

Dmytro CHUMACHENKO<sup>1</sup>, Tetiana DUDKINA<sup>1</sup>, Tetyana CHUMACHENKO<sup>2</sup>

<sup>1</sup> National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine

<sup>2</sup> Kharkiv National Medical University, Kharkiv, Ukraine

## ASSESSING THE IMPACT OF THE RUSSIAN WAR IN UKRAINE ON COVID-19 TRANSMISSION IN SPAIN: A MACHINE LEARNING-BASED STUDY

*COVID-19 pandemic has significantly impacted the world, with millions of infections and deaths, healthcare systems overwhelmed, economies disrupted, and daily life changed. Simulation has been recognized as a valuable tool in combating the pandemic, helping to model the spread of the virus, evaluate the impact of interventions, and inform decision-making processes. The accuracy and effectiveness of simulations depend on the quality of the underlying data, assumptions, and modeling techniques. Ongoing efforts to improve and refine simulation approaches can enhance their value in addressing future public health emergencies. The Russian full-scale military invasion of Ukraine on February 24, 2022, has created a significant humanitarian and public health crisis, with disrupted healthcare services, shortages of medical supplies, and increased demand for emergency care. The ongoing conflict has displaced millions of people, with Spain ranking 5th in the world for the number of registered refugees from Ukraine. The research aims to estimate the impact of the Russian war in Ukraine on COVID-19 transmission in Spain using means of machine learning. The research is targeted at COVID-19 epidemic process during the war. The research subjects are methods and models of epidemic process simulation based on machine learning. To achieve the study's aim, we used forecasting methods and built a model of COVID-19 epidemic process based on the XGBoost method. As a result of the experiments, the accuracy of forecasting new cases of COVID-19 in Spain for 30 days was 99.79 %, and the death cases of COVID-19 in Spain – were 99.86 %. The model was applied to data on the incidence of COVID-19 in Spain for the first 30 days of the war escalation (24.02.2022 – 25.03.2022). The calculated forecasted values showed that the forced migration of the Ukrainian population to Spain, caused by the full-scale invasion of Russia, is not a decisive factor affecting the dynamics of the epidemic process of COVID-19 in Spain. Conclusions. The paper describes the results of an experimental study assessing the impact of the Russian full-scale war in Ukraine on COVID-19 dynamics in Spain. The developed model showed good performance to use it in public health practice. The analysis of the obtained results of the experimental study showed that the forced migration of the Ukrainian population to Spain, caused by the full-scale invasion of Russia, is not a decisive factor affecting the dynamics of the epidemic process of COVID-19 in Spain.*

**Keywords:** epidemic model; epidemic process; epidemic simulation; simulation; COVID-19; XGBoost; war.

### Introduction

COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, has profoundly impacted the world since its emergence in late 2019. The virus quickly spread globally, resulting in millions of infections and deaths, overwhelming healthcare systems, and disrupting economies and daily life [1]. The pandemic has highlighted inequalities and vulnerabilities in societies, with marginalized populations bearing a disproportionate disease burden. Scientists and public health experts have worked tirelessly to develop vaccines, treatments, and strategies to mitigate the spread of the virus [2]. However, challenges such as vaccine hesitancy, new variants, and ongoing political and social tensions complicate the response [3]. The pandemic has underscored the critical importance of global cooperation and preparedness in addressing public health emergencies.

Multiple approaches have been used to combat COVID-19 pandemic, including non-pharmaceutical interventions such as mask-wearing, social distancing, quarantine measures, vaccination campaigns, contact tracing, and testing [4]. Various pharmaceutical interventions have been developed, including antiviral drugs and monoclonal antibodies, to treat COVID-19 patients [5]. The development and distribution of vaccines have been critical tools in controlling the spread of the virus, with numerous vaccines now available globally [6]. Public health communication and education campaigns have also been crucial in disseminating information about the virus, promoting adherence to prevention measures, and addressing vaccine hesitancy [7]. International collaboration and cooperation have been essential in addressing the global nature of the pandemic, and ongoing efforts must combat the virus's evolving variants and ensure equitable access to interventions.

The pandemic has stimulated the development of data-driven medicine field. Researchers worldwide have developed methods for medical data analysis [8], analysis of COVID-19 in other fields of human beings [9], social factors impact estimation [10], medical image processing [11], automated diagnosis [12], investigation of virus behavior [13], infodemics [14], and infectious dynamics simulation [15], etc.

Simulation has been recognized as a valuable tool in combating COVID-19 pandemic. Simulation-based approaches have been used to model the spread of the virus, evaluate the impact of interventions, and inform decision-making processes [16]. These simulations have been used to forecast the number of infections and deaths, estimate the resource needs of healthcare systems, and predict the effects of policy interventions. Simulation-based training has also prepared healthcare workers and emergency responders for the challenges of caring for COVID-19 patients. By allowing stakeholders to test different scenarios and interventions, simulation has provided critical insights into the pandemic and has helped guiding response efforts. However, the accuracy and effectiveness of simulations depend on the quality of the underlying data, assumptions, and modeling techniques and the limitations of simulations must be recognized [17]. Simulation has been an essential tool in combating COVID-19 pandemic, and ongoing efforts to improve and refine simulation approaches can enhance their value in addressing future public health emergencies.

On February 24, 2022, Russia launched a full-scale invasion of Ukraine, leading to significant violence and casualties. Russian war has destroyed cities, displaced millions, and created a severe humanitarian crisis in Ukraine. The impact of the military invasion has also extended to the public health sector, with challenges including disrupted healthcare services, shortages of medical supplies, and increased demand for emergency care [18]. The escalation of war has had a profound impact on the lives of civilians in the region, and ongoing efforts are needed to address the war's humanitarian and public health consequences.

More than 8 million refugees from Ukraine are recorded across Europe, and more than 5 million people are estimated to be displaced in Ukraine, according to the United Nations [19]. Spain ranks 5th in the world regarding the number of registered refugees from Ukraine. The number of refugees in Spain at the beginning of 2023 is 170 thousand [20].

Thus, the paper estimates the impact of the Russian war in Ukraine on COVID-19 transmission in Spain using means of machine learning. This research is targeted at COVID-19 epidemic process during the war. The research subjects are methods and models of epidemic process simulation based on machine learning.

To achieve the aim of the research following tasks have been formulated:

1. Methods and models of COVID-19 epidemic process analysis in Spain should be analyzed.
2. Data on COVID-19 pandemic in Spain should be analyzed.
3. Impact of the Russian war on public health should be analyzed.
4. A simulation model of COVID-19 epidemic process based on the XGBoost method should be developed.
5. Tuning and verification of the proposed model should be done.
6. Estimating COVID-19 epidemic process dynamics in Spain during the first 30 days of the full-scale Russian war in Ukraine should be done.
7. Results of the experimental study should be analyzed.

The respective contribution of the research is three-fold. First, developing models based on the XGBoost method will allow estimating the accuracy of ensemble machine learning methods applied to the simulation of COVID-19 epidemic process. Second, using the actual data on COVID-19 epidemic process dynamics in Spain will allow us to investigate the pandemic character. Third, the proposed methodology will allow us to assess the impact of migration flows caused by the Russian war in Ukraine on COVID-19 dynamics in Spain.

