ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2026, Vol. 87, No. 6, pp. 529–541.

Forecasting Epidemic Outbreaks of Acute Respiratory Viral Infections with Early Warning Signals and Machine Learning

M. Yu. Lazutov[*,a] **and V. N. Leonenko**[*,**,b]

[*]*ITMO University, St. Petersburg, Russia*
[**]*Smorodintsev Research Institute of Influenza, St. Petersburg, Russia*
e-mail: [a]*marklazutov@gmail.com,* [b]*vnleonenko@yandex.ru*

Received April 20, 2025
Revised October 120, 2025
Accepted November 6, 2025

Abstract—Forecasting critical transitions in epidemic incidence is important for epidemic surveillance. Earlier detection can support a more timely public-health response. This paper studies anomaly detection methods based on artificial intelligence and early warning signals (EWSs) in incidence time series. The goal is to predict transitions from seasonal infections to epidemic outbreaks. Two approaches are examined. The first is a classification approach that estimates whether a critical transition is near. The second is a regression approach that forecasts future infection dynamics. Several machine learning models are applied to two types of data. The models include ensemble methods (Easy Ensemble, RUSBoost, and Balanced Bagging) and deep learning architectures (Early Warning Signal Network (EWSNet), LSTM, and GRU). The first dataset contains influenza incidence with epidemic periods labeled by expert criteria. The second dataset contains unlabeled COVID-19 incidence. The results show that Easy Ensemble and EWSNet provide the best balance between precision and recall. Recurrent neural networks model the dynamics of mean values effectively, whereas variance forecasting remains more difficult. The results show that classical EWS methods and machine learning can be combined to improve epidemic forecasting and support public-health decision making.

Keywords: epidemic forecasting, early warning signals, machine learning, time-series analysis, critical transitions

DOI: 10.7868/S1608303226060056

1. INTRODUCTION

Early detection and forecasting of critical phenomena are important in many scientific fields [1], including epidemiology [2]. This task can also be viewed as a signal-recognition problem [3]. A relevant example is the emergence of a new respiratory viral strain. At early stages, such strains may be difficult to distinguish from seasonal acute respiratory viral infections in aggregated surveillance data. For this reason, specialized statistical indicators are used together with more general methods. These indicators are known as early warning signals (EWSs). They help estimate whether a system is approaching a transition point, or critical point. Many EWS variants have been proposed. They include spectral EWSs based on the spectrum of a time series and EWSs based on critical slowing down [4]. The latter group includes indicators based on increases in variance, autocorrelation, skewness, and related statistics. For epidemic processes, many studies have tested EWSs on both synthetic and real data [5].

This study compares methods for early detection of incidence outbreaks caused by pandemic respiratory viruses. The analysis uses incidence data for acute respiratory viral infections, including seasonal influenza and COVID-19. We compare specialized EWSs based on critical slowing down with more general statistical approaches. We also evaluate whether these methods can support epidemic-process management and control. Two forecasting settings are considered. In the first setting, an epidemic threshold is available. Then the task is binary classification: the critical transition point is either "near" or not near. The proximity interval is selected according to epidemic-surveillance requirements. In this study, it is one month before the critical point. In the second setting, epidemic-threshold labels are unavailable. This is the case for COVID-19, because the disease has not circulated long enough to estimate a multi-year baseline. Then outbreak detection is formulated as forecasting selected EWSs. The absolute indicator values are predicted, and the time to the critical point is estimated from the forecast incidence. This setting is a special case of predictive statistical methods based on multi-year incidence data (see, for example, [6]). Such methods, together with SIR models, are among the most widely used forecasting models in epidemiology.

2. BACKGROUND AND RELATED APPROACHES

Machine learning is now widely used in epidemic forecasting. This development follows broader progress in machine learning and artificial intelligence. Traditional models, such as SIR and SEIR models, are useful but have limitations. They may be too rigid for nonlinear factors and for the early stages of pandemics caused by novel viruses [2]. Classical machine learning algorithms can also be effective. For example, Random Forest and support vector machines have been used for early epidemic outbreak detection [7]. Recurrent neural networks have also shown good results in disease forecasting; see, for example, [8]. At the same time, many studies continue to use the theory of complex systems and EWSs. In these studies, EWSs may be statistical indicators or explanatory models, such as a risk index based on social-distancing measures [28]. Table 1 lists studies that use different EWS-based approaches to epidemic forecasting. Some studies use a hybrid approach that combines machine learning with EWSs. However, comparative studies of such combinations are still limited. To address this gap, this paper considers several models and two problem formulations: classification and regression.

Table 1. Related studies

	2025 Zhao et al. [9]	2024 Bizzotto et al. [10]	2025 Gao et al. [11]	2024 Drake et al. [12]	2021 O'Brien [13]
Context	31 diseases, 134 regions	COVID-19, Italy	COVID-19, Singapore, Canada, Hong Kong	SARS-CoV-2, UK	COVID-19, 24 countries
Method	Critical-slowing-down resilience indicators + time-to-event analysis	Nowcasting reproduction number method (R_e)	Feature-based time-series classification (ML)	Phylogenetic logistic growth rate (LGR) analysis	Sequential EWS assessment using GAMs
Best-performing EWSs	Variance, ACF	Nowcasted R_e	Variance, ACF, skewness	Maximum LGR of dominant clusters	Concatenation of variance, ACF, and skewness
Forecast horizon	18–21 days	6–23 days	14 days	6–7 days	10–14 days

3. DATA

Two input datasets were used in this study. The first dataset was obtained from the database of the Research Institute of Influenza. It contains weekly incidence of acute respiratory viral infections in St. Petersburg from January 1990 through October 2020. The dataset includes the absolute number of cases. It also includes the estimated city population for each year of the study period. This information makes it possible to convert incidence to relative measures. The dataset also indicates whether the seasonal epidemic threshold was exceeded. The threshold was calculated according to the methodology of the Research Institute of Influenza. This indicator separates periods with an officially declared epidemic situation from non-epidemic periods. The general form of the time series is shown in Fig. 1. Since the data contain epidemic and non-epidemic labels, they are suitable for an EWS-based classification problem.

