Input space optimization for neural network–based forecasting of non-stationary time series[*]

Oleksandr Kovaliv[1,†], Ievgen Sidenko[1,*,†], Yuriy Zhukov[2,†] and Yuriy Kondratenko[1,†]

[1] Petro Mohyla Black Sea National University, Mykolaiv, Ukraine
[2] C-Job Nikolayev, Mykolaiv, Ukraine

Abstract

Forecasting time series under instability remains a challenging task due to structural discontinuities, nonstationarity, and the influence of external factors. This paper proposes a forecasting approach focused on optimizing the model's input space rather than architectural changes to the forecasting algorithm itself. This study examines the impact of exogenous variables on neural network-based forecasting using data on infectious disease incidence in Ukraine. A baseline neural network forecasting model based solely on lagged target values is compared with models incorporating exogenous information. A systematic procedure for constructing and optimizing the input space is proposed, including evaluating various subsets of exogenous variables for fixed model architecture, training parameters, and forecast horizon. Experimental results demonstrate that the proposed input space optimization strategy leads to a significant reduction in forecast error, and these improvements are achieved without increasing model complexity or changing the training procedure.

Keywords

input space optimization, exogenous variables, infectious diseases; forecasting; time series; machine learning; neural networks

1. Introduction

Economic crises, armed conflicts, pandemics, climate change, and other social upheavals that lead to abrupt changes in the dynamics of observed processes significantly complicate time series forecasting, demonstrating low robustness and limited accuracy when using classical methods. These methods are primarily designed for stationary or quasi-stationary time series and are poorly adapted to structural breaks and external disturbances. Recent research shows that modern approaches to time series forecasting increasingly focus on accounting for external factors and improving model adaptability to non-stationary data [1].

Modern machine learning methods and neural networks [2-6] provide more flexible tools for analyzing complex time dependencies and nonlinear patterns. Neural models are capable of automatically extracting hidden patterns from data and adapting to changing conditions, making them promising for forecasting in unstable environments. In recent years, various neural network architectures for time series forecasting have been actively studied, including recurrent neural networks, temporal convolutional networks, transformers, and specialized deep learning models. However, many existing studies have focused primarily on architectural improvements, while the impact of external factors on forecasting performance remains poorly understood [7].

One of the key reasons for the decline in forecasting accuracy in unstable conditions is the neglect of external factors that significantly influence the dynamics of the processes being studied. Such factors can be represented as exogenous variables external factors that affect the target time series but are not generated by its internal dynamics. Such variables include calendar

---

[*] *CMIS 2026: International Workshop on Computer Modeling and Intelligent Systems, May 5, 2026, Zaporizhzhia, Ukraine*

[1*] Corresponding author.

[†] These authors contributed equally.

kizutoamazed@gmail.com (O. Kovaliv); ievgen.sidenko@chmnu.edu.ua (Ie. Sidenko); yuriy.zhukov@nuos.edu.ua (Y. Zhukov); yuriy.kondratenko@chmnu.edu.ua (Yu. Kondratenko)

0009-0001-2047-6319 (O. Kovaliv); 0000-0001-6496-2469 (Ie. Sidenko); 0000-0002-7454-8007 (Y. Zhukov); 0000-0001-7736-883X (Yu. Kondratenko)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

characteristics, socioeconomic indicators, crisis-related events, and environmental factors. Integrating them into forecasting models expands the information context and improves the model's resilience to sudden changes in conditions.

This paper examines the integration of exogenous variables in public health, specifically the spread of infectious diseases, which is highly sensitive to external factors, including seasonal factors, population mobility, demographic changes, public health measures, and emergencies [8]. Under these conditions, disease dynamics can deviate significantly from historical trends, making traditional forecasting approaches unreliable.

Therefore, optimization in this study is understood not as increasing model complexity, but as the targeted development of the model's input space and the rational use of external information that influences disease spread.

The problem statement. The scientific challenge is to develop optimization-oriented forecasting approaches that improve accuracy and reliability through rational structuring of model inputs and integration of relevant exogenous variables, without increasing model complexity. The application of this approach will be explored in epidemiological forecasting to support evidence-based decision-making in healthcare systems.

Analysis of recent studies and publications. In recent years, research in time series forecasting has increasingly focused on the use of neural network models capable of accounting for nonlinearities, structural shifts, and complex temporal dependencies characteristic of real-world data. Deep learning architectures designed for the direct modeling of temporal signals without explicitly specifying seasonal or trend components have made a significant contribution to the development of this area. For example, the N-BEATS architecture was proposed in [9-11] as a system of residual blocks capable of automatically learning to decompose a time series and, in certain configurations, providing partial interpretability of the forecasted components. The development of this approach has led to the emergence of models aimed at improving the stability and accuracy of long-term forecasting, including N-HiTS, where hierarchical interpolation and multi-level scaling improved scalability and stability over long forecasting horizons [12-14].

