Republic; shahd.m.hammoud2002@gmail.com
* Correspondence: rvysotskiy@itmo.ru
† Presented at the 12th International Conference on Time Series and Forecasting, Canaria, Spain, 15–17 July 2026.

Abstract

Forecasting chaotic and nonlinear time series in financial and economic domains remains an open challenge, as conventional statistical and machine learning methods struggle to capture their complex dynamics. We propose ESN-F, an Echo State Network enhanced with a Fourier transform layer and polynomial feature expansion, designed to improve forecasting accuracy for processes exhibiting chaotic, non-stationary, and non-normal behaviour. The model is evaluated on 1500 series drawn from the M4 Competition dataset and on a real-world Moscow real estate price index. On the M4 subset, ESN-F achieves a mean MAPE of 3.44% on predictable series and 5.26% on chaotic series, outperforming standard ESN, Light Gradient-Boosting Machine (LGBM), Prophet, and Singular Spectrum Analysis (SSA). On the real estate index, ESN-F reaches a cross-validated mean MAPE of 2.56%, compared with 3.19% for ESN and 7.18% for LGBM. In addition, we show that a set of dynamical features—including the Hurst exponent, Kolmogorov–Sinai entropy, correlation dimension, largest Lyapunov exponent, and dominant Fourier harmonics—can identify time series for which reservoir computing is most beneficial, enabling model selection prior to training. These results confirm that the proposed Fourier-augmented reservoir architecture provides a practical and computationally efficient approach to chaotic time-series forecasting.

Keywords: time-series; chaotic dynamics; reservoir computing; Fourier transformation; forecasting

Check for updates

Academic Editors: Olga Valenzuela, Fernando Rojas, Luis Javier Herrera, Hector Pomares and Ignacio Rojas

Published: 10 October 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

Chaotic and nonlinear time series arise naturally in financial and economic systems, where sensitivity to initial conditions, irregular scaling, and non-normal return distributions make reliable forecasting exceptionally difficult [1,2]. Standard statistical models such as ARIMA assume linearity and normality that these series routinely violate, while most deep learning methods require large training sets and substantial computational resources. The practical demand for accurate long-term predictions in domains such as asset pricing and real estate nevertheless persists, motivating the search for methods that can exploit the specific structure of chaotic dynamics.

Reservoir computing, and in particular the Echo State Network (ESN), is a recurrent neural network paradigm that maintains a fixed, randomly initialised internal reservoir

and trains only a linear readout layer [3]. This design combines rich nonlinear dynamics with very low training cost, making it attractive for chaotic systems. However, standard ESNs do not explicitly extract spectral or polynomial features from the input, which may limit their accuracy on quasi-periodic and strongly nonlinear processes. Kovantsev and Gladilin [4] further showed that predictability varies widely across time series and that selecting the right model class before fitting can substantially reduce forecasting error.

In this paper, we address both gaps. We propose ESN-F, an Echo State Network augmented with a Fourier transform layer and polynomial feature expansion, and we investigate which time-series characteristics determine when reservoir computing outperforms alternatives. Our main contributions are: (i) the ESN-F architecture that enriches reservoir states with explicit frequency-domain and polynomial representations; (ii) an empirical evaluation on 1500 series from the M4 Competition dataset and on a real-world Moscow real estate price index; and (iii) a characterisation of the dynamical features—Hurst exponent, Kolmogorov–Sinai entropy, correlation dimension, largest Lyapunov exponent, and dominant Fourier harmonics—that predict whether ESN-F will outperform gradient-boosting and other baselines.

The remainder of the paper is organised as follows. Section 2 reviews related work on chaotic systems and reservoir computing. Section 3 describes the ESN-F architecture. Section 4 presents the datasets and predictability features. Section 5 reports experimental results. Section 6 concludes.

2. Related Works

Research on chaotic systems established the theoretical basis for forecasting complex processes. Lorenz [1] showed how sensitivity to initial conditions limits long-term predictability, while Mandelbrot [2] connected irregular scaling and fractal structure to natural and financial series. Broad overviews by Loskutov [5] and Kapitsa et al. [6] emphasised the ubiquity of nonlinear interactions in complex systems. Quantitative descriptors then became central, including the Hurst effect [7], the Moses effect [8], and entropy-based measures such as Kolmogorov–Sinai entropy. Recent work improved KSE estimation for high-dimensional data [9], discussed practical limitations [10], and generalised the concept for broader use [11]. At the same time, attractor-based deep learning approaches have shown promise for modelling chaotic trajectories [12].