The second time series was obtained from open data on COVID-19 incidence in St. Petersburg from March 2020 to March 2024 [14]. The general form of the time series is shown in Fig. 2. These incidence data were collected daily. For COVID-19, it is more difficult to define an "epidemic outbreak" than for seasonal and pandemic influenza. The reason is that coronavirus infection does not have clearly expressed seasonality. In this study, outbreaks are defined as sustained increases in incidence followed by declines. These increases form epidemic waves in the data and are associated with the emergence of new coronavirus strains.

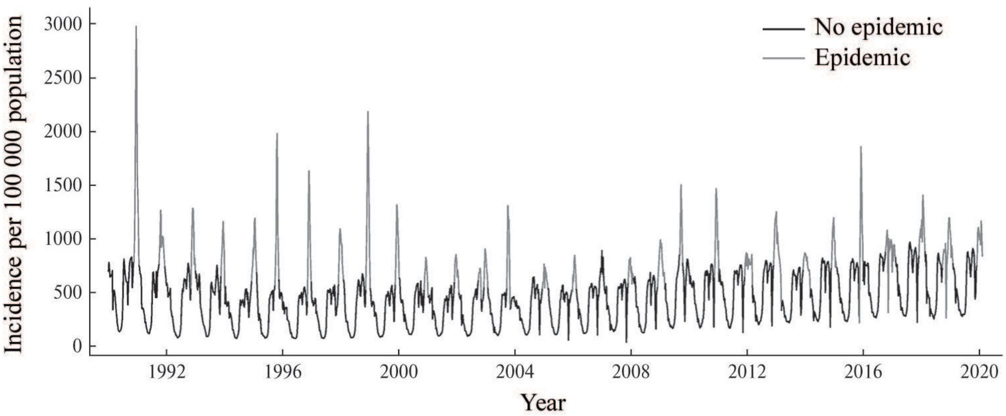

Fig. 1. Time series with labeled epidemics.

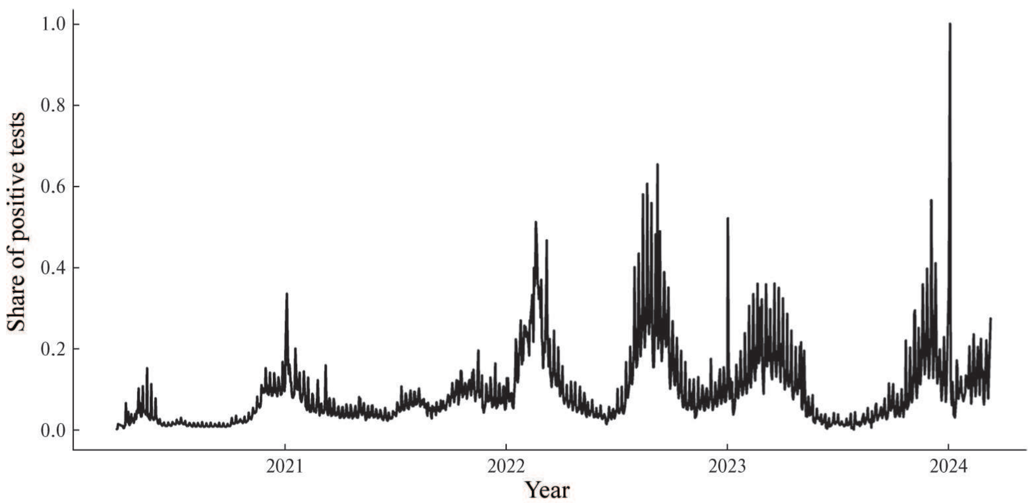

Fig. 2. Time series without labeled epidemics.

4. THEORETICAL BACKGROUND

We selected EWSs that have shown good performance for epidemic time series. According to Table 1 and [15], most studies use a forecasting horizon of about one to two weeks. Variance and the mean value are among the most useful indicators of proximity to a critical point. These indicators often increase when epidemic onset approaches. Figure 3 shows EWS calculations for acute respiratory viral infection incidence. Two cases are shown: one near influenza epidemic onset and one far from it. In this example, the most informative indicator is variance with a window length of 30. In the first case, this indicator increases monotonically. This is a sign of critical slowing down and may indicate an approaching phase transition [4]. In the second case, the indicator decreases monotonically. This means that the effect is absent, which is the expected behavior of an EWS far from a transition. Variance with window lengths of 15 and 20 does not show the same pattern. Thus, variance is sensitive to the window length. The window parameter should therefore be selected empirically before the indicator is used in epidemic-surveillance practice. In the classification task, we compare variance and skewness as EWSs. In the regression task, we forecast the variance and the mean incidence value and use them to predict epidemic increases. We selected longer decision horizons than those used in the reviewed studies: 4 and 5 weeks for classification and 14 and 28 days for regression.