In parallel, Transformer-based architectures have also been actively explored; originally introduced for natural language processing tasks [15-16], they were subsequently adapted for time series analysis. Informer proposed a sparse attention mechanism focused on the efficient processing of long sequences. Autoformer extended this approach by integrating explicit time series decomposition and an autocorrelation mechanism to better account for periodicity and recurring patterns in temporal data [17-19]. Frequency-based modifications of Transformer architectures, such as FEDformer, employ spectral representations of time series to improve forecasting performance and robustness in long-horizon prediction tasks [20]. Studies in [21-23], including PatchTST and related token-based approaches, have shown that aggregating temporal observations into larger segments facilitates more robust modeling of long-term dependencies while reducing computational complexity.

Despite significant progress in the development of architectural solutions, the contemporary literature increasingly emphasizes that increasing model complexity does not guarantee a proportional improvement in forecasting quality. Studies [24-26] demonstrate that the effectiveness of neural network models is largely determined by the quality of the input space and the structure of the features used, rather than solely by architectural design. In this context, growing attention is being paid to data-centric and input-oriented approaches focused on optimizing data representation and selecting informative features to improve forecast robustness under non-stationary conditions.

Studies [27-30] demonstrate that the inclusion of exogenous factors (calendar effects, climatic conditions, population mobility indicators, and socioeconomic variables) can significantly improve forecasting accuracy and model robustness in the presence of structural shifts. As a result, architectures explicitly designed to integrate exogenous information have attracted considerable attention. For example, the Temporal Fusion Transformer [31, 32] combines attention mechanisms with the separation of static features, known future inputs, and dynamic temporal components, while ensuring interpretability of individual feature contributions. An extension of the N-BEATS architecture to support exogenous variables (N-BEATSx) confirmed that appropriate integration of external factors can substantially improve forecasting accuracy while maintaining the stability of the base model [33, 34]. More recent approaches, such as TimeXer and related models [35, 36], further formalize forecasting with exogenous variables as a distinct research direction, proposing specialized mechanisms for fusing endogenous and exogenous signals.

In the field of epidemiological forecasting, neural network methods gained widespread adoption during the COVID-19 pandemic, when traditional statistical models proved insufficiently flexible to capture abrupt regime shifts, reporting delays, and weak or unstable seasonality [37-40]. A substantial portion of applied research relies on a limited and heuristically selected set of exogenous factors or focuses primarily on comparing architectural solutions, without addressing the problem of systematic input space optimization [41-43]. Under conditions of high data instability, such approaches limit model interpretability and reduce robustness to changes in the external environment.

Despite significant progress in neural time series forecasting, an unsolved problem remains the development of optimization-oriented approaches that systematically integrate exogenous variables while maintaining model stability and interpretability.

The research goal. Given the limitations of traditional forecasting methods in unstable environments and the insufficient consideration of external factors in existing neural network-based approaches, the aim of this paper is to develop and evaluate an optimization-based framework for forecasting the spread of infectious diseases. The proposed framework focuses on constructing the model's input structure by integrating exogenous variables to improve forecast accuracy and robustness without increasing model complexity.

2. Theoretical backgrounds

Let the target modeling process be represented by a time series y_t, t = 1, 2, ..., T, where y_t, – the number of registered infection cases at time t.

In the classical formulation of time series forecasting, future values are estimated solely based on past observations of the target variable [23]. This approach can be expressed as:

$$\hat{y_{t+h}} = f(y_t, y_{t-1}, \dots, y_{t-W}), \tag{1}$$

where h is the forecasting horizon, W is the length of the historical window.

This formulation implicitly assumes that the statistical properties of the process remain relatively stable over time. However, under crisis conditions and external shocks, such assumptions are frequently violated, resulting in structural breaks and degraded forecasting performance.

From the non-stationary process theory, the dynamics of y_t are influenced not only by internal temporal dependencies but also by external factors [24]. These external influences are represented by exogenous variables. The set of available exogenous variables describing external conditions affecting the target process:

$$Z = (z^{(1)}, z^{(2)}, \dots, z^{(K)}). \tag{2}$$

The observed values of these variables at time t are represented by the vector:

$$z_t = (z_t^{(1)}, z_t^{(2)}, \dots, z_t^{(K)}), \tag{3}$$

where each component represents an external factor affecting disease dynamics.