In this paper, section 1, namely, the current research analysis provides the current state of COVID-19 epidemic process simulation methods and models applied to the pandemic in Spain. Section 2, namely Data, describes the background of COVID-19 epidemic process in Spain, relation of the Russian war and public health, and analyses the data sources of COVID-19 morbidity applicable for given research. Section 3, namely Materials and Methods, provides a methodology of research, an overview of the XGBoost method, and metrics of the model's performance evaluation. Section 4 provides the results of forecasting COVID-19 confirmed and deaths in Spain by the developed model. The discussion section discusses the obtained results from the epidemiology view. Conclusions describe the outcomes of the investigation.

## 1. Current Research Analysis

Mathematical and computational models have been used extensively to investigate the epidemic processes of infectious diseases. These models provide a quantitative understanding of how diseases spread, how interventions can control transmission, and how public health policies can be optimized. Models are built on assumptions and simplifications, ranging from simple compartmental models to complex, agent-based simulations. Model outputs can make predictions, test hypotheses, and inform decision-making processes. In the case of COVID-19

pandemic, models have played a crucial role in forecasting the number of cases and deaths, assessing the impact of non-pharmaceutical interventions, evaluating vaccine distribution strategies, and identifying populations at high risk [21]. While modeling has limitations, it remains an essential tool for understanding the dynamics of infectious diseases and designing effective control strategies.

Since the beginning of the pandemic, many researchers have used the simulation tool to study the epidemic process of COVID-19 in Spain.

Paper [22] uses a SIRD model to analyze the evolution of COVID-19 in Spain and make reliable short-term predictions. The model confirms quarantine's effectiveness and shows the infection's evolution in various stages. The analysis provides valuable information to influence the epidemic's evolution and perform actions to reduce its harmful effects. Article [23] also uses the compartment approach and proposes a new SEIRS dynamical model with a vaccine rate to investigate the impact of vaccines on COVID-19 pandemic. The model is formulated with integer order and generalized with Atangana-Baleanu fractional derivative. The existence and stability of the solution were analyzed using the Predictor-Corrector algorithm. The simulations are based on actual data from Spain and demonstrate the effectiveness of vaccines in controlling the spread of COVID-19.

The paper [24] discusses the emergence of COVID-19 coronavirus variants and the need to estimate the case fatality ratio during such epidemics. The authors use non-parametric approaches to estimate odds ratios and confidence intervals for COVID-19 data and apply a non-parametric cure model to estimate cure rates. Simulations based on R-software predict the effects of patients based on covariates in different scenarios and identify risk factors associated with COVID-19 mortality, including older age, male sex, comorbidities, and access to timely healthcare. The authors emphasize the importance of accurate estimates and consistent messaging to avoid inducing panic in significant outbreaks of COVID-19.

The paper [25] applies resampling techniques to a modified compartmental SEIR model to account for undetected infected individuals during an epidemic. The authors conducted numerical simulations on the first wave of COVID-19 in Spain, showing that point estimates obtained through bootstrap samples improve upon the original data. The relative error of the detected currently infected people is reduced from 0.061 to 0.0538 for the mean overall bootstrap estimated series. Using resampling techniques provides a more accurate estimation of epidemic parameters and can be helpful in decision-making processes during pandemics. The authors of the paper [26] present a new dynamic probabilistic model based on simulation models to estimate the prevalence of SARS-CoV-2 infections in different compartments. The

proposed model is optimized using gradient-free stochastic black-box optimization techniques simulating multiple trajectories of the infection dynamics and is applied to the first and successive waves of the pandemic in Spain. The model confirms the accuracy of estimating the seroprevalence and allows for dynamically examining the population's exposure to the virus.

Research [27] proposes a deep learning approach to model network contagion dynamics by learning effective local mechanisms from time series data. The graph neural network architecture used in this approach makes minimal assumptions about the dynamics and can simulate arbitrary network structures. The accuracy of the proposed approach is demonstrated using different contagion dynamics, and the applicability is illustrated using actual data from COVID-19 outbreak in Spain. The paper suggests that deep learning offers a new and complementary perspective to model contagion dynamics on networks. The paper [28] presents an agent-based COVID-19 model to test various sampling strategies in a scenario including the 62 largest cities in Spain. The evaluation analyzes the impact of different parameters, such as testing frequency, quarantine time, percentage of quarantine breakers, or group testing, on sampling efficacy. The most effective strategies are pooling, rapid antigen test campaigns, and requiring negative testing for access to public areas.

The authors of [29] propose a mathematical approach to analyze the geographical distribution of COVID-19 cases during the first wave of the pandemic in Galicia, Spain. The approach assigns a risk factor for each municipality based on geographic and demographic information to assess the risk of future outbreaks. The results are compared with empirical data and demonstrate the method's validity. The paper [30] describes an SEIR-type model that analyzes the interaction between different vaccination scenarios and levels of protection measures on disease incidence. The model incorporates a range of factors such as protection levels, quarantine measures, and vaccination efficacy of four types of vaccines. The model was verified and evaluated using a response timeline and vaccination strategies in the Basque Country, Spain. The simulations showed significant differences in disease prevalence and cumulative mortality between vaccination scenarios for low and medium-level protection measures. However, any vaccine scenario limits virus transmission and disease mortality for high-level protection measures.

The paper [31] analyzes the spread of the Omicron COVID-19 variant in Spain using an agent-based simulator EpiGraph. The simulation includes a population of the 63 most populated cities of Spain. The study analyzes the propagation of Omicron under different scenarios and transportation intensities, concluding that the Omicron

variant spreads quickly throughout the country, regardless of the initial location and transportation conditions. The study [32] explores the fractional dynamics of COVID-19 outbreaks in Spain between February and October 2020 using the Atangana-Baleanu non-integer order derivative. The authors analyzed optimal control problems in a fractional sense and provided graphical simulations using the Predictor-Corrector method. They also analyzed the existence of a unique solution and the stability of the proposed scheme.

The paper [33] presents a discrete event simulation model for short-term planning of hospital resource needs, specifically ICU beds, during outbreaks such as COVID-19 pandemic. The simulation model consists of stochastic modeling of patient admission and patient flow processes, with a Gompertz growth model used to represent the patient arrival process. The model was tested in two Autonomous Regions of Spain during the two COVID-19 waves experienced in 2020, with daily predictions informing regional logistic healthcare planning teams. The Gompertz model better fits pandemic-related data and has superior prediction capacity compared to other sigmoid models based on Richards, Logistic, and Stannard functions. The study [34] presents a Bayesian inference framework to quantify uncertainties in epidemiological models, specifically SEIJR and SIJR models involving populations of susceptible, exposed, infective, diagnosed, dead, and recovered individuals, in response to lockdown measures. The study tracks the time evolution of Covid-19 in Spain using official data. It shows that transmission and recovery rates vary in response to facts,

and the diagnosis rate is relatively low, leading to large amounts of undiagnosed infective individuals. The paper concludes that confinement can delay but not stop the spread of the virus unless a large enough fraction of the population is confined until the asymptomatic population is depleted and can be replaced by solid distancing through masks.

Research [35] identifies critical communities more likely to suffer COVID-19 outbreaks in Spain and estimate the impact of different mobility restrictions using an SEIR compartmental model. They derive a mathematical descriptor based on the linear stability theory that quantifies the potential impact of becoming a hotspot. They find that the confinement of economically non-active individuals may significantly reduce risk, whose effects are equivalent to the confinement of the total population. The study was conducted across the totality of municipalities in Spain during the second wave from August 1st to November 30th, 2020. The paper [36] presents a new model for predicting the spread of COVID-19 at the municipal level, incorporating the inherent stochasticity of the virus and using a chemical reaction-based approach. The model considers variable infection rates based on mitigation policies, multiple infections from a single infected person, and viral load data from wastewater treatment plants. The model was successfully employed for predicting the pandemic's evolution in small and medium-sized municipalities in Galicia, Spain.