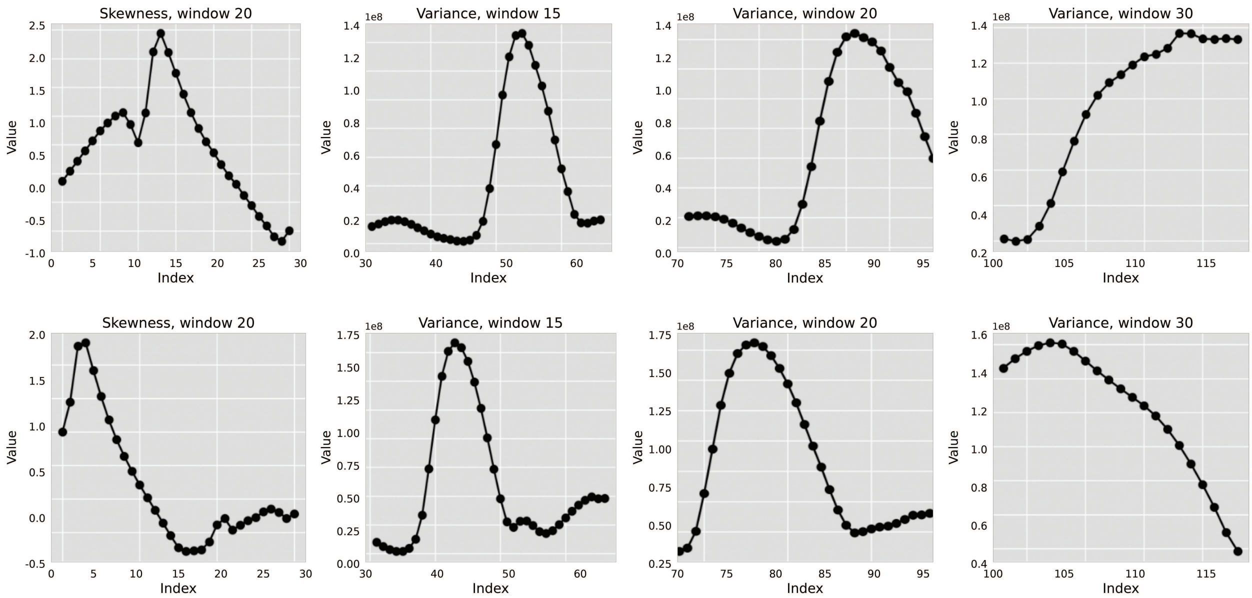

Fig. 3. EWSs near (upper row) and far from (lower row) the critical transition.

The next section describes the methods used to solve the classification and regression problems. The EWSs were calculated using the Python package `ewstools` [16].

4.1. Classification Approach

The data were preprocessed before model training and testing. First, the time series was divided into segments without epidemics. Second, a time interval t before epidemic onset was specified. This interval defines the warning lead time. After that, the training sample was constructed.

Data preprocessing was performed as follows. The time series with the number of incidence cases is represented as a vector $X = \{x_k\}$, $k = \overline{1, |X|}$ and is supplemented with a time series of binary labels Y:

$$Y = \begin{cases} 1, & \text{epidemic,} \\ 0, & \text{no epidemic.} \end{cases} \tag{1}$$

All non-epidemic segments are then selected:

$$X' = \{X_i \mid \forall j \in [0,\ |X_i| - 1]\ Y_j = 0\}, \quad i = \overline{0, M-1},$$

where X_i is a subset of X and M is the number of non-epidemic segments.

The obtained segments are then transformed using the sliding-window rule:

$$X'' = \begin{pmatrix} X_0^{(0)} & X_0^{(1)} & \dots & X_0^{(w-1)} \\ X_0^{(1)} & X_0^{(2)} & & X_0^{(w)} \\ \vdots & & \ddots & \vdots \\ X_{M-1}^{N-w} & X_{M-1}^{N-w+1} & \dots & X_{M-1}^{N-1} \end{pmatrix}, \tag{2}$$

where $N = |X_{M-1}|$ and w is the window length. In this study, $w = 30$ was used.

The choice of the threshold value t is important. This value determines when the model should report an approaching epidemic. According to epidemic-surveillance specialists, a warning should be received no later than one month before epidemic onset. This corresponds to $t = 4$ or $t = 5$. The incidence time series for seasonal acute respiratory viral infections is recorded weekly. Therefore, $t = 4$ and $t = 5$ correspond to 28 and 35 days. The following binary array describes whether a given time-series point is close to epidemic onset:

$$Y'_l = \begin{cases} 1, & \text{if } E_l - I_l \le t, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Here $l = \overline{0, K}$, and K is the number of rows in matrix X''. The value I_l is the index of the last element of the lth time series in X''. The value E_l is the first index of the nearest epidemic incidence value, where $E_l \in [0, |X|]$.

For each row in the obtained matrix, EWSs were calculated with specified values of the window w_{ews}. In this work, variance was calculated for $w_{ews} = \{15,\ 20, 30\}$. Skewness was calculated for $w_{ews} = 20$. The calculated EWS values were then combined into the final matrix:

$$X''' = \begin{pmatrix} S_0^{(0)} & S_0^{(1)} & \dots & V_0^{\prime,(0)} & V_0^{\prime,(1)} & \dots & {V''}_0^{(0)} & \dots & {V'''}_0^{(29)} \\ \vdots & & \ddots & & & & \vdots & & \\ S_{M-1}^{(0)} & S_{M-1}^{(1)} & \dots & V_{M-1}^{\prime,(0)} & V_{M-1}^{\prime,(1)} & \dots & {V''}_{M-1}^{(0)} & \dots & {V'''}_{M-1}^{(29)} \end{pmatrix}. \tag{4}$$

Here $S_i^{(j)}$ is the skewness value with index j for time series i. The value $V_i^{\prime,(j)}$ is the variance value with $w_{ews} = 15$ and index j for time series i. The values ${V''}_i^{(j)}$ and ${V'''}_i^{(j)}$ are variance values with $w_{ews} = 20$ and $w_{ews} = 30$, respectively.