Incorporating exogenous information extends (1) to:

$$\hat{y}_{t+h} = f\left(y_t, y_{t-1}, \dots, y_{t-p}, z_t, z_{t-1}, \dots, z_{t-q}\right), \tag{4}$$

where q – the historical depth of exogenous inputs.

With a neural network framework, the forecasting function is approximated by a parameterized model f_θ, where θ is the set of learnable parameters. Model training is formulated as a minimization problem:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^{N} l\left(y_{t_i+h}, \hat{y}_{t_i+h}\right), \tag{5}$$

where l is a loss function, and N is the number of training samples.

Forecasting accuracy depends not only on the model parameters θ, but also on the structure and informativeness of the input space.

In this study, optimization is interpreted as a data-centric problem focused on the rational construction of the model input space rather than on increasing architectural complexity. Specifically, the optimization problem consists in selecting an informative subset $Z^* \subset Z$ that minimizes forecasting error:

$$Z^* = L\left(f_\theta(y, Z)\right). \tag{6}$$

The values of exogenous variables are not optimized directly; instead, the optimization is performed over the structure of the input space.

3. Research methods

The proposed approach focuses on the organization and selection of input information rather than on modifying model architecture or training procedures.

The experimental base of the study consists of time series data. Data preprocessing includes the removal of missing or anomalous values, temporal alignment of observations, normalization of numerical values, and transformation of the data into a format suitable for neural network training. All time series are aligned on a unified temporal scale and divided into training and testing subsets. Additional input information is represented by a set of available variables (2), where each variable corresponds to an independent source of information potentially affecting the dynamics of the target time series. Observed values of these variables at time t are represented by the vector (3). The complete input representation available to the forecasting model at time t is defined as.

$$X_t = \left(y_t, y_{t-1}, \dots, y_{t-W+1}, z_t, z_{t-1}, \dots, z_{t-W+1}\right), \tag{7}$$

where y_t the target time series, W – the length of the historical window.

Using all available variables does not necessarily lead to improved forecasting performance, as redundant or weakly informative inputs may increase noise and reduce model stability. Therefore, the structure of the model input space is defined by selecting a subset, which specifies a particular configuration of additional input variables.

$$Z' \subset Z. \tag{8}$$

For a given configuration Z', the corresponding input representation is expressed as

$$X_t^Z = \left(y_t, \dots, y_{t-W+1}, z_t^{Z'}, \dots, z_{t-W+1}^{Z'}\right), \tag{9}$$

where $z_t^{(Z')}$ contains only the variables included in the selected subset Z'.

The optimization strategy is based on the principles of data-centric machine learning and treats the structure of the model input space as the primary object of optimization. Instead of modifying the forecasting model itself, the approach systematically evaluates alternative configurations of input variables.

A set of candidate input space configurations $S=\{Z_1^{'}, Z_2^{'}, \ldots, Z_M^{'}\}$ is constructed. For each configuration Z, a forecasting model with fixed architecture and training parameters is trained and evaluated. The forecasting task is solved repeatedly at this stage; however, it serves exclusively as an internal assessment mechanism for measuring the informativeness of each input space configuration rather than as the final objective of the study. The optimal structure of the model input space is defined as (6).

Algorithm for constructing the model input space:

Step 1. Define the target time series y_t and historical window length W.

Step 2. Specify the set of available additional variables Z.

Step 3. Construct a set of candidate input space configurations S.

Step 4. For each configuration $Z^{'}_{i}$, form the corresponding input representation $X_t^{(Z^{'}_{i})}$.

Step 5. Train a forecasting model using a fixed architecture and experimental setup and evaluate forecasting quality using selected metrics.

Step 6. Compare the results across all configurations and determine the optimized input space structure Z^*.

The workflow of the proposed approach is illustrated in Figure 1, which demonstrates the difference between the conventional forecasting pipeline with a fixed input space and the proposed optimization-oriented procedure for constructing the model input space.

Forecasting quality is assessed using standard quantitative metrics.

The search for the optimal model input space structure Z^* is formulated as a subset selection problem over the set of available additional variables Z. The search strategy is implemented as a two-stage procedure. In the first stage, a preliminary reduction of the original variable set Z is performed, resulting in a reduced subset Z_{top} obtained using fast screening criteria. This step is introduced to limit computational complexity. In the second stage, a wrapper-based search is applied, using forward selection to iteratively construct candidate input space configurations and identify the configuration that yields the best forecasting performance.

The optimal configuration Z^* is selected based on an aggregated forecasting quality measure computed across multiple error metrics and multiple forecasting horizons. When several configurations produce similar error values, preference is given to the configuration that demonstrates higher stability, reflected by lower variability of forecasting errors across different time segments or validation folds.