Table 1 presents the analysis of the models of COVID-19 propagation in Spain.

Table 1

The overview of COVID-19 epidemic process models in Spain

Paper	Task	Method	Findings
Martinez V. [22]	To forecast the evolution of COVID-19 during first months of pandemic in Spain.	Compartmental approach	The proposed model allows adjusting the parameters in real time when the disease is latent
Kumar P., et. al. [23]	To justify the role of vaccines in combating the pandemic.	Compartmental approach	The authors exemplified that the given model works well to show the role of vaccine on the dynamics of COVID-19 in the population. We observed that in both cases (disease-free equilibrium and endemic equilibrium) vaccine is effective and when the vaccine rate increases then the population goes into a comfortable environment.
Ghosh S., et. al. [24]	To estimate relation between mortality due to COVID-19 and age specific gender using non-parametric approach.	Machine learning (logistic regression)	The impact of sex on mortality from disease epidemics is important, but studies using individual-level data are needed to confirm an interaction between age and sex in COVID-19 mortality in order to guide clinical care personal and to address questions of whether men require additional surveillance, prevention, and earlier intervention than women.

Continuation of Table 1

Paper	Task	Method	Findings
Morillas F., et. al. [25]	To apply resampling to a compartmental model of differential equations to improve the estimates of the variables in the presence of uncertainty.	Compartmental approach	The proposed procedure allows managing the uncertainty of the data in a more appropriate way than using the punctual estimation by considering only observed data. The approximations performed using the wavelet graduation and the bootstrap technique are more consistent or robust
Matabuena M., et. al. [26]	To retrospectively reconstruct the dynamics of infections while estimating seroprevalence in the different compartments of the model to better characterize the concrete mechanisms of virus transmission in the territories analyzed.	Compartmental approach	Research illustrates the potential of combining simulation models with a Black-Box optimization techniques to obtain new epidemic models. More reliable inferences by fitting the model at simultaneous locations, for example, through multilevel simulation models can be obtained.
Murphy C., et. al. [27]	To provide predictions for previously unseen network structures, therefore allowing the exploration of the properties of the learned dynamics beyond the training data.	Deep learning (graph neural network)	Many standard graph neural network architectures do not correctly handle the problem of learning contagion dynamics from time series. However, the proposed approach performs better when trained on data whose underlying network structure is heterogeneous, which could prove useful in real-world applications of the proposed method given the ubiquitousness of scale-free networks
Guzman-Merino M., et. al. [28]	To understand the effects of different possible sampling strategies that contribute to design policies that can slow down the propagation and reduce the number of affected individuals.	Agent-based approach	Design and implementation of a more refined social interaction model, which reflects profession-dependent connections to increase the realism of population mixing. The authors also used contact matrices extracted from public surveys to reflect age-dependent interactions of each individual.
Carballosa A., et. al. [29]	To show the difference that geographical placement and demographic circumstances has on the impact of the virus spread.	Compartmental approach	The developed tools can be potentiality used as a workbench to test different policies and their implications in the risk factors at a node level to prepare for future COVID-19 waves and new pandemics.
Ganga A., Bidegain G. [30]	To analyze the impact of the interaction between different vaccination scenarios, regarding vaccination rate and efficacy, and different levels of nonpharmaceutical protection measures (from the use of mask to lockdown) on disease incidence and mortality.	Compartmental approach	The results suggest that in an initial vaccination phase (30%–50% of the population is vaccinated), COVID19 incidence, as measured in daily cases and cumulative mortality, importantly decreases when vaccination and a high level of nonpharmaceutical interventions are in place. That is, in the first vaccination phase, with vaccination, strong mobility restrictions and closure of indoor facilities such as public spaces, restaurants, and bars are critical to significantly control disease outbreaks.
Merino M.G., et. al. [31]	To evaluate the effectiveness of different policies that restrict the movement of people between different cities in Spain for the time interval comprised between May 2021 and March 2022.	Agent-based approach	Independently of the initial location of the Omicron variant, and the existing transportation conditions, the high transmissibility of Omicron variant -about 2 times larger than Delta, and roughly the 3.6 times more transmissible than the initial COVID-19 strain, allows it to quickly spread through the country and become the dominant strain in all territories.

End of Table 1

Paper	Task	Method	Findings
Kumar, P., et al. [32]	To analyze the structure of the novel coronavirus in Spain at different transmission rates and to suggest optimal control measures to combat the disease.	Compartmental approach	A brief idea of the structure of novel coronavirus in Spain by using different infection transmission rates and to indicate the danger of these deadly diseases for future precautions with the introduction of some healthcare measures has been given.
Garcia-Vicuna D., et al. [33]	To enable tactical decisions for the provision of specialized health resources during the current outbreak with a focus on the transition period to generate a short-term projection of the current state of the hospital.	Population growth model	Paper shows how operations research can contribute to a rapid respond to a healthcare crisis by reporting on a successful real-world application of the simulation model to support a decision-making process of crucial importance to the health of patients in two Autonomous Regions of Spain.
Carpio A., Pierret E. [34]	To analyze variations in the model rates and in the distributions of the different populations, a time grows because of the measures implemented, including undiagnosed infected and asymptomatic individuals.	Compartmental approach	Confinement can delay but not stop the spread of the virus unless a large enough fraction of the population is confined until the asymptomatic population is depleted and can be replaced by strong distancing through masks.
Carballosa A., et al. [35]	To evaluate the spatial behavior of the virus across any region by considering different administrative units such as municipalities in the Spanish case.	Compartmental approach	The results show how the confinement of the economically non-active individuals results in a significant reduction of risk, whose effects are equivalent to the confinement of the total population.
Pajaro M., et al. [36]	To predict the evolution at the local level (municipalities), and for short term time horizons (7–10 days maximum).	Compartmental approach	The model can be used to test different scenarios and for forecasting in small and medium size municipalities with Wastewater Treatment Plants in time horizons of 7–10 days.

Although most models of COVID-19 dynamics in Spain are based on a compartmental approach, the methodology proposed in this study is based on machine learning methods. Machine learning models have advantages over compartmental models in infectious disease – spreading simulations.

Machine learning models are more flexible and can handle more complex data structures, such as unstructured and heterogeneous data, which makes them suitable for analyzing data from multiple sources [37]. Machine learning models can learn patterns from data and make predictions based on these patterns. They can identify hidden patterns that may not be visible in compartmental models. Machine learning models can handle large amounts of data and can be easily scaled up to handle bigger or more complex models. Machine learning models are less sensitive to outliers and can handle missing data more effectively than compartmental models. Machine learning models can improve the accuracy of predictions by learning from past data and using multiple data sources. This makes them better suited for predicting

infectious disease outbreaks and determining the best course of action.

## 2. Data

### 2.1. Background on COVID-19 Pandemic in Spain

Spain has had over 13 million confirmed cases and almost 120 thousand deaths from COVID-19 as of the beginning of 2023 [38]. The virus was first confirmed in Spain in January 2020, and community transmission began by mid-February [39]. A lockdown was imposed in mid-March, and by late March, Madrid had the most cases and deaths in the country [40]. Medical professionals and retirement home residents were at a higher risk of infection. Spain had the most COVID-19 deaths in a single day by any country in early April. The daily death toll fell below 100 for the first time in May 2020, and Spain lifted the state of alarm in June. However, some restrictions were later reimposed due to increased cases [41].