Machine learning methods are increasingly used to model nonlinear and chaotic series. Reviews of financial forecasting confirm that deep models can capture complex dependencies [13], while economic studies based on nonlinear dynamics and bifurcations illustrate how instability emerges in real systems [14,15]. Among neural approaches, reservoir computing is especially attractive because it combines rich internal dynamics with limited training cost; its advantages over conventional neural networks in financial applications were highlighted by Georgopoulos et al. [3]. Related studies improved sequence modelling with chaos-informed LSTMs [16], new prediction schemes for chaotic processes [17], and Fourier-based architectures capable of representing both simple and complex periodic structures [18].

A practical forecasting workflow also requires diagnosing predictability before model selection. Kovantsev et al. [4] showed that features such as correlation dimension, Kolmogorov–Sinai entropy, and noise measures help match time series to appropriate models, substantially reducing selection effort. A broader review of predictability analysis for time series is given by Bezbochina et al. [19]. These studies motivate our use of dynamical features together with reservoir computing for chaotic forecasting.

Taken together, this body of work motivates the choice of the ESN as the core architecture: its fixed random reservoir provides inherently rich nonlinear dynamics suitable for

https://doi.org/10.3390/cmsf2026016020

chaotic systems, its training cost is orders of magnitude lower than that of deep recurrent networks (since only the linear readout is fitted), and prior work [3] has already demonstrated its competitiveness in financial forecasting. The gap we address is the absence of explicit frequency-domain and polynomial representations in standard ESNs, which we hypothesise limits accuracy on quasi-periodic and strongly nonlinear processes.

3. Design of Reservoir Network with Fourier Transformation Layer and Polynomial Features

The ESN-F model is an enhancement of the standard ESN specifically formulated to improve predictive accuracy on chaotic and nonlinear time series by integrating reservoir computing dynamics with polynomial and Fourier-based feature expansions. The core of the model is the reservoir, a large and sparsely connected recurrent network. Its internal state $s_t \in \mathbb{R}^N$ evolves according to nonlinear Equation (1):

$$s_t = (1-\alpha)s_{t-1} + \alpha \tanh\big(W_{\text{in}}[1; x_t] + W_{s_{t-1}}\big), \tag{1}$$

where $\alpha \in (0,1]$ is the leaking rate, $W_{\text{in}} \in \mathbb{R}^{N\times(d+1)}$ is the input weight matrix connecting the input vector $x_t \in \mathbb{R}^d$, and $W \in \mathbb{R}^{N\times N}$ is the reservoir's internal recurrent weight matrix, typically sparse and scaled by its spectral radius to ensure appropriate dynamic behaviour. The spectral radius controls the stability of the reservoir state and balances the retention of past information against the incorporation of new inputs. The hyperbolic tangent ($tanh(\cdot)$) introduces nonlinearity and helps capture complex temporal dynamics.

In order to meticulously encapsulate nonlinear interrelationships within the input data, polynomial feature expansions of the input vector x_t are calculated (2). With respect to a specified polynomial degree p, the resulting polynomial feature vector $\phi(x_t)$ encompasses all monomial terms of the input features up to the degree p:

$$\phi(x_t) = \text{Poly}_p(x_t), \tag{2}$$

where $\text{Poly}_p(\cdot)$ denotes the polynomial expansion operator. This expansion enhances the model's ability to capture smooth nonlinear dependencies.

Additionally, Fourier-based expansions explicitly extract periodic features from the input. For each frequency component $k = 1, 2, \ldots, K$, Fourier characteristics $\psi(x_t)$ are computed from the discrete Fourier transform coefficients as (3):

$$\psi(x_t) = \big[a_k, b_k\big]_{k=1}^{K}, \tag{3}$$

where the discrete Fourier transform coefficients are defined by

$$a_k = \frac{2}{N}\sum_{n=0}^{N-1} x_n \cos\left(\frac{2\pi k n}{N}\right), \qquad b_k = \frac{2}{N}\sum_{n=0}^{N-1} x_n \sin\left(\frac{2\pi k n}{N}\right), \tag{4}$$

where N is the window length and x_n are the values of the input time series on this window. Here, K represents the number of retained Fourier harmonics, and the corresponding magnitude is computed as $A_k = \sqrt{a_k^2 + b_k^2}$, enabling the model to capture periodic or quasi-periodic patterns in the data.