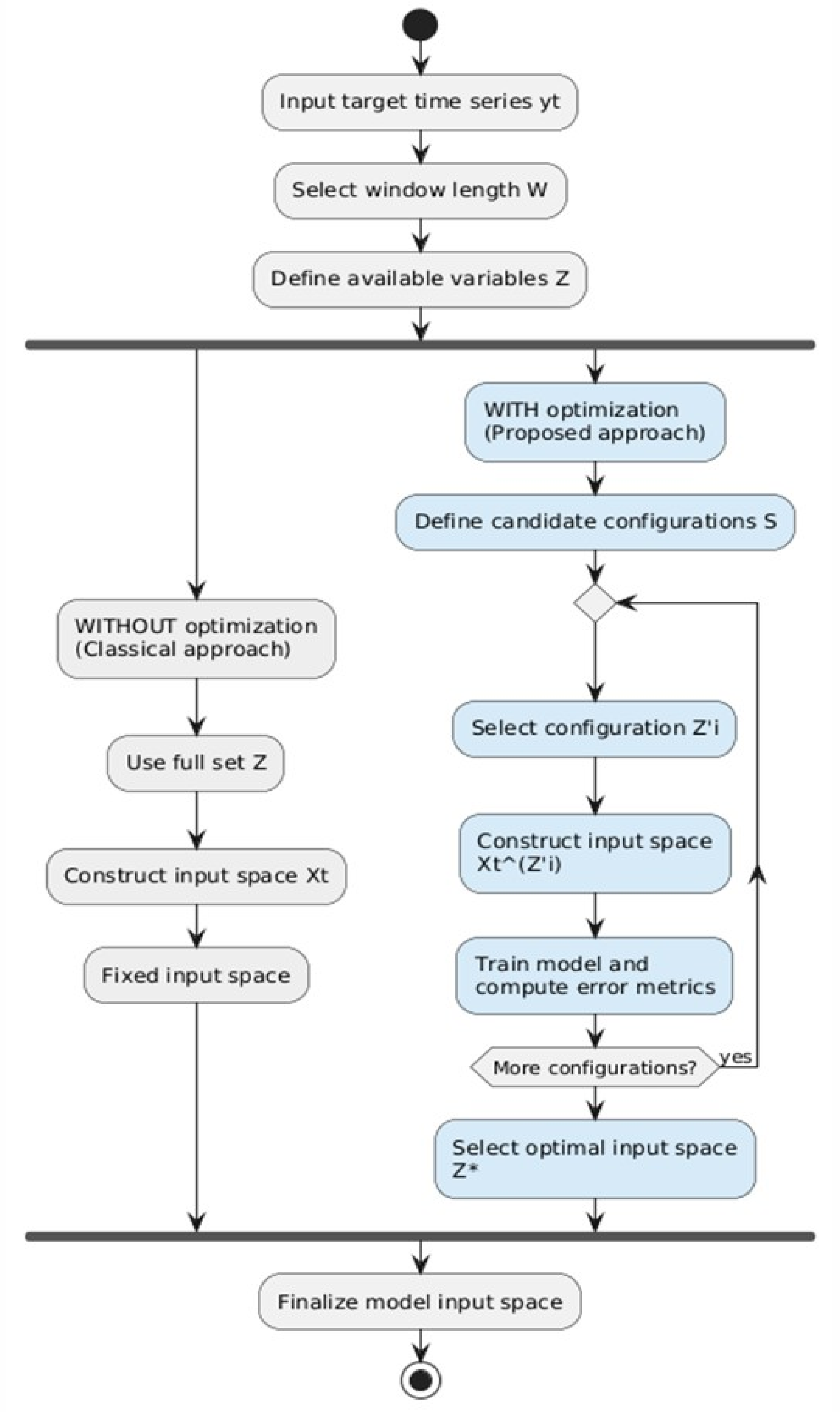

Figure 1: Workflow for constructing the model input space with and without optimization.

4. Results and discussion

In [44], the task of neural network based forecasting of infectious disease spread in Ukraine was addressed using normalized monthly time series constructed from reports of the Public Health Center of the Ministry of Health of Ukraine (reporting form No. 1) for the period from December 2016 to January 2024.

In the present study, a comparison is performed between the baseline forecasting approach and the approach with an optimized model input space, while all key experimental settings remain unchanged. In particular, the same neural network architecture (N-BEATS), temporal resolution (monthly time step), historical window length $W=24$ months, forecasting horizon $h=6$ months, data preprocessing and normalization procedures, training strategy, and evaluation metrics are preserved. This ensures that any differences in forecasting performance can be attributed solely to the optimization of the model input space.