The number of infections and deaths may have been underestimated due to a lack of testing and reporting. A government seroprevalence study in July 2020 showed that 5.2 % of the population, or about two million people, could have been infected during the pandemic [42]. Spain was the second country in Europe to record half a million cases and had over one million COVID-19 cases by October 2020.

COVID-19 outbreak in Spain resulted in several cities and regions being put on lockdown in March 2020, with the Catalan government quarantining four municipalities after cases were reported at a hospital [43]. On 14 March, a state of alarm was declared across the country, which led to the central government taking control and implementing measures such as banning nonessential activities and large gatherings. However, citizens could still travel to work and buy essential items. Autonomous communities also announced their emergency measures, such as population confinement and closure of ports and airports. On 25 March, the parliament approved the government's request to extend the state of alarm until 11 April. Some nonessential workers were allowed to return to work on 13 April, and the government attempted to reduce the spread of the virus through contact tracing and to distribute masks [44].

In March 2020, the Government of Spain implemented several travel restrictions in response to COVID-19 pandemic [45]. On 16 March, the Spanish borders were closed, allowing only Spanish citizens and those needing to enter the country. Flights from Italy to Spain were canceled until 25 March, and all traffic between Morocco and Spain was suspended. The Balearic Islands closed all airports and ports, except for a few exceptions, and the Canary Islands restricted flights between the peninsula and its islands. Additionally, movement between provinces was forbidden until at least the end of June.

Spain has a high percentage of its population vaccinated against COVID-19, 87 % fully vaccinated, and 88% with the first dose. Spain is also a country most in favor of COVID-19 vaccines [46].

## **2.2. Relations between the Russian War and the Public Health**

On February 24, 2022, Russia launched a full-scale military invasion of Ukraine for no reason. Russia uses a military strategy that targets civilians and civilian infrastructure, as exemplified by Russian attacks in Ukraine and documented by the media and the United Nations [47]. Almost one year after the start of the war escalation, the Office of the UN High Commissioner for Human Rights recorded 7155 civilian fatalities and 11 662 injuries in Ukraine. However, the actual figures are likely to be higher [48].

The quality of healthcare and service delivery in some areas of Ukraine has declined due to the war, with facilities lacking running water and electricity. Hospitals in eastern Ukraine cannot provide essential care, causing patients to travel further for treatment or not seek care at all [49]. Childhood vaccination coverage has declined, posing a risk for the spread of diseases such as polio and measles. The conflict has also made it harder to manage chronic diseases, with many displaced people stopping their medications. Psychological trauma has been reported among civilians and soldiers, affecting mental health. It is unclear how access to HIV services and other essential medicines is affected in regions under temporary Russian occupation.

Over 760 attacks in Ukraine have been verified by World Health Organization Surveillance System for Attacks on Health Care as of the beginning of 2023, with almost 700 attacks impacting health facilities [50]. The situation is complicated in areas close to the front line or those that the Government of Ukraine has recently retaken. In these areas, most health facilities are not operating. The health system struggles to meet the complex needs of those who remain, such as the elderly, people with low mobility, and those with physical disabilities [51].

As of January 2023, the United Nations High Commissioner for Refugees has recorded 8.0 million refugees from Ukraine in Europe [48]. 4.8 million were registered for temporary protection or similar national protection schemes in Europe. For cross-border movements, 18.2 million crossings out of Ukraine and 10.0 million crossings into Ukraine were recorded. These latter figures do not reflect individuals. The International Organization for Migration estimates that 5.4 million people are internally displaced across Ukraine as of January 2023 [52].

Forced refugees from the Russian war in Ukraine face challenges that impact public health. Displacement and overcrowding in refugee shelters and evacuation transport can increase the risk of infectious disease outbreaks. The lack of access to clean water, sanitation, and hygiene facilities can also contribute to the spreading of diseases such as dysentery. Furthermore, refugees may have limited access to healthcare and medications, increasing the risk of untreated chronic and infectious diseases. Mental health issues are also a concern, as refugees may experience trauma and stress related to conflict and displacement. Additionally, refugees face language and cultural barriers that can make accessing healthcare and other essential services difficult.

Although Spain does not border Ukraine, the country ranks 5th in the world regarding the number of registered refugees from Ukraine. The number of refugees in Spain at the beginning of 2023 was 170 thousand [20].

### 2.3. Data Sources Analysis

Throughout COVID-19 pandemic, many researchers worldwide have used the Johns Hopkins University & Medicine Coronavirus Resource Center [53] as their primary data source to model the dynamics of COVID-19. The Johns Hopkins University & Medicine Coronavirus Resource Center dashboard is an up-to-date source of information on COVID-19 pandemic, offering real-time data on confirmed cases, deaths, recoveries, and country-specific information. The dashboard is frequently updated, and the resource center provides expert insights, infographics, and articles on various aspects of the pandemic.

However, data on the incidence of COVID-19 in Ukraine after February 24, 2022, are not available on the portal. Therefore, in this study, we used data obtained from the WHO Coronavirus (COVID-19) Dashboard of World Health Organization data [54].

The WHO Coronavirus (COVID-19) Dashboard is an online platform. The World Health Organization (WHO) provides to monitor the global spread of COVID-19. The dashboard provides real-time updates on the number of confirmed cases, deaths, and recoveries reported by countries worldwide. The data are presented in easy-to-read graphs and maps that allow users to track the progression of the pandemic in different regions of the world. The dashboard also provides information on the testing capacity and measures taken by countries to contain the spread of the virus. It is regularly updated with the latest information from the WHO and other reliable sources.

We used cumulative and daily data on reported cases and deaths from COVID-19.

## 3. Materials and Methods

### 3.1. Methodology

The methodology proposed in this study makes it possible to assess changes in the dynamics of COVID-19 epidemic process caused by the forced migration of the population of Ukraine due to Russia's military invasion of Ukraine. The assessment uses a predictive model of morbidity and mortality from COVID-19 in the selected territory, which is applied to different periods: 30 days before the start of the escalation of the Russian war in Ukraine and 30 days after.

As a forecasting method, the machine learning approach was chosen as the most accurate one. Machine learning methods are increasingly being used to forecast the dynamics of infectious diseases. These methods involve algorithms that can automatically learn patterns from data and make predictions based on these patterns.

In infectious disease forecasting, machine learning methods typically use data on past disease outbreaks to train models that can predict future outbreaks. These models can consider a range of factors, such as demographics, climate, and human behavior, that may influence the spread of the disease. By analyzing large amounts of data, machine learning models can identify complex patterns that may not be clear to humans and provide accurate forecasts that can inform public health decision-making.

For this, the ensemble method of machine learning XGBoost was applied. Ensemble machine learning methods can improve the accuracy and robustness of infectious disease forecasting by combining multiple models and their predictions. By leveraging the strengths of various models and integrating their outputs, ensemble methods can reduce individual model biases and errors and provide more accurate predictions [55]. Ensemble methods can provide better uncertainty estimates and identify areas of agreement and disagreement among models, allowing for more informed decision-making by public health officials.

The proposed methodology consists of the following steps:

1. Construction of a machine learning model for forecasting the dynamics of the epidemic process of COVID-19.

2. Tuning and verifying the model on the test sample from January 25, 2022, to February 23, 2022, of morbidity and mortality from COVID-19 in Spain.