Thus, the final dataset is obtained as

$$D_c = \left\{X_i''',\ Y_i'\right\}, \quad i = \overline{1, K}. \tag{5}$$

The resulting dataset was divided into training, validation, and test samples. The split was 70, 10, and 20%, respectively. The sample elements were not mixed because the data are ordered in time. The dataset is imbalanced: class 0 samples outnumber class 1 samples for all values of t. In this task, false negatives are especially important. A false negative means that the model incorrectly reports no epidemic threat. Therefore, recall plays the primary role:

$$Recall = \frac{TP}{TP + FN}. \tag{6}$$

Here TP is the proportion of true positives, and FN is the proportion of false negatives. Accuracy should also remain adequate, for example above 0.5. This condition helps ensure that the models do not classify randomly:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}. \tag{7}$$

Here FP is the proportion of false positives, and TN is the proportion of true negatives. The balance between FN and FP is also important. Therefore, the $F1$ score is used:

$$F1 = 2\frac{Precision \cdot Recall}{Precision + Recall}. \tag{8}$$

The area under the precision–recall curve (PR-AUC) is also useful for model evaluation. It does not depend on the number of true negatives. It also summarizes model performance over a range of threshold values. Therefore, it can be more informative than the $F1$ score.

The following models were used for classification: EWSNet, Easy Ensemble [19], RUSBoost [20], Balanced Bagging classifier [21], Decision Tree, and Random Forest [22]. Deep neural networks with convolutional and recurrent layers can be effective when enough training data are available. Such networks can classify upcoming critical transition points [17, 18]. For this reason, we tested the specialized EWSNet method [18] together with more general approaches. EWSNet combines two independently operating networks: a long short-term memory (LSTM) network and a fully convolutional neural network (FCNN). The authors trained EWSNet on a large synthetic dataset with several dynamical systems. The model shows good accuracy on empirical test data. However, it can still produce frequent false alarms. In the cited study, its test accuracy was about 75%. Its main advantage over classical machine learning models is invariance to the length of the input time series. The comparison results are presented in the next section.

4.2. Regression Approach

For unlabeled COVID-19 data, the task was to forecast an epidemic increase in incidence. The time series of the share of positive COVID-19 tests was represented as an array $X = \{x_k\}$, $k = \overline{1, |X|}$. The time series was split into training, validation, and test sets in proportions of 70, 10, and 20%.

Let $T = \lceil 0.7 \cdot |X| \rceil$ and $V = \lceil 0.1 \cdot |X| \rceil$. Then

$$X_{train} = \{x_k\}, \quad k = \overline{1, T}, \tag{9}$$

$$X_{val} = \{x_k\}, \quad k = \overline{T+1, T+V}, \tag{10}$$

$$X_{test} = X \setminus (X_{train} \cup X_{val}). \tag{11}$$

A set of EWSs can then be calculated. These EWSs form additional time series obtained with a sliding window $w_{ews} = 14$. Each point x_k in X is matched to a point e_X^k in the EWS time series E_X. This EWS point is calculated on the interval $[x_{k-w_{ews}},\ x_{k-1}]$. The modeling task is to calculate $\{e_X^{T+V+1}, e_X^{T+V+2}, \dots\}$. In this task, we use only the time series of mean values $E_{X_{mean}}$ and variances $E_{X_{var}}$.

For model training, the resulting time series must be transformed into matrices using a sliding window m. This parameter can be determined by evaluating metrics on the validation set. The metrics are described below.

Six metrics are used to evaluate modeling quality: MSE, $RMSE$, MAE, $NRMSE$, $NMAE$, and R^2:

$$MSE = \frac{1}{|X_{test}|} \sum_{k=1}^{|X_{test}|} \left(e_X^{T+V+k} - x_{T+V+k}\right)^2, \tag{12}$$

$$RMSE = \sqrt{MSE}, \tag{13}$$

$$MAE = \frac{1}{|X_{test}|} \sum_{k=1}^{|X_{test}|} |e_X^{T+V+k} - x_{T+V+k}|, \tag{14}$$

$$NRMSE = \frac{RMSE}{\max(X_{test}) \ - \min(X_{test})}, \tag{15}$$

$$NMAE = \frac{MAE}{\max(X_{test}) \ - \min(X_{test})}, \tag{16}$$

$$R^2 = 1 - \frac{\sum_k \left(x_{T+V+k} - e_X^{T+V+k}\right)^2}{\sum_k \left(x_{T+V+k} - \mathrm{mean}(X_{test})\right)^2}. \tag{17}$$

The MSE was selected as the loss function in the models. The MAE was also calculated because it is easier to interpret from the subject-area perspective. It directly shows the magnitude of forecast errors. The $RMSE$ is stricter because it is more sensitive to outliers than MAE. This study analyzes time series with different scales, namely the mean and variance. Therefore, normalized metrics, such as $NRMSE$ and $NMAE$, were used. These metrics are invariant to scale changes. The coefficient of determination R^2 was used to estimate the proportion of variance in the reference values explained by the forecast.