The dataset contains the disease category, the time index, and the incidence indicator per 100,000 population: x_1 – infectious category (disease), total number of categories – 57; values encoded as integers from 1 to 57; x_2 – date, observation period – from December 2016 to January 2024; x_3 – incidence level (cases);

The target time series y_t is formed from the incidence indicator x_3 for each disease.

Prior to model training, the target time series were scaled using min-max normalization to the [0, 1] interval. Exogenous variables defined on bounded scales were used without additional normalization, while variables with arbitrary ranges were normalized separately. This preprocessing ensured numerical stability while preserving the semantic meaning of each variable. The normalization was performed separately for each time series over the entire observation period:

$$y_t^{norm} = \frac{y_t - min(y)}{max(y) - min(y)}. \qquad (10)$$

In the experiment, normalization parameters were computed using the full observation period. While this ensures consistent scaling across the time series, it may introduce a potential risk of information leakage, as future values can influence the scaling of earlier observations. However, since normalization is applied uniformly across all compared configurations, it does not affect the validity of the comparative analysis.

For example, consider the normalization of data for the disease Influenza. The analysis showed that over the entire observation period $min(y) = 0$, $max(y) = 444.6$. In February 2017 the original incidence value was $y_{2017.02} = 55.18$.

$$y_{2017.02}^{norm} = \frac{55.18}{444.6} \approx 0.124. \qquad (11)$$

The same normalization procedure was applied to all time points of the analyzed time series. It should be noted that relatively low normalized values observed in the early years are explained by the presence of higher epidemic peaks in later periods, which determine the global maximum of the time series.

In the experimental part of the study, a limited set of exogenous data is considered to extend the model input space. These variables are not generated by the internal dynamics of the disease incidence time series and are not used in the baseline neural network model without optimization. All exogenous variables are deterministic, derived from calendar and demographic information, and available at the forecasting origin, which reduces the risk of information leakage from future periods, as all variables are available at the forecasting origin. The selection of these variables is motivated by their known relevance in epidemiological forecasting, where seasonal patterns and population-related factors are among the primary drivers of disease dynamics.

The variable z_1 represents the calendar month index (from 1 to 12) and is used to account for intra-annual seasonal patterns in disease incidence. The variable z_2 denotes the season of the year (encoded as 1–4) and reflects broader seasonal effects. The variable z_3 corresponds to the normalized population size and captures demographic influence on disease dynamics.

The specified exogenous variables are not included in the model input space simultaneously by default. Instead, they are treated as elements of a candidate set from which different input space configurations are constructed. These configurations are evaluated sequentially within the optimization procedure, and the optimal set of exogenous variables is selected based on forecasting performance while keeping the neural network architecture, training parameters, historical window length, and forecasting horizon unchanged.

In order to identify the most informative configuration of input data, a systematic enumeration of all possible subsets of the set, in Table 1. In general, the input space optimization may be implemented as a two-stage procedure (screening followed by forward selection). However, in this study the candidate set was limited to three exogenous variables (|Z|=3), therefore the screening step was trivial and the optimization was implemented as an exhaustive enumeration of all subsets, which is equivalent. This design choice ensures a controlled experiment and allows for an exact solution to the input selection problem, avoiding approximation errors associated with heuristic search strategies.

Table 1
Selection of the optimal input space configuration

#	Configuration Z'			Average MAPE, %
	z_1	z_2	z_3	
0	-	-	-	82.92
1	+	-	-	79.10
2	-	+	-	80.35
3	-	-	+	79.80
4	+	+	-	77.95
5	**+**	**-**	**+**	**73.56**
6	-	+	+	78.40
7	+	+	+	78.41

As shown in Table 1, the lowest forecasting error is achieved when using the configuration $Z^* = \{z_1, z_3\}$. Including all exogenous variables without prior selection does not yield the best result, indicating the presence of redundant features in the input space and confirming the necessity of the optimization procedure.

In Table 2, the in-sample training performance of the N-BEATS model is reported for the selected input space configurations.

Table 2
Performance of neural networks trained on pooled time series

Model configuration	Input space	Average MAPE, %	Average MAE	Average RMSE
N-BEATS without optimization [13]	$y_{t:t-W+1}$	63.56	1140.8	1322.6
N-BEATS with all exogenous variables	$y_{t:t-W+1} + \{z_1, z_2, z_3\}$	60.20	1095.4	1268.1
N-BEATS with optimized input space	$y_{t:t-W+1} + \{z_1, z_3\}$	**57.10**	**1038.6**	**1201.9**

The results presented in Table 2 indicate that the inclusion of exogenous variables reduces approximation error on the training data. However, the greatest improvement is observed only when using the optimized input space configuration, which suggests that not all exogenous factors contribute equally to modeling the time series dynamics. These findings demonstrate the effectiveness of a data-centric approach to improving neural network forecasting performance without increasing model architectural complexity.