All these different features—the reservoir states s_t, polynomial features $\phi(x_t)$, and Fourier features $\psi(x_t)$—are concatenated into a single feature vector (5):

$$H_t = \big[s_t; \phi(x_t); \psi(x_t)\big], \tag{5}$$

https://doi.org/10.3390/cmsf2026016020

The feature vector H_t is then standardised component-wise to yield a transformed vector ${H_t}^s$. Specifically, each element $H_t^{(j)}$ is scaled as $H_t^{s(j)} = \left(H_t^{(j)} - \mu_j\right)/\sigma_j$, where μ_j and σ_j are the mean and standard deviation of the j-th feature computed over the training window. A Ridge regression readout then maps these transformed features linearly to produce the final prediction $\widehat{y}_t$ (6):

$$\hat{y}_t = W_{\text{out}} {H_t}^s. \tag{6}$$

Here, W_{out} are the regression weights learned, obtained by minimising a regularised least-squares criterion during training.

We adopt an iterative (recursive) multi-step forecasting strategy. At each step beyond the training horizon, the model's own prediction $\widehat{y}_t$ is appended to the input window and used as the observation for the next step. Concretely, the reservoir state is warm-started by running the network on the training data without weight updates, and then the readout W_{out} is applied to generate forecasts step by step. This strategy most faithfully mimics real-world forecasting conditions, where future observations are unavailable, and allows us to assess the model's ability to sustain coherent dynamics over many time steps. The training procedure minimises a Ridge-regularised least-squares objective, so only W_{out} is fitted; all reservoir weights W and W_{in} remain fixed after initialisation.

4. Dataset and Time-Series Predictability

4.1. Dataset

For our investigation we use the "M4 Forecasting Competition Dataset" (The Makridakis competition on Kaggle for benchmarking modern ML methods for forecasting: https://www.kaggle.com/datasets/yogesh94/m4-forecasting-competition-dataset (accessed on 12 April 2026)). This dataset consists of many real-life time series with varying frequencies, including daily, monthly, quarterly, etc. Each time series relates to topics such as micro, industry, macro, finance, demographic, and others. In total, there are 100,000 time series, but for our study, we used only 1500 randomly selected series with a length of 1000 points, as many of them have missed values.

4.2. Predictability Metrics

Because reservoir computing is not a universal forecasting technique, it is important to identify in advance the time series for which it is likely to be productive. We characterise each series using a set of features that quantify predictability and the degree of chaotic behaviour in the underlying dynamical system. Most of these features were identified in [4] as effective indicators for model selection.

The Hurst exponent represents a methodology to quantify autocorrelation and the long-term memory characteristics of time series (7).

$$H = \frac{\frac{\ln R(\tau)}{S(\tau)}}{\ln(\alpha\tau)} \tag{7}$$

where $\tau \in [3, N]$, $\alpha = 0.5$ is the Anis–Lloyd constant; $R(\tau)$ is a maximum difference between the accumulated deviations of the time series from the average in the subinterval of length n, which can be expressed by (8); and $S(n)$ is the standard deviation of the time series in the subinterval of length n, represented by (9):

$$R(\tau) = \max_{1 \le t \le \tau} \sum_{i=1}^{t} (x_i - \overline{x}_t) - \min_{1 \le t \le \tau} \sum_{i=1}^{t} (x_i - \overline{x}_t), \tag{8}$$

$$S(\tau) = \sqrt{\frac{1}{\tau}\sum_{t=1}^{\tau}(x_t - \overline{x}_\tau)^2}. \tag{9}$$