3. Construction of a retrospective forecast of the dynamics of morbidity and mortality from COVID-19 in Spain from February 24, 2022, to March 25, 2022.

4. Calculation of the deviation of the forecast morbidity and mortality from COVID-19 in Spain from the actual dynamics of the epidemic process of COVID-19 from February 24, 2022, to March 25, 2022.

5. To analyze the differences in the model's performance for the periods January 25, 2022 - February 23, 2022, and February 24, 2022 - March 25, 2022, and evaluate the factors influencing changes in the dynamics of the epidemic process from the viewpoint of view of epidemiology.

### 3.2. XGBoost Model

Boosting is a machine learning technique that combines weak classifiers to form a more accurate predictor [56]. The basic idea is to train a sequence of weak learners, which are simple models that perform slightly better than random guessing, and then combine their predictions to create a more robust model. Each weak learner is trained on a modified version of the training data that emphasizes the examples that the previous models have classified incorrectly. This approach allows the ensemble

to focus on the most challenging examples to classify and gradually improve the model's overall accuracy.

When applied to forecasting the dynamics of infectious diseases, boosting can improve the accuracy of predictions by reducing errors and capturing complex non-linear relationships between variables. Boosting algorithms are known for handling noisy and high-dimensional data, making them suitable for analyzing complex epidemiological datasets. Boosting models can adjust and update their predictions in real time, allowing for more timely and accurate forecasting of disease dynamics.

Extreme Gradient Boosting or XGBoost is a popular machine learning algorithm for forecasting time series, including the dynamics of infectious diseases. It is an ensemble method that combines the predictions of many individual models to improve overall accuracy. XGBoost is a tree-based algorithm that builds decision trees iteratively, refining each tree's predictions based on previous tree errors [57].

The first step in XGBoost is to build a decision tree that predicts the outcome variable based on input variables. The algorithm then calculates the errors of the tree and assigns weights to the data points based on the basis of their errors. It then builds another tree that predicts the previous tree's errors and updates the data points' weights. This process is repeated until the desired number of trees is built, or the error rate reaches a minimum.

The second step in XGBoost is to combine the predictions of the individual trees. This is done by assigning weights to each tree based on its accuracy and using them to calculate a weighted average of the predictions. The result is a highly accurate model that can be used to forecast the dynamics of infectious diseases. XGBoost also includes a regularization term that penalizes complex models, preventing overfitting and improving generalization performance.

The function for XGBoost optimization is

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t), \quad (1)$$

where  $l$  is loss function;

$y_i$  is value of the  $i$ -th element of the training sample;

$x_i$  is set of features of the  $i$ -th element of the training sample;

$f_t$  is function we want to train on step  $t$ ;

$f_t(x_i)$  is forecasting on the  $i$ -th element of the training sample;

$\Omega(f)$  is regularization of function  $f$ .

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2, \quad (2)$$

where  $T$  is number of nodes in the tree;

$w$  is values in leaves;

$\gamma$  and  $\lambda$  are regularization parameters.

The optimized function  $L^{(t)}$  is approximated using the Taylor expansion to the second term by the following expression:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + g_i f_t(x_i) + 0.5 h_i f_t^2(x_i)) + \Omega(f_t), \quad (3)$$

where

$$g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}, \quad (4)$$

$$h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial^2 \hat{y}_i^{(t-1)}}. \quad (5)$$

The minimum of this expression with respect to  $f_t(x_i)$  is located at the point

$$f_t(x_i) = \frac{-g_i}{h_i}. \quad (6)$$

Each individual tree of the ensemble  $f_t(x_i)$  is trained by a standard algorithm.

XGBoost has several advantages that make it an effective method for the task of infectious disease simulation:

1. It can handle large datasets with high dimensionality and many features. This is important for infectious disease simulation as many variables may need to be considered when predicting disease spread.

2. XGBoost has high accuracy and can make exact predictions, which is essential for forecasting the dynamics of infectious diseases.

3. It can handle missing data, a common issue in infectious disease data.

However, XGBoost also has some disadvantages:

1. It can be computationally intensive and requires much memory and processing power, which may limit its use in some settings.

2. It may be prone to overfitting if not properly optimized, which can lead to inaccurate predictions.

3. XGBoost requires careful tuning of hyperparameters, which can be time-consuming and may require expert knowledge to optimize.

The computational complexity of XGBoost depends on various factors such as the number of trees, depth of the trees, number of features, and size of the dataset. The time complexity of XGBoost is  $O(n * d^2 * k)$ , where  $n$  is the number of samples,  $d$  is the maximum depth of the tree, and  $k$  is the number of trees. Therefore, increasing the number or depth of the trees can significantly increase the computational complexity. However,

XGBoost is designed to be highly optimized and efficient, and it provides several options for controlling the computational complexity, such as subsampling, regularization, and early stopping. Overall, XGBoost is a relatively fast and scalable algorithm that can handle large datasets and complex models.

Nevertheless, this method is suitable for solving the problem posed in the framework of this study since high prediction accuracy is required. Simultaneously, there is no need to use the proposed model in real-time, so the relatively high computational complexity is not a disadvantage in our case.

### 3.3. Model Verification Methods

The model accuracy estimate is calculated using mean absolute percentage error:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|, \quad (7)$$

where  $A_t$  is the actual value;

$F_t$  is the forecasted value.

MAPE is a commonly used metric for evaluating the accuracy of forecasting models, including those used for infectious disease forecasting. The advantage of using MAPE is that it provides a measure of the relative error of the forecast, which is particularly useful when dealing with time series data that exhibit trends or seasonality. This means that the metric considers the actual values' magnitude, providing a more accurate representation of the forecasting error.

To assess the forecast of the incidence of COVID-19 for the period of the war, the deviations of the forecast data from the registered statistics were calculated:

$$D = |F_t - A_t|, \quad (8)$$

where  $A_t$  is the actual value;  $F_t$  is the forecasted value.

## 4. Results

### 4.1. Model Tuning

The XGBoost model was implemented in the Python programming language to verify and adjust the model for predicting the epidemic process of COVID-19. The model's accuracy was evaluated for 3, 7, 14, 21, and 30 days. Data on morbidity and mortality from COVID-19 in Spain obtained from the World Health Organization COVID-19 Dashboard were used for verification. As a test sample, data on incidence and mortality from COVID-19 in Spain from January 25, 2022, to February 23, 2022, i.e., within 30 days before the start of a full-scale Russian invasion of Ukraine.

Figure 1 presents the results of the retrospective forecast of cumulative new cases of COVID-19 in Spain

from January 25, 2022, to February 23, 2022.

Figure 2 presents the results of the retrospective forecast of cumulative deaths of COVID-19 in Spain from January 25, 2022, to February 23, 2022.

Table 2 shows the model accuracy rates for cumulative new cases and deaths of COVID-19 in Spain from January 25, 2022 to February 23, 2022.