Figure 2 shows a neural network forecast for one infectious disease – influenza. As can be seen, the forecast generated using the optimized input space more accurately matches the observed dynamics than the baseline model, which is consistent with the quantitative improvements in forecasting accuracy noted in the experimental results.

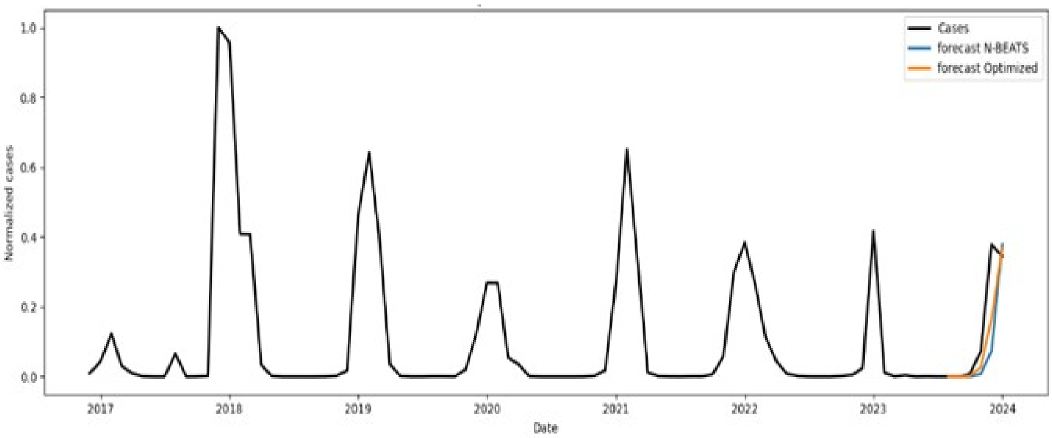

Figure 2: Forecasting with optimized input space (Influenza).

5. Conclusions and prospects for further research

This study investigated the improvement of neural network–based time series forecasting under non-stationary conditions through optimization of the model input space rather than through modifications of network architecture. Using infectious disease incidence data, the baseline N-BEATS model demonstrated limited predictive performance, with an average forecasting error of 82.92% MAPE.

The proposed input space optimization approach, based on systematic selection of exogenous variables, led to a substantial improvement in forecasting accuracy. The inclusion of all available exogenous variables reduced the average out-of-sample error to 78.41% MAPE, whereas the optimized input space achieved a further reduction to 73.56% MAPE. This corresponds to an overall relative improvement of approximately 11.3% compared to the baseline model.

In-sample evaluation also confirmed the positive effect of input space optimization. The average training error decreased from 63.56% MAPE for the baseline configuration to 57.10% MAPE for the optimized model, indicating improved approximation quality without signs of excessive overfitting. Importantly, these gains were achieved without changing the N-BEATS architecture, training procedure, or forecasting horizon.

The obtained results demonstrate that data-centric optimization of input variables is an effective and computationally efficient strategy for enhancing forecasting accuracy in unstable time series. This approach improves model generalization and practical relevance while avoiding unnecessary increases in model complexity.

Further research will focus on exploring automated and adaptive strategies for exogenous variable selection, and integrating dynamic input space optimization into real-time forecasting systems.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] J. Kim, et al. A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges, Artif. Intell. Rev. 58 (2025) 216. doi:10.1007/s10462-025-11223-9.

[2] I. Atamanyuk, Y. Kondratenko, Calculation Method for a Computer's Diagnostics of Cardiovascular Diseases Based on Canonical Decompositions of Random Sequences, in: CEUR Workshop Proceedings, Vol. 1356, 2015, pp. 108-120. URL: http://ceur-ws.org/Vol-1356/paper_28.pdf.

[3] O. Kozlov, Y. Kondratenko, O. Skakodub, Comparative Analysis of Parametric Optimization Techniques: Fuzzy DSS for Medical Diagnostics, in: S. N. Shahbazova, et al. (Eds.), Recent Developments and the New Directions of Research, Foundations, and Applications. Studies in Fuzziness and Soft Computing, vol. 423, Springer, Cham, 2023, pp. 193-201. doi:10.1007/978-3-031-23476-7_17.

[4] I. Sidenko, et al., Artificial Intelligence Technologies for Efficient Solving of Recognition Tasks, in book: Artificial Intelligence: Achievements and Recent Developments, 2025, pp. 145–196. doi:10.1201/9788743800989-6.

[5] O. Kovaliv, et. al., Neural Network Architectures for Recognizing Military Objects on Satellite Images, in: Proceedings of the IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 07-09 September 2023. doi:10.1109/IDAACS58523.2023.10348905.