Many studies use reservoir computing for short-term forecasting [23–25]. This forecast is short-term because the models are trained only on segments without critical transitions. If transition segments are included, the dynamics may be modeled incorrectly. For example, in [30], recurrent neural networks were used to forecast critical transitions in ecological systems. The networks were trained on raw time series without EWSs, and not all models produced correct estimates. In our case, information is available about some critical transitions that have already occurred. This information can be used for longer-term forecasting. Therefore, we used an LSTM network and a GRU network. GRU is a simplified version of LSTM with one fewer component in its cell. LSTM networks are considered robust and accurate for forecasting chaotic dynamical systems [31]. In the experiments, both LSTM and GRU networks had 128 cells. The selected sliding-window width m was 20. This corresponds to 20 days, or 20 input neurons. In all cases, the models were trained for 100 epochs with the ADAM optimizer.

5. EXPERIMENTAL RESULTS

5.1. Classification Results

Tables 2 and 3 present the test-set metrics for threshold values $t = 4$ and $t = 5$, respectively. The metrics were obtained with the optimal hyperparameters. Confusion matrices were calculated in the same way (Fig. 4).

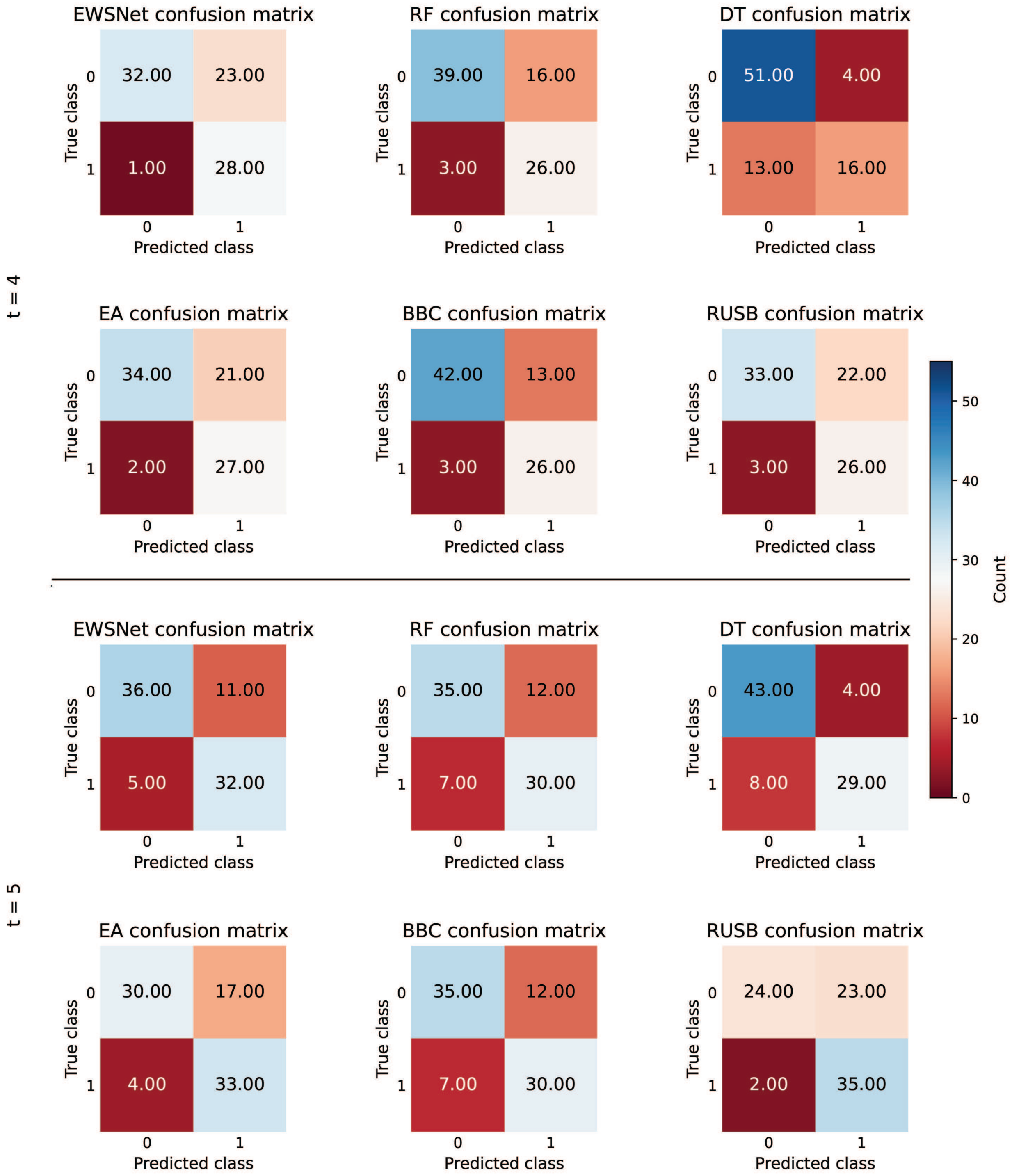

Fig. 4. Confusion matrices at $t = 4$ (two upper rows) and $t = 5$ (two lower rows).

The experiments support the following conclusions:

- For Decision Trees, accuracy varies from moderate to high values, approximately 0.75–0.85. However, recall and the $F1$ score vary strongly depending on the threshold value. This method shows unsatisfactory recall at the lower time-threshold value ($t = 4$). Its lower PR-AUC compared with other ensemble methods indicates weaker performance in ranking positive samples.