Kolmogorov–Sinai entropy (KSE) is a measure of the chaotic component and complexity of a dynamic system. It is calculated as a metric invariant of the dynamical system generating the time series. Following [4], this entropy can be calculated based on the entropy rate of finite partitions of the state space of the series. Given finite k embeddings of the phase space M of the dynamical system generating the time series $\xi = \{C_{(1)}, C_{(2)}, ..., C_{(k)}\}$, one can calculate the probability measure of the i-element, $\mu(T^{-1}C_{i_1} \cap \cdots \cap T^{-n}C_{i_n})$, where T is a measurable transformation of the phase space M onto itself with preservation of the measure μ, $n \in (-\infty, \infty)$ and $i \in [1, k]$. The entropy value $h_\mu\ (T, \xi)$ is a limit that can be estimated as follows (10):

$$\begin{aligned} h_\mu(T, \xi) = -\lim_{n\to\infty}\frac{1}{n}\sum_{i_1,\ldots,i_n}\mu(T^{-1}C_{i_1} \cap \cdots \cap T^{-n}C_{i_n}) \\ \times \ln\mu(T^{-1}C_{i_1} \cap \cdots \cap T^{-n}C_{i_n}). \end{aligned} \tag{10}$$

For ${h_\mu}^{KS}(T)$ estimation, we take the upper bound of the entropy rate $h_\mu(T, \xi)$ over all finite partitions (11):

$${h_\mu}^{KS}(T) = sup_\xi h_\mu(T, C). \tag{11}$$

Since considering all possible partitions presents certain technical difficulties, for our practical needs, we estimate Kolmogorov–Sinai entropy as per Shiozawa and Tokuda [9] by a partitioned estimator for KSE in high-dimensional systems.

Correlation dimension is a measure of the dimensional properties of the space occupied by a set of random points; it is also often described as a form of fractal dimension. This measurement is more robust when the dataset contains a small number of points, and it often agrees with other dimensional computations. The correlation dimension is represented by the slope of the logarithmic graph of the correlation integral and can be determined using the following limit (12):

$$d_k = \lim_{r\to 0}\lim_{m\to\infty}\frac{\ln C(r)}{\ln r}, \tag{12}$$

where the correlation integral is estimated by (13):

$$C(r) = \sum_{i=1}^{m}\sum_{j=i+1}^{m}\frac{\theta(r - \rho(i, j))}{m(m-1)}, \tag{13}$$

where $\rho_k(i, j) = \sqrt{\sum_{l=1}^{k}(x_{i-k+l} - x_{j-k+l})^2}$ is the Euclidean distance between the delay vectors, r is the cell size of the characteristic phase space, m is the number of delay vectors in the reconstructed phase space and $\theta(x)$ is the Heaviside function (14):

$$\theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{14}$$

Fourier transform helps characterise the internal structure of a time series through the amplitudes and phases of its frequency components. In this study, we use a single Fourier-based feature: the number of harmonics whose amplitudes are not smaller than the mean amplitude over all frequencies. This value, denoted by N_{Fh}, is computed as (15).

$$N_{Fh} = \sum_{i=1}^{k} \theta(A_i - \bar{A}), \tag{15}$$

where A_i is the amplitude of the i-component of Fourier transform, $\bar{A}$ is the mean and θ is the Heaviside function (14), which is equal to 1 if its argument greater than 0 and 0 otherwise.

The noise measure feature is predicated upon the comparison of the standard deviation of a time series to that of its first-order differences, as articulated below (16). The objective is to identify high-frequency noise that amplifies the relative deviation of value alterations at each temporal increment; thus, a higher value of this metric signifies a reduced influence of noise on the time series. This metric was successfully used in [4].

$$F_N = 1 - \sqrt{\frac{N \times \sum_{i=1}^{N-1}(x'_i - x')^2}{(N-1) \times \sum_{i=1}^{N}(x'_i - x')^2}}. \tag{16}$$

Lyapunov exponents quantify the rate of divergence of nearby trajectories in phase space. From a geometric perspective, they delineate the extent of stretching and contraction along stable and unstable directions. We use the maximal Lyapunov exponent (MLE) as formulated in Loskutov [20]. A positive MLE indicates exponential divergence of trajectories and serves as a hallmark of chaotic dynamics; values near zero suggest marginal or periodic behaviour.