[6] D. Chumachenko, et. al., Methodology for assessing the impact of emergencies on the spread of infectious diseases, Radioelectronic and computer systems 3(111) (2024) 6–26. doi:10.32620/reks.2024.3.01.

[7] R. D. T. de Lima, S. M. M. Fernandes, e S. M. L. de Lima, Time series forecasting with exogenous variables: a literature review to identify promising gaps in computational research, Rev. Principia 62 (2025). doi:10.18265/2447-9187a2025id8575.

[8] E. Etu, et al. Forecasting pediatric emergency department arrivals: Evaluating the role of exogenous variables using deep learning models, Intelligence-Based Medicine 12 (2025) 100313. doi:10.1016/j.ibmed.2025.100313.

[9] A. K. Shaikh, A. Nazir, I. Khan, A. S. Shah, Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series, Scientific reports 12(1) (2022) 22562.

[10] C. Challu, et al., NHITS: Neural Hierarchical Interpolation for Time Series Forecasting, in: Proceedings of the AAAI Conference on Artificial Intelligence, 37(6) (2023) 6989-6997. doi:10.1609/aaai.v37i6.25854.

[11] H. Hewamalage, C. Bergmeir, K. Bandara, Recurrent neural networks for time series forecasting: Current status and future directions, International Journal of Forecasting 37(1) (2021) 388-427. doi:10.1016/j.ijforecast.2020.06.008.

[12] R. Magallanes-Quintanar, et al., Neural Hierarchical Interpolation for Standardized Precipitation Index Forecasting. Atmosphere 15(8) (2024).

[13] S. F. Stefenon, et al., Neural Hierarchical Interpolation Time Series (NHITS) for Reservoir Level Multi-Horizon Forecasting in Hydroelectric Power Plants, IEEE Access 13 (2025) 54853-54865. doi: 10.1109/ACCESS.2025.3554446.

[14] K. Bandara, C. Bergmeir, H. Hewamalage, LSTM-MSNet: Leveraging Forecasts on Sets of Related Time Series With Multiple Seasonal Patterns, IEEE Transactions on Neural Networks and Learning Systems 32(4) (2021) 1586-1599. doi: 10.1109/TNNLS.2020.2985720.

[15] O. Bulgakova, V. Zosimov, V. Perederyi, AI Scheduling with Contextual Transformers, International Journal of Intelligent Systems and Applications 17(4) (2025). doi:10.5815/ijisa.2025.04.05.

[16] S. Ahmed, et al., Transformers in Time-Series Analysis: A Tutorial, Circuits Syst Signal Process 42 (2023) 7433–7466. doi:10.1007/s00034-023-02454-8.

[17] H. Wu, J. Xu, J. Wang, M. Long, Autoformer: decomposition transformers with auto-correlation for long-term series forecasting, in: Proceedings of the 35th International Conference on Neural Information Processing Systems (NIPS '21). Curran Associates Inc., Red Hook, NY, USA, 1717, 22419–22430, 2021.

[18] R. P. Masini, M. C. Medeiros, E. F. Mendes, Machine learning advances for time series forecasting, J. Econ Surv 37 (2023) 76–111. doi:10.1111/joes.12429.

[19] Y. Wang, et al., TimeXer: empowering transformers for time series forecasting with exogenous variables, in: Proceedings of the 38th International Conference on Neural Information Processing Systems (NIPS '24). Curran Associates Inc., Red Hook, NY, USA, 15, 469-498, 2024.

[20] T. Zhou, et al., Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting, in: International conference on machine learning, 27268-27286, 2022.

[21] Y. Nie, A Time Series is Worth 64Words: Long-term Forecasting with Transformers, 2022. doi:10.48550/arXiv.2211.14730.

[22] A. Zeng, M. Chen, L. Zhang, Q. Xu, Are Transformers Effective for Time Series Forecasting?, in: Proceedings of the AAAI Conference on Artificial Intelligence, 37(9), 11121-11128, 2023. doi:10.1609/aaai.v37i9.26317.

[23] Y. Liu, et al., Transformer: Inverted transformers are effective for time series forecasting, arXiv:2310.06625.

[24] I. V. Sidenko, et. al., Hybrid neural network and genetic algorithm combination for recognition and avoidance of vehicle obstacles, in: Proceedings of the Industrial Simulation Conference'2024, EUROSIS-ETI, Valencia, Spain, 2024.

[25] I. Sidenko, G. Kondratenko, O. Heras, Y. Kondratenko, Neural Technologies for Objects Classification with Mobile Applications, Journal of Mobile Multimedia 20(03) (2024) 727-748. doi:10.13052/jmm1550-4646.2039.