Table 2. Metric values for the four-week horizon ($t = 4$)

Classifier	Accuracy	Recall	F1 score	PR-AUC
Decision Tree	0.79	0.55	0.65	0.75
Easy Ensemble	0.72	0.93	0.70	**0.87**
RUSBoost	0.70	0.89	0.67	0.78
Balanced Bagging	**0.80**	0.89	**0.76**	0.84
EWSNet	0.71	**0.96**	0.70	0.86
Random Forest	0.77	0.89	0.73	0.85

Table 3. Metric values for the five-week horizon ($t = 5$)

Classifier	Accuracy	Recall	F1 score	PR-AUC
Decision Tree	**0.85**	0.78	**0.82**	0.89
Easy Ensemble	0.78	0.89	0.78	**0.92**
RUSBoost	0.75	**0.90**	0.73	0.90
Balanced Bagging	0.77	0.81	0.75	0.86
EWSNet	0.80	0.86	0.80	0.88
Random Forest	0.77	0.81	0.75	0.90

- Easy Ensemble shows consistently good results for different threshold values in terms of recall, PR-AUC, and the $F1$ score. Adaptive boosting combined with sampling effectively handles class imbalance. High recall indicates clear identification of minority-class samples. However, this sometimes reduces precision, so the $F1$ score is not always the highest.
- RUSBoost provides balanced results with good recall and PR-AUC values. The cost of the method's simplicity and robustness under class imbalance is a slightly lower accuracy at lower threshold values ($t = 4$) compared with Easy Ensemble.
- Balanced Bagging can be considered a reliable classification tool with moderate recall, PR-AUC, and $F1$ score values. The use of oversampling methods such as SMOTE at $t = 4$ improves the detection of minority-class samples.
- The specialized EWSNet method demonstrates high recall, but is slightly inferior in accuracy and the $F1$ score. It works reliably at $t = 4$, reaching the maximum recall value in the experiment, 0.965, but its lower precision limits its competitiveness in terms of the $F1$ score.
- Random Forest shows stable results. It is structurally simpler and easier to use than several competing methods. It also requires a limited number of features and trees. However, its PR-AUC usually lags behind ensemble methods such as Easy Ensemble and RUSBoost.

5.2. Regression Results

Tables 4 and 5 present the test-set metrics for COVID-19 incidence, which is the second time series. The results are shown for two forecasting horizons: 14 and 28 days. The tables correspond to the GRU and LSTM networks, respectively. The forecasts are shown in Fig. 5.

Table 4. Results for GRU

	14 days		28 days	
	Mean	Variance	Mean	Variance
$NMAE$	0.038	0.058	0.074	0.077
$NRMSE$	0.057	0.125	0.113	0.139
R^2	0.960	0.697	0.838	0.494

Table 5. Results for LSTM

	14 days		28 days	
	Mean	Variance	Mean	Variance
$NMAE$	0.032	0.063	0.054	0.077
$NRMSE$	0.055	0.131	0.094	0.123
R^2	0.961	0.652	0.899	0.578