5. Experimental Results and Discussion

We conducted two experiments. In the first, we compare ESN-F against four baselines—ESN, LGBM, SSA, and Prophet—on 1500 randomly selected M4 series of length 1000. Each series is split into a training prefix and a 15-step test horizon; the models predict iteratively (feeding their own outputs back as inputs) and MAPE is computed on the held-out steps. We chose 15 steps as the evaluation horizon because pilot experiments showed that iterative prediction errors accumulate noticeably at this length, making model differences clearly visible. Model accuracy is estimated via a rolling-origin evaluation: for each series, the cut-off point is shifted 16 times (one step per iteration), the model is retrained from scratch at each origin, a 15-step-ahead forecast is generated iteratively, and the reported MAPE is the mean over all 16 runs. The ESN and ESN-F models use a reservoir of $N = 100$ nodes, spectral radius $\rho = 0.9$, connection sparsity of 10%, and leaking rate $\alpha = 1.0$ (equivalent to a standard ESN update without memory decay). The Ridge regularisation coefficient is $\lambda = 1.0$, the polynomial expansion degree is $p = 2$, and the number of retained Fourier harmonics is $K = 5$. MAPE distributions across all 1500 series are shown in Figure 1; ESN-F achieves the lowest median error, with LGBM as its closest competitor.

For the next experiment, we divide the dataset into two clusters according to their time-series features, applying the approach described in [4]. These features, which we treat as predictability characteristics, are the Hurst exponent, noise factor, correlation dimension, largest Lyapunov exponent, Kolmogorov–Sinai entropy, and number of prevailing Fourier harmonics, all described in Section 4.2. The clustering yields a *Predictable* group of 1176 series and a *Chaotic* group of 324 series that exhibit signs of complex, hard-to-predict dynamics. Table 1 reports mean MAPE for all five models in each group.

Several observations stand out from Table 1. On the Predictable cluster, all three reservoir and boosting models achieve low errors (3.4–3.7%), confirming that these series are well-behaved. Prophet performs roughly twice as poorly (6.86%), consistent with its design for trend-plus-seasonality patterns rather than complex dynamics; SSA is the weakest baseline (18.03%) in this group, likely because its window-based decomposition is sensitive

to the varied spectral structure of M4 series. On the Chaotic cluster, the gap between ESN-F and SSA widens further (5.26% vs. 20.05%), demonstrating that decomposition-based methods deteriorate sharply on irregular series, while reservoir networks retain robustness. Notably, plain ESN edges out ESN-F by 0.07 percentage points on the Chaotic cluster, suggesting that the Fourier and polynomial features provide marginal overhead on the most erratic series; this trade-off is further examined below.

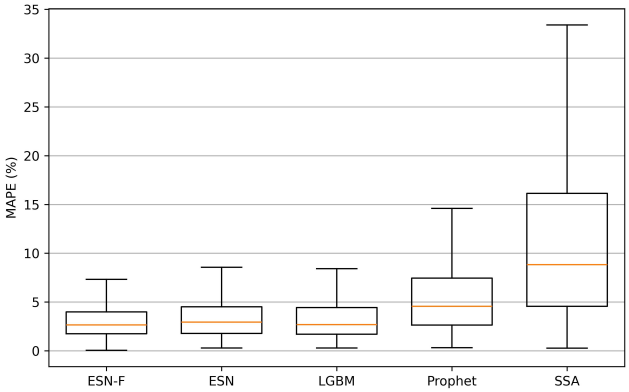


Figure 1. Comparison of MAPE distributions for different forecasting models. Orange lines denotes the median values.

Table 1. Mean MAPE (%) values for the models in different data clusters. Predictable: A total of 1176 series assigned by the clustering procedure of [4] to the low-complexity group; Chaotic: A total of 324 series assigned to the high-complexity group. Bold indicates the best result per row.

Cluster	ESN-F	ESN	LGBM	Prophet	SSA
Predictable	**3.44**	3.56	3.72	6.86	18.03
Chaotic	5.26	**5.19**	5.39	8.57	20.05

We focus on the 324 Chaotic series to understand when ESN-F specifically outperforms or underperforms LGBM. We treat a difference of more than one percentage point as practically significant. ESN-F shows better results in 27% of cases and LGBM is better in 15%; the remaining 58% are effectively tied. We then examine the time-series features of the subsets in which each model excels. Their distributions are shown as box plots in Figure 2.