[26] V. Stepashko, O. Bulgakova, V. Zosimov, Numerical study of the generalized iterative algorithm GIA GMDH with active neurons, in: 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 2017, pp. 496-500. doi: 10.1109/STC-CSIT.2017.8098836.

[27] B. Adhikari, X. Xu, N. Ramakrishnan, B. A. Prakash, EpiDeep: Exploiting Embeddings for Epidemic Forecasting, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). Association for Computing Machinery, New York, NY, USA, 577–586. doi:10.1145/3292500.3330917.

[28] State-Space Models, in: P. J. Brockwell, R.A. Davis, (Eds.), Introduction to Time Series and Forecasting, Springer Texts in Statistics, Springer, New York, NY. doi:10.1007/0-387-21657-X_8.

[29] O. Ryan, J. M. B. Haslbeck, L. J. Waldorp, Non-Stationarity in Time-Series Analysis: Modeling Stochastic and Deterministic Trends. Multivariate Behavioral Research 60(3) (2025) 556-588. doi:10.1080/00273171.2024.2436413.

[30] V. Zosimov, O. Bulgakova, Application of Personalized Ranking Models Based on Expert Evaluations for Sorting Goods on E-commerce Web Resources, in: IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT), Zbarazh, Ukraine, 2020, pp. 42-45. doi:10.1109/CSIT49958.2020.9321902.

[31] B. Lim, S. Ö. Arık, N. Loeff, T. Pfister, Temporal fusion transformers for interpretable multi-horizon time series forecasting, International journal of forecasting 37(4) (2021) 1748-1764.

[32] S. Ö. Arik, T. Pfister, TabNet: Attentive Interpretable Tabular Learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, 35(8), 6679-6687, 2021. doi:10.1609/aaai.v35i8.16826.

[33] K. G. Olivares, et. al., Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with NBEATSx, International Journal of Forecasting 39(2) (2023) 884-900. doi:10.1016/j.ijforecast.2022.03.001.

[34] J. Gasthaus, et. al., Probabilistic forecasting with spline quantile function RNNs, in: The 22nd international conference on artificial intelligence and statistics, pp. 1901-1910, 2019.

[35] Q. Huang, Z. Zhou, K. Yang, Y. Wang, Exploiting Language Power for Time Series Forecasting with Exogenous Variables, in: Proceedings of the ACM on Web Conference, pp. 4043-4052, 2025.

[36] H. Zhou, et. al., Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting, in: Proceedings of the AAAI Conference on Artificial Intelligence, 35(12), 11106-11115, 2021. doi:10.1609/aaai.v35i12.17325.

[37] S. Shastri, et. al., Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study, Chaos, Solitons & Fractals 140 (2020) 110227. doi:10.1016/j.chaos.2020.110227.

[38] V. K. R. Chimmula, L. Zhang, Time series forecasting of COVID-19 transmission in Canada using LSTM networks, Chaos, solitons & fractals 135 (2020) 109864. doi:10.1016/j.chaos.2020.109864.

[39] A. S. Kwekha-Rashid, H. N. Abduljabbar, B. Alhayani, Coronavirus disease (COVID-19) cases analysis using machine-learning applications, Appl Nanosci 13 (2023) 2013-2025. doi:10.1007/s13204-021-01868-7.

[40] Y. Cheng, et. al., Machine Learning Techniques Applied to COVID-19 Prediction: A Systematic Literature Review, Bioengineering 12(5) (2025) 514. doi:10.3390/bioengineering12050514.

[41] A. Adugna, D. Abebaw, A. Abebaw. M. Jemal, Deep learning architectures for influenza dynamics and treatment optimization: a comprehensive review, Front. Artif. Intell. 8 (2025) 1521886. doi: 10.3389/frai.2025.1521886.

[42] Y. Wang, Y., Wang, Y., Hu, J. Wang, Biology-Informed Matrix Factorization: An AI-Driven Framework for Enhanced Drug Repositioning, Biology 14(5) (2025) 549. doi:10.3390/biology14050549.

[43] O. Bulgakova, V. Zosimov, Modeling the Impact of External Factors on E-Commerce Consumer Behavior, in: IEEE 19th International Conference on Computer Science and Information Technologies (CSIT), 2024, pp. 1-4, doi: 10.1109/CSIT65290.2024.10982577.

[44] O. Kovaliv, Y. Kondratenko, I. Sidenko, G. Kondratenko, D. Chumachenko, Improving the Accuracy of Infectious Disease Forecasts Based on Comparing Neural Network Architectures, Computation 14 (2026) 54. doi:10.3390/computation14020054.