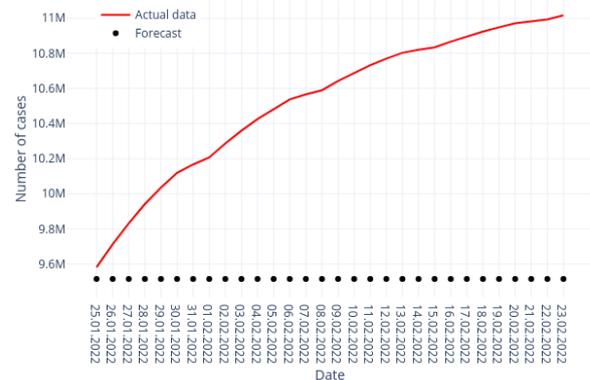


Fig. 1. Forecasting of COVID-19 cumulative new cases (25.01.22 – 23.02.2022)

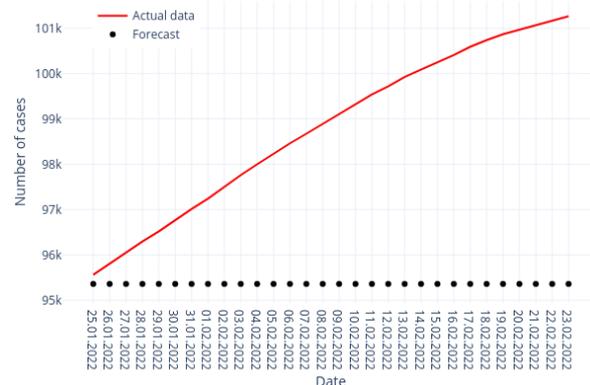


Fig. 2. Forecasting of COVID-19 cumulative death cases (25.01.22 – 23.02.2022)

Table 2

MAPE of cumulative forecast for 25.01.22 – 23.02.22

Duration of forecast (days)	New cases	Death cases
3 days (train)	0.007644 %	0.002714 %
3 days (test)	0.019822 %	0.004682 %
7 days (train)	0.016716 %	0.005538 %
7 days (test)	0.03972 %	0.009623 %
14 days (train)	0.029444 %	0.01026 %
14 days (test)	0.062766 %	0.018175 %
21 days (train)	0.041337 %	0.014896 %
21 days (test)	0.079319 %	0.026013 %
30 days (train)	0.238342 %	0.020345 %
30 days (test)	0.094509 %	0.034428 %

Figure 3 presents the results of the retrospective forecast of daily new cases of COVID-19 in Spain from January 25, 2022, to February 23, 2022.

Figure 4 presents the results of the retrospective forecast of daily death cases of COVID-19 in Spain from January 25, 2022, to February 23, 2022.

Table 3 shows model accuracy rates for daily new cases and deaths of COVID-19 in Spain from January 25, 2022 to February 23, 2022.

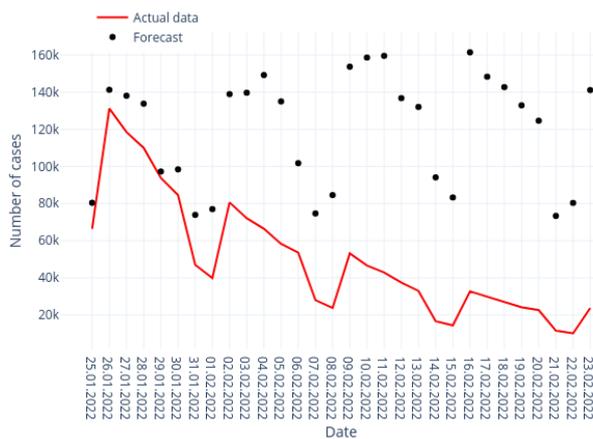


Fig. 3. Forecasting of COVID-19 daily new cases (25.01.22 – 23.02.2022)

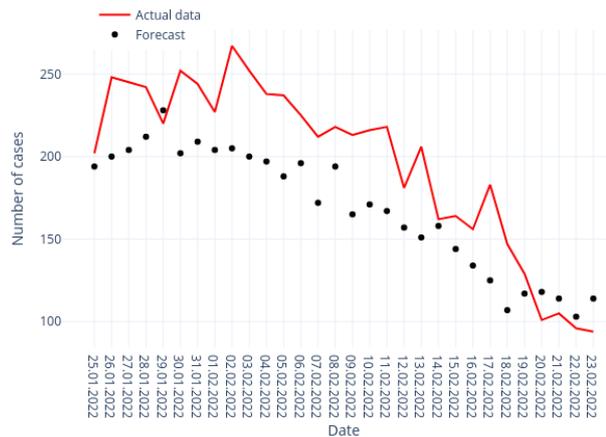


Fig. 4. Forecasting of COVID-19 daily death cases (25.01.22 – 23.02.2022)

Tuning and verifying the model on a sample from January 25, 2022, to February 23, 2022, showed high accuracy in forecasting the epidemic process of COVID-19 in Spain for both new and lethal cases.

### 4.2. Experimental Study

In order to assess the impact of a full-scale invasion, forecasts were made for the dynamics of COVID-19 epidemic process in Spain from February 24, 2022, to March 25, 2022, i.e., for the first 30 days of a full-scale Russian invasion of Ukraine.

Table 3

MAPE of daily forecast for 25.01.22 – 23.02.22

Duration of forecast (days)	New cases	Death cases
3 days (train)	0.36929 %	0.207202 %
3 days (test)	0.150603 %	0.133261 %
7 days (train)	0.393667 %	0.222863 %
7 days (test)	0.205718 %	0.129178 %
14 days (train)	0.495801 %	0.27154 %
14 days (test)	0.654244 %	0.153076 %
21 days (train)	0.573341 %	0.384405 %
21 days (test)	1.383668 %	0.159227 %
30 days (train)	0.67663 %	0.5238 %
30 days (test)	2.408559 %	0.161292 %

Figure 5 presents the results of the retrospective forecast of cumulative new cases of COVID-19 in Spain from February 24, 2022, to March 25, 2022.

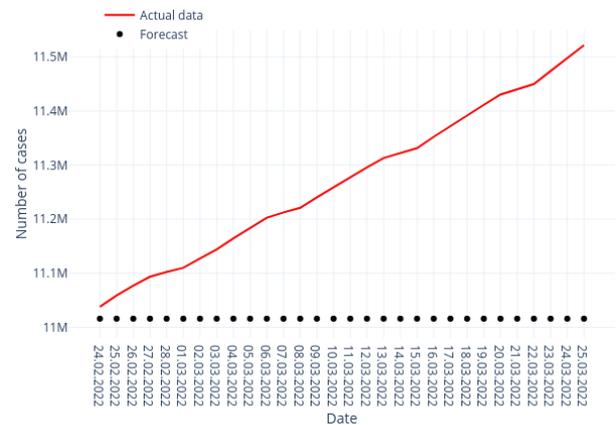


Fig. 5. Forecasting of COVID-19 cumulative new cases (24.02.22 – 25.03.2022)

Figure 6 presents the results of the retrospective forecast of cumulative death cases of COVID-19 in Spain from February 24, 2022, to March 25, 2022.

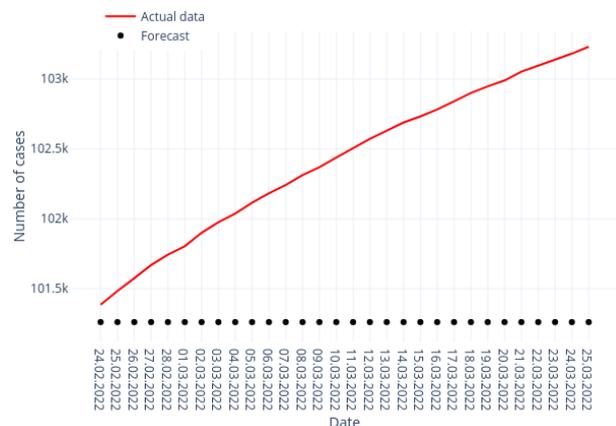


Fig. 6. Forecasting of COVID-19 cumulative death cases (24.02.22 – 25.03.2022)

Table 4 shows model accuracy rates for cumulative new cases and deaths of COVID-19 in Spain from February 24, 2022, to March 25, 2022.