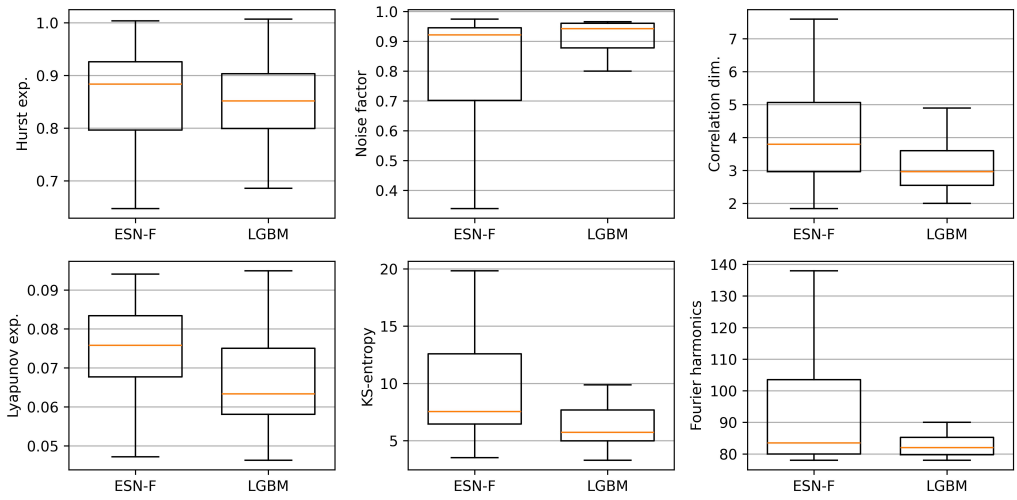


Figure 2. Time-series feature distributions of the bad predictability cluster, where the models performed with better results.

We may notice there that ESN-F manages to operate with a series with a wider range of characteristics, especially the Hurst exponent, noise, and Kolmogorov–Sinai entropy. The fact that the series with the Hurst exponent close to 0.5 performs better for ESN-F confirms our hypothesis that reservoir networks might be good for processes with chaotic dynamics because this value is considered the indicator of chaos in the dynamic system.

During our research, we encountered an opportunity to apply the proposed methodology to a real-world task: forecasting real estate prices in Moscow (the Moscow Real Estate Domclick Index is a composite index of the Moscow Real Estate market: https://www.moex.com/en/index/MREDC (accessed on 12 April 2026)).

The task was to make long-term forecasts with low error. As we had already compared the model, we chose the three best models according to Table 1. The data consist of 418 points from 28 December 2016 to 25 December 2024 with a weekly frequency. The original series was smoothed using the average value of 13 points (quarter). Due to the small data size, to estimate the models' quality, we used cross-validation with expanding training size. The initial training size was 104 observations and the forecast range was 52 points (one year for weekly sampling). In Figure 3 estimated errors are shown. Our designed model ESN-F outperformed the ESN and LGBM models.

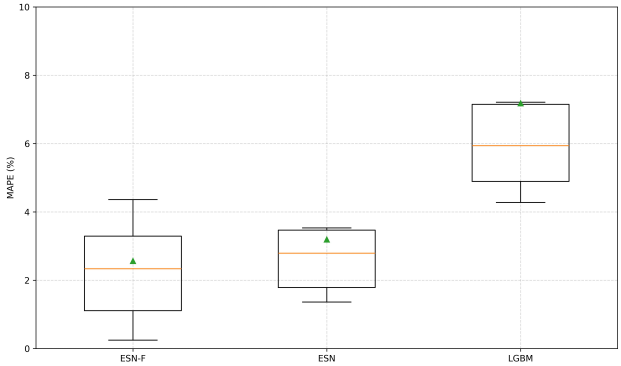


Figure 3. Comparison of MAPE distributions of models' long-term forecast for real estate prices. The triangles show the mean value.

In this forecasting task, LGBM has a bigger difference compared to the ESN-F and ESN models. The chaotic properties of the time series are as follows: Hurst exponent: 0.65; noise factor: 0.99; correlation dimension: 1.33; largest Lyapunov exponent: 0.01; Kolmogorov–Sinai entropy: 1.84; number of prevailing Fourier harmonics: 30. The time series displays weakly chaotic behaviour with persistent memory and low stochastic influence. The moderate Kolmogorov–Sinai entropy and 30 dominant Fourier harmonics indicate complex deterministic–stochastic dynamics, which proves the need for hybrid modelling approaches for accurate forecasting.