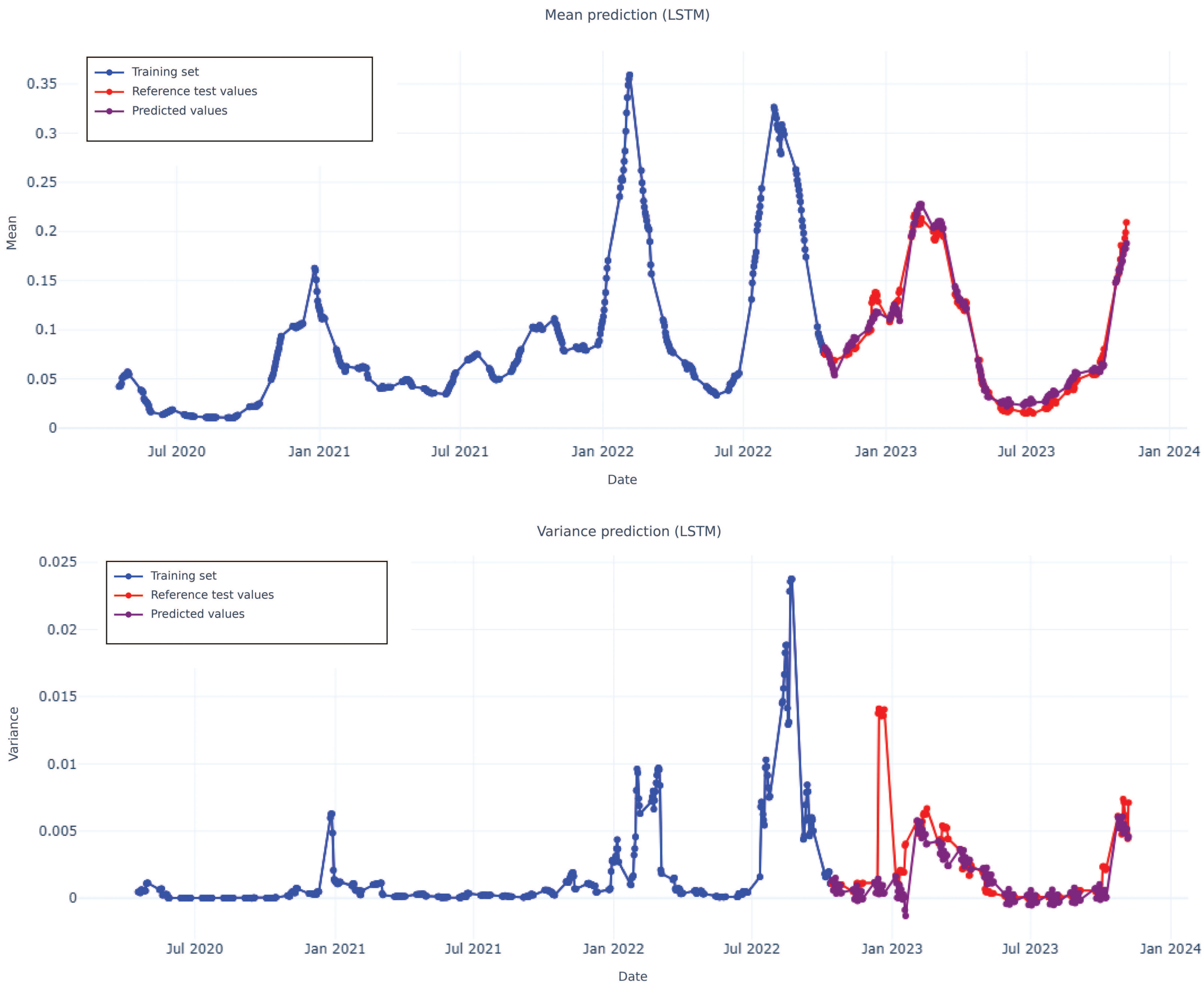

Fig. 5. Mean and variance prediction results for the COVID-19 incidence time series.

Tables 4 and 5 show that LSTM performs better than GRU. However, the difference between the models is small, so both approaches can be used. The mean value is forecast more accurately than the variance. As expected, the 28-day horizon gives worse results than the 14-day horizon. Nevertheless, the results remain satisfactory for mean-value prediction.

6. CONCLUSION

This paper considered methods for forecasting critical transitions in epidemic time series. The methods combine early warning signals with machine learning. Two main approaches were identified. The first is binary classification, which determines whether a critical transition point is near. The second is predictive modeling of the mean and variance of incidence indicators. In epidemic surveillance systems, predictive modeling is most useful at the early stages of circulation of novel pandemic strains. At this stage, the seasonality of the infection may not yet be known. Individual epidemic waves may also be difficult to distinguish. In this case, the next increase in incidence can be inferred from interval forecasts of the mean and variance of new infection cases. Later, a new disease may become seasonal, with clear periods of epidemic increase and decline. Then it becomes possible to estimate the threshold incidence that separates epidemic waves. Retrospective data can then be used to label time-series segments. In this case, we propose switching to simpler classification-based outbreak detection algorithms.

The experimental results show that Easy Ensemble and EWSNet are the most effective models for the classification task. Their metrics change only slightly across the considered threshold values. They also show high recall and satisfactory precision. For incidence forecasting, recurrent neural networks, including LSTM and GRU, model the mean value effectively. However, they have difficulty predicting variance. Variance is less smooth and more sensitive to fluctuations, so it is a harder target for these models. The values of $NRMSE$ and R^2 indicate satisfactory prediction of mean-value dynamics. LSTM slightly outperforms GRU.

This study shows that EWS calculation can be combined with machine learning to forecast critical transitions in epidemic systems. The results may be useful not only in public health but also in other fields where critical phenomena must be understood and forecast. Future work will expand the test datasets and examine improved model architectures. This should increase the robustness and applicability of the proposed methods in other scenarios. One promising direction is the use of multivariate machine learning methods, including methods based on multimodal data. Such models may account for temperature and humidity [29], disease-related user search queries [26], and retrospective incidence data. They may also include data on telemedicine consultations for cough and emergency-department visits for bronchiolitis [27]. In Russian epidemic surveillance, these approaches are limited by restricted data availability and low data quality. However, the models can already be implemented and tested on synthetic data. This would prepare them for future use as data collection improves and new demographic and epidemic datasets become available.

FUNDING

This work was supported by the Russian Science Foundation under Agreement no. 24-25-20104 dated April 12, 2024. It was also supported by the St. Petersburg State Autonomous Institution "Fund for the Support of Scientific, Scientific-Technical, and Innovative Activities" through a regional grant for co-financing the project under Agreement no. 24-25-20104 dated May 22, 2024, agreement no. 169 dated June 3, 2024.

REFERENCES

1. Ramadevi, B. and Bingi, K., Chaotic time series forecasting approaches using machine learning techniques: A review, *Symmetry*, 2022, vol. 14, no. 5, article no. 955.

2. Leonenko, V.N., When a pandemic comes to town: Combating propagation of novel viruses in Russian cities using statistical and mathematical modeling, in *Proc. 2024 IEEE 3rd Int. Conf. on Problems of Informatics, Electronics and Radio Engineering (PIERE)*, IEEE, 2024, pp. 990–995.

3. Bulychev, Yu.G., Signal Recognition without State Space Expansion Based on Observations Containing a Singular Interference: The Case of Nonlinear Parameters of Basis Functions, *Autom. Remote Control*, 2024, vol. 85, no. 2, pp. 147–161.

4. Bury, T., Bauch, C.T., and Anand, M., Detecting and distinguishing tipping points using spectral early warning signals, *J. R. Soc. Interface*, 2020, vol. 17, no. 170, article no. 20200482.

5. Southall, E., Brett, T.S., Tildesley, M.J., et al., Early warning signals of infectious disease transitions: A review, *J. R. Soc. Interface*, 2021, vol. 18, no. 182, article no. 20210555.

6. Zhang, X., Yang, L., Chen, T., et al., Predicting influenza-like illness trends based on sentinel surveillance data in China from 2011 to 2019: A modelling and comparative study, *Infect. Dis. Model.*, 2024, vol. 9, no. 3, pp. 816–827.

7. Alfred, R. and Obit, J.H., The roles of machine learning methods in limiting the spread of deadly diseases: A systematic review, *Heliyon*, 2021, vol. 7, no. 6, article no. e07371.