Table 4

MAPE of cumulative forecast for 24.02.22 – 25.03.22

Duration of forecast (days)	New cases	Death cases
3 days (train)	0.007106 %	0.002408 %
3 days (test)	0.003804 %	0.002171 %
7 days (train)	0.015247 %	0.004909 %
7 days (test)	0.00638 %	0.00384 %
14 days (train)	0.0274 %	0.009304 %
14 days (test)	0.011209 %	0.0064 %
21 days (train)	0.038839 %	0.013469 %
21 days (test)	0.016065 %	0.008685 %
30 days (train)	0.053692 %	0.019581 %
30 days (test)	0.022448 %	0.011264 %

Figure 7 presents the results of the retrospective forecast of daily new cases of COVID-19 in Spain from February 24, 2022, to March 25, 2022.

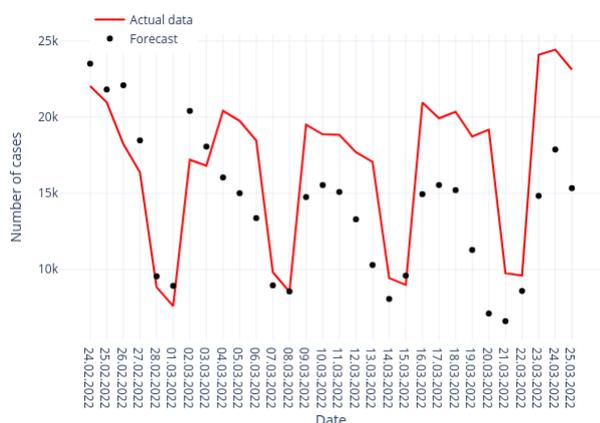


Fig. 7. Forecasting of COVID-19 daily new cases (24.02.22 – 25.03.2022)

Figure 8 presents the results of the retrospective forecast of daily death cases of COVID-19 in Spain from February 24, 2022, to March 25, 2022.

Table 5 shows model accuracy rates for daily new cases and deaths of COVID-19 in Spain from February 24, 2022, to March 25, 2022.

The results of assessing the accuracy of the retrospective forecast of the epidemic process of COVID-19 in Spain, built from February 24, 2022, to March 25, 2022, show the high accuracy of the model's performance. The accuracy of the forecast has not deteriorated compared to the forecast built from January 25, 2022, to March 25, 2022. This suggests that the forced migration of the Ukrainian population to Spain, caused by the full-scale invasion of Russia, is not a decisive factor affecting

the dynamics of the epidemic process of COVID-19 in Spain.

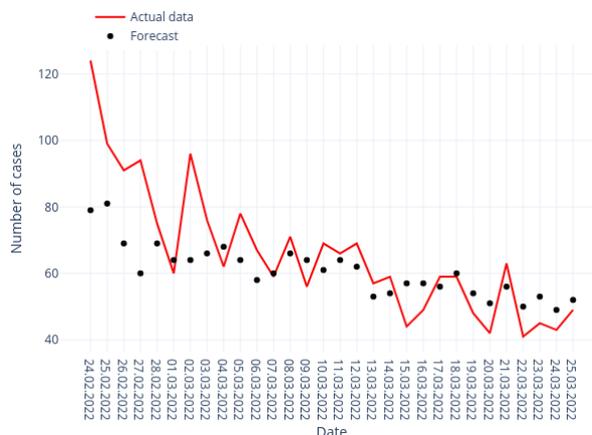


Fig. 8. Forecasting of COVID-19 daily death cases (24.02.22 – 25.03.2022)

Table 5

MAPE of daily forecast for 24.02.22 – 25.03.22

Duration of forecast (days)	New cases	Death cases
3 days (train)	0.255853 %	0.203772 %
3 days (test)	0.106256 %	0.260878 %
7 days (train)	0.295687 %	0.226502 %
7 days (test)	0.127042 %	0.229637 %
14 days (train)	0.491176 %	0.294949 %
14 days (test)	0.144826 %	0.169625 %
21 days (train)	0.615931 %	0.332027 %
21 days (test)	0.16897 %	0.153512 %
30 days (train)	1.100953 %	0.650314 %
30 days (test)	0.215537 %	0.144842 %

Figure 9 shows the deviation of daily reported new cases of COVID-19 in Spain from the forecasted values for the period from February 24, 2022, to March 25, 2022.

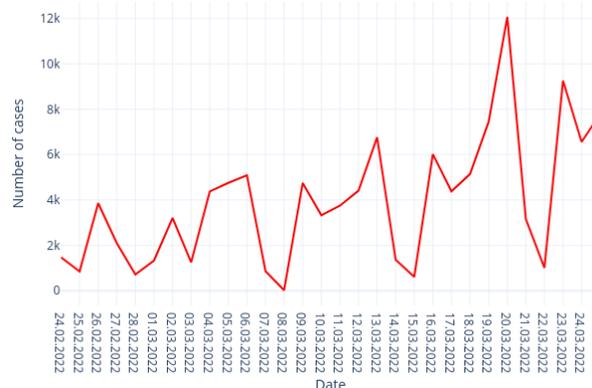


Fig. 9. Deviation of COVID-19 daily new cases from forecasted values (24.02.22 – 25.03.2022)

Figure 10 shows the deviation of daily reported death cases of COVID-19 in Spain from the forecasted values for the period from February 24, 2022, to March 25, 2022.

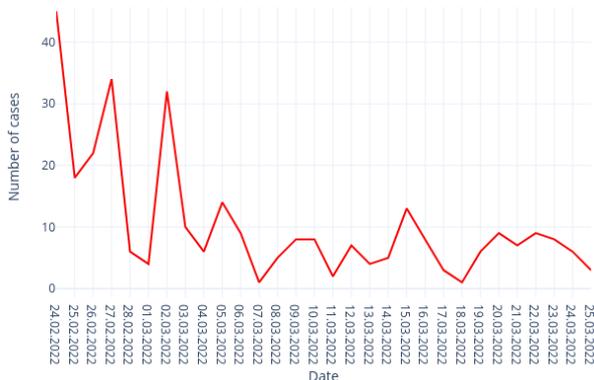


Fig. 10. Deviation of COVID-19 daily death cases from forecasted values (24.02.22 – 25.03.2022)

## 5. Discussion

Russia's treacherous attack on Ukraine on February 24, 2022, occurred during the height of COVID-19 pandemic. At this time, the peak incidence of COVID-19 was observed in Ukraine. On February 23, 2022, 25,062 cases were registered. At the same time, vaccination coverage against COVID-19 was low. 36.96% were vaccinated with two doses, and 38.24% with one dose. Children under 12 years of age were not subject to vaccination, and vaccination of children from 12 was allowed only from January 2022. Therefore, the epidemic situation for COVID-19 in the country was unfavorable.

The epidemic situation of COVID-19 in Spain was also tense. The epidemic process of COVID-19, which proceeded in waves at that time, was on the rise in incidence. The maximum incidence was registered in January 2022. On January 12, 2022, the number of cases was 179125. January 2022 was the month with the highest average cases in Spain. In February, the incidence was decreasing on the eve of the Russian-Ukrainian war. On February 23, the daily number of cases was 33,911.