The average MAPE cross-validation of each of the models is calculated in Table 2.

Table 2. Mean MAPE (%) values of models in cross-validation for real estate price forecasting. Cross-validation uses an expanding training window; initial train size: 104 weeks; forecast horizon: 52 weeks.

ESN-F	ESN	LGBM
2.56	3.19	7.18

The Python program code with the data examples for these experiments is stored in the GitHub repository: https://github.com/ATHigh/Refinements-of-reservoir (accessed on 16 June 2026).

https://doi.org/10.3390/cmsf2026016020

6. Conclusions

We proposed ESN-F, an Echo State Network augmented with a Fourier transform layer and polynomial feature expansion, and evaluated it on 1500 M4 Competition series and a real-world real estate price index. The experiments show that ESN-F achieves lower forecasting errors than standard ESN and LGBM on series with weakly chaotic, persistent-memory dynamics, while remaining competitive on more regular series. Among the five methods tested—ESN-F, ESN, LGBM, Prophet, and SSA—LGBM is the closest competitor to the reservoir models. Prophet performs roughly twice as poorly on both cluster groups, reflecting its strong prior towards trend-plus-seasonality structure. SSA shows the largest errors overall, particularly on the Chaotic series cluster, which is consistent with its sensitivity to window length and its assumption of quasi-periodic components. These results suggest that reservoir computing occupies a practical middle ground: richer in internal dynamics than boosting methods, yet far cheaper to train than deep recurrent networks.

We also demonstrated that a pre-modelling assessment based on dynamical features—the Hurst exponent, noise factor, correlation dimension, largest Lyapunov exponent, Kolmogorov–Sinai entropy, and number of dominant Fourier harmonics—can identify the time series for which ESN-F is most likely to outperform alternatives. This enables model selection prior to fitting and reduces wasted computation.

On the real estate forecasting task, ESN-F achieves a cross-validated mean MAPE of 2.56%, versus 3.19% for ESN and 7.18% for LGBM, confirming the benefits of the Fourier and polynomial augmentations on series with mixed deterministic–stochastic structure. Future work will address statistical significance testing, ablation of individual feature components, and integration of ESN-F as a specialised forecasting agent within a multi-agent system for socio-economic analysis.

Author Contributions: Conceptualisation, A.K.; methodology, A.K. and R.V.; software, R.V.; validation, S.H. and R.V.; formal analysis, S.H.; investigation, R.V.; data curation, S.H.; writing—original draft preparation, R.V.; writing—review and editing, A.K. and S.H.; supervision, A.K.; project administration, R.V.; funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Economic Development of the Russian Federation (IGK 000000C313925P4C0002), agreement No 139-15-2025-010.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in "The Makridakis competition on Kaggle for benchmarking modern ML methods for forecasting" at https://www.kaggle.com/datasets/yogesh94/m4-forecasting-competition-dataset (accessed on 12 April 2026) and in the "Moscow Real Estate Domclick Index" at https://www.moex.com/en/index/MREDC (accessed on 12 April 2026).

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

Abbreviations

The following abbreviations are used in this manuscript:

ESN	Echo state network (reservoir neural network)
ESN-F	Echo state network with the Fourier layer and polynomial expansion
KSE	Kolmogorov–Sinai entropy
LGBM	Light gradient-boosting machine
MAS	Multi-agent system
SSA	Singular spectrum analysis