8. Adugna, A., Abebaw, D., Abebaw, A., and Jemal, M., Deep learning architectures for influenza dynamics and treatment optimization: A comprehensive review, *Front. Artif. Intell.*, 2025, vol. 8, article no. 1521886.

9. Zhao, Q., Moniz, N., Korotasz, A., et al., Early warning signals of emerging infectious diseases, *bioRxiv*, 2025, article no. 2025.03.03.641350.

10. Bizzotto, A., Guzzetta, G., Marziano, V., et al., Increasing situational awareness through nowcasting of the reproduction number, *Front. Public Health*, 2024, vol. 12, article no. 1430920.

11. Gao, S., Chakraborty, A.K., Greiner, R., et al., Early detection of disease outbreaks and non-outbreaks using incidence data: A framework using feature-based time series classification and machine learning, *PLOS Comput. Biol.*, 2025, vol. 21, no. 2, article no. e1012782.

12. Drake, K.O., Boyd, O., Franceschi, V.B., et al., Phylogenomic early warning signals for SARS-CoV-2 epidemic waves, *EBioMedicine*, 2024, vol. 100, article no. 104939.

13. O'Brien, D.A. and Clements, C.F., Early warning signals predict emergence of COVID-19 waves, *medRxiv*, 2021, article no. 2021.06.24.21259444.

14. Kouprianov, A., COVID-19.SPb.monitoring. Monitoring COVID-19 epidemic in St. Petersburg, Russia: Data and scripts, 2021. https://github.com/alexei-kouprianov/COVID-19.SPb.monitoring.

15. Brett, T., Ajelli, M., Liu, Q.H., et al., Detecting critical slowing down in high-dimensional epidemiological systems, *PLOS Comput. Biol.*, 2020, vol. 16, no. 3, article no. e1007679.

16. Bury, T.M., ewstools: A Python package for early warning signals of bifurcations in time series data, *J. Open Source Softw.*, 2023, vol. 8, no. 82, article no. 5038.

17. Bury, T., Sujith, R.I., Pavithran, I., et al., Deep learning for early warning signals of tipping points, *PNAS*, 2021, vol. 118, no. 39, article no. e2106140118.

18. Deb, S., Sidheekh, S., Clements, C.F., Krishnan, N.C., and Dutta, P.S., Machine learning methods trained on simple models can predict critical transitions in complex natural systems, *R. Soc. Open Sci.*, 2022, vol. 9, no. 2, article no. 211475.

19. Liu, X.Y., Wu, J., and Zhou, Z.H., Exploratory undersampling for class-imbalance learning, *IEEE Trans. Syst. Man Cybern. B Cybern.*, 2009, vol. 39, no. 2, pp. 539–550.

20. Seiffert, C., Khoshgoftaar, T.M., Van Hulse, J., and Napolitano, A., RUSBoost: A hybrid approach to alleviating class imbalance, *IEEE Trans. Syst. Man Cybern. A Syst. Hum.*, 2010, vol. 40, no. 1, pp. 185–197.

21. Maclin, R. and Opitz, D., An empirical evaluation of bagging and boosting, in *Proc. Fourteenth Natl. Conf. on Artificial Intelligence*, 1997, pp. 546–551.

22. Bishop, C., *Pattern Recognition and Machine Learning*, Berlin: Springer, 2006.

23. Deb, S., Sidheekh, S., Clements, C.F., Krishnan, N.C., and Dutta, P.S., Machine learning methods trained on simple models can predict critical transitions in complex natural systems, *R. Soc. Open Sci.*, 2022, vol. 9, no. 2, article no. 211475.

24. Lim, S., Giorgini, L., Moon, W., and Wettlaufer, J.S., Predicting critical transitions in multiscale dynamical systems using reservoir computing, *Chaos*, 2020, vol. 30, no. 12, article no. 123126.

25. Kong, L.-W., Fan, H.W., Grebogi, C., and Lai, Y.C., Machine learning prediction of critical transition and system collapse, *Phys. Rev. Res.*, 2021, vol. 3, no. 1, article no. 013090.

26. Su, K., Xu, L., Li, G., et al., Forecasting influenza activity using self-adaptive AI model and multi-source data in Chongqing, China, *EBioMedicine*, 2019, vol. 47, pp. 284–292.

27. Morbey, R.A., Todkill, D., Moura, P., et al., Using machine learning to forecast peak health care service demand in real-time during the 2022–23 winter season: A pilot in England, UK, *PLOS One*, 2025, vol. 20, no. 1, article no. e0292829.

28. Cho, G., Park, J.R., Choi, Y., et al., Detection of COVID-19 epidemic outbreak using machine learning, *Front. Public Health*, 2023, vol. 11, article no. 1252357.

29. Amendolara, A.B., Sant, D., Rotstein, H.G., and Fortune, E., LSTM-based recurrent neural network provides effective short term flu forecasting, *BMC Public Health*, 2023, vol. 23, no. 1, article no. 1788.

30. Lapeyrolerie, M. and Boettiger, C., Limits to ecological forecasting: Estimating uncertainty for critical transitions with deep learning, *Methods Ecol. Evol.*, 2022, vol. 14, no. 3, pp. 785–798.

31. Sangiorgio, M. and Dercole, F., Robustness of LSTM neural networks for multi-step forecasting of chaotic time series, *Chaos Solitons Fractals*, 2020, vol. 139, article no. 110045.

This paper was recommended for publication by A.I. Mikhalski, a member of the Editorial Board