Russia's aggression led to significant destruction and loss of life and caused massive migration to other regions of Ukraine and abroad. One of the countries where refugees from Ukraine recovered was Spain. According to the Spanish Ministry of the Interior, from March 2022 to mid-July 2022, 130,000 Ukrainian refugees found refuge in Spain [58].

With the beginning of military aggression, the life of Ukrainians changed a lot. In territories subjected to rocket attacks, artillery, and tank shelling, people were forced to hide in shelters, the subway, the basements of houses, and other premises. There was high population overcrowding, and distancing as a measure to contain the spread of the pathogen was not applicable. People have

stopped wearing face masks. In the first place was the desire to save one's life under shelling. Insufficient vaccination coverage contributed to the fact that there were many people susceptible to SARS-CoV-2 among people in shelters. All these factors contributed to the intensive circulation of coronavirus among people. People got infected and got sick. However, as a rule, only people with severe forms of the disease applied for medical help, and even then, not consistently. This was because the medical infrastructure was also affected by the war. Many hospitals were destroyed or damaged, and the number of medical personnel decreased due to the death, injury, or evacuation of hospital personnel. Also, in several localities, laboratories stopped working due to the evacuation of personnel and violations of logistics, which limited the possibility of obtaining test systems and testing cases. The population had to seek medical help by phone or the Internet, but the connection was unstable in territories under military attack or temporarily occupied. Thus, there were many sources of COVID-19 infection among the refugees. When leaving territories dangerous to life, people used overcrowded transport in trains heading to the west of the country and carrying several times the number of people exceeding the standards. There were no ventilation and primary sanitary and hygienic conditions. All this also contributed to the intensification of COVID-19 epidemic process.

When traveling to Spain, many COVID-19 restrictions were lifted for Ukrainian refugees. They were allowed to board flights to Spain without the obligatory medical uniform QR code (SpTH) for other passengers.

Therefore, it was expected that a new rise in the incidence of COVID-19 would begin in Spain.

However, this did not happen. At the beginning of the migration of refugees from Ukraine to Spain, there was a relatively high coverage of full vaccination of the country's population, which averaged 79.1%, with fluctuations in different territories. Ukrainian refugees were primarily sent to areas where vaccination coverage exceeded the national average. In Spain, strict restrictive measures were in place, people were required to wear masks in public places and transport, non-essential retail outlets were closed, and public events were not held.

Although air passengers from Ukraine were allowed not to have a medical form QR code (SpTH), vaccination certificate, or COVID-19 test results, such passengers were required to undergo a medical check for symptoms of COVID-19, such as fever, nasal congestion, cough, shortness of breath, and headache, etc., they underwent a rapid test for the virus antigen. They applied the Strategy for early detection, surveillance, and control of COVID-19 [59].

Thus, COVID-19 control and prevention measures – distancing, restrictive measures, strict mask regime,

high vaccination coverage of the Spanish population, and barriers to entry into the country were deterrents for circulating COVID-19 pathogen. Despite the high probability of the presence of sources of infection, Ukrainian refugees did not affect the worsening of the epidemic situation in Spain.

The results of modeling and retrospective forecasting of the epidemic process of COVID-19 in Spain confirm the epidemiological analysis and show that the forced migration of Ukrainian citizens caused by Russia's full-scale military invasion of Ukraine is not a critical factor influencing the dynamics of the incidence of COVID-19.

### Conclusions

This study is devoted to an experimental study of the impact of Russia's military invasion of Ukraine on the dynamics of COVID-19 epidemic process in Spain. For this, a methodology based on the ensemble method of machine learning XGBoost was proposed. Cumulative and daily morbidity and mortality data from COVID-19 in Spain obtained from the World Health Organization COVID-19 Dashboard were used for modeling. Data from January 25, 2022, to February 23, 2022, were used for testing. Data from February 24, 2022, to March 25, 2022, were used for the experimental study.

The scientific novelty of the article is the application of the ensemble machine learning method XGBoost to the modeling of emerging diseases using the example of COVID-19. The experimental results showed high accuracy of the model for various forecasting periods, including a long-term forecast for 30 days. The model was verified using cumulative and daily morbidity and mortality data.

The practical novelty of this study lies in assessing the impact of Russia's full-scale invasion of Ukraine on the dynamics of COVID-19 in Spain. The modeling results make it possible to substantiate the necessary control measures to curb the epidemic increase in the incidence of COVID-19. The proposed solution can be applied to other territories for which this task is relevant.

The developed model showed the high accuracy of forecasting new cases of COVID-19 in Spain for 30 days, which was 99.79%, and the death cases of COVID-19 in Spain – were 99.86%. The model was applied to data on the incidence of COVID-19 in Spain for the first 30 days of the war escalation (02/24/2022 – 03/25/2022). The calculated forecasted values showed that the forced migration of the Ukrainian population to Spain, caused by the full-scale invasion of Russia, is not a decisive factor affecting the dynamics of the epidemic process of COVID-19 in Spain.

**Contributions of authors:** conceptualization – **Dmytro Chumachenko**; methodology – **Dmytro Chumachenko**; formulation of tasks – **Dmytro Chumachenko, Tetyana Chumachenko**; analysis – **Dmytro Chumachenko, Tetiana Dudkina, Tetyana Chumachenko**; development of model – **Dmytro Chumachenko, Tetiana Dudkina**; software – **Tetiana Dudkina**; verification – **Dmytro Chumachenko, Tetiana Dudkina**; analysis of results – **Dmytro Chumachenko, Tetyana Chumachenko**; visualization – **Dmytro Chumachenko, Tetiana Dudkina**; writing – original draft preparation – **Dmytro Chumachenko**; writing – review and editing – **Tetiana Dudkina, Tetyana Chumachenko**.

**Acknowledgment.** The study was funded by the Ministry of Health of Ukraine in the framework of the research project 0123U100184 on the topic “Analysis of the impact of war and its consequences on the epidemic process of widespread infections on the basis of information technologies”.

### References

1. Hu, B., Guo, H., Zhou, P. and Shi, Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, 2020, vol. 19, iss. 19, pp. 1–14. DOI: 10.1038/s41579-020-00459-7.
2. Greco, M., Campagna, M., Cricelli, L., Grimaldi, M. and Strazzullo, S. COVID-19-related innovations: A study on underlying motivations and inter-organizational collaboration. *Industrial Marketing Management*, 2022, vol. 106, pp. 58–70. DOI: 10.1016/j.indmarman.2022.07.014.
3. Hossain, M., Abdulla, F. and Rahman, A. Challenges and difficulties faced in low- and middle-income countries during COVID-19. *Health Policy OPEN*, 2022, vol. 3, article no. 100082. DOI: 10.1016/j.hpopen.2022.100082.
4. Zhang, S.-X., Yang, M., Zheng, J.-X., Zhang, B.-Q., Pan, C.-H. and Tian, L.-G. Comprehensive strategies and measures to control COVID-19. *Infectious Diseases of Poverty*, 2022, vol. 11, iss. 1, article no. 67. DOI: 10.1186/s40249-022-00994-w.
5. Kim, S. COVID-19 Drug Development. *Journal of Microbiology and Biotechnology*, 2022, vol. 32, iss. 1, pp. 1-5. DOI: 10.4014/jmb.2110.10029.
6. Vitiello, A., Ferrara, F., Troiano, V. and La Porta, R. COVID-19 vaccines and decreased transmission of