References

1. Lorenz, E.N. Dimension of weather and climate attractors. *Nature* **1991**, *353*, 241–244. https://doi.org/10.1038/353241a0.
2. Mandelbrot, B.B.; Hudson, R.L. *(Mis)behavior of Market: A Fractal View of Risk, Ruin and Reward*; Basic Books: New York, NY, USA, 2004.
3. Georgopoulos, S.P.; Tziatzios, P.; Stavrinides, S.G.; Antoniades, I.P.; Hanias, M.P. Reservoir computing vs. neural networks in financial forecasting. *Int. J. Comput. Econ. Econom.* **2022**, *13*, 127283. https://doi.org/10.1504/ijcee.2023.127283.
4. Kovantsev, A.; Gladilin, P. Analysis of multivariate time series predictability based on their features. In *2020 International Conference on Data Mining Workshops (ICDMW)*; IEEE: Piscataway, NJ, USA, 2020; pp. 348–355. https://doi.org/10.1109/ICDMW51313.2020.00055.
5. Loskutov, A.Y. *Problems of Nonlinear Dynamics*; Chaos, I., Ed.; Vestnik Moskovskogo Universita. Ser. 3 Fizika Astronomiya; Moscow University: Moscow, Russia, 2001. (In Russian)
6. Kapitsa, S.; Kurdyumov, S.; Malinetsky, G. *Synergetics and Forecasts of the Future*; Editorial URSS: Moscow, Russia, 2001. (In Russian)
7. Hurst, H.E. Long-Term Storage Capacity of Reservoirs. *Trans. Am. Soc. Civ. Eng.* **1951**, *116*. https://doi.org/10.1061/taceat.0006518.
8. Chen, L.; Bassler, K.E.; McCauley, J.L.; Gunaratne, G.H. Anomalous scaling of stochastic processes and the Moses effect. *Phys. Rev. E* **2017**, *95*, 1. https://doi.org/10.1103/PhysRevE.95.042141.
9. Shiozawa, K.; Tokuda, I.T. Estimating Kolmogorov–Sinai entropy from time series of high-dimensional complex systems. *Phys. Lett. A* **2024**, *510*, 129531. https://doi.org/https://doi.org/10.1016/j.physleta.2024.129531.
10. Dünki, R.M. The estimation of the Kolmogorov entropy from a time series and its limitations when performed on EEG. *Bull. Math. Biol.* **1991**, *53*, 665–678. https://doi.org/10.1007/BF02461547.
11. Amigó, J.M. The equality of Kolmogorov-Sinai entropy and metric permutation entropy generalized. *Phys. D: Nonlinear Phenom.* **2012**, *241*, 789–793. https://doi.org/10.1016/j.physd.2012.01.004.
12. Dutta, A.; Harshith, J.; Ramamoorthy, A.; Lakshmanan, K. Attractor Inspired Deep Learning for Modelling Chaotic Systems. *Hum.-Centric Intell. Syst.* **2023**, *3*, 461–472. https://doi.org/10.1007/s44230-023-00045-z.
13. Jiang, W. Applications of deep learning in stock market prediction: Recent progress. *Expert Syst. Appl.* **2021**, *184*, 115537. https://doi.org/10.1016/j.eswa.2021.115537.
14. Vosvrda, M. Bifurcation routes and economic stability. In *Proceedings of the 7th International Conference of the Society for Computational Economics, New Haven, CT, USA*; Society for Computational Economics (SCE): Dordrecht, The Netherlands, 2001.
15. Hanias, M.; Magafas, L.; Stavrinides, S. "Reverse Engineering" in Econophysics. *Int. J. Product. Manag. Assess. Technol.* **2019**, *7*, 36–49. https://doi.org/10.4018/ijpmat.2019010103.
16. Parreño, S.J. Chaos Theory Enhanced LSTM Model of the Philippine Stock Exchange Index. *Sn Comput. Sci.* **2024**, *5*, 779. https://doi.org/10.1007/s42979-024-03160-1.
17. Jiang, J.; Chen, L.; Ke, L.; Dou, B.; Zhu, Y.; Shi, Y.; Qiu, H.; Zhang, B.; Zhou, T.; Wei, G.W. Machine learning predictions from unpredictable chaos. *arXiv* **2025**, arXiv:2503.14956. https://doi.org/10.48550/arXiv.2503.14956.
18. Dong, Y.; Li, G.; Tao, Y.; Jiang, X.; Zhang, K.; Li, J.; Deng, J.; Su, J.; Zhang, J.; Xu, J. FAN: Fourier Analysis Networks. *arXiv* **2025**, arXiv:2410.02675. https://doi.org/10.48550/arXiv.2410.02675.
19. Bezbochina, A.; Stavinova, E.; Kovantsev, A.; Chunaev, P. Enhancing Predictability Assessment: An Overview and Analysis of Predictability Measures for Time Series and Network Links. *Entropy* **2023**, *25*, 1542. https://doi.org/10.3390/e25111542.
20. Loskutov, A.Y. Fascination of chaos. *Physics-Uspekhi* **2010**, *53*, 1257–1273. https://doi.org/10.3367/UFNe.0180.201012d.1305.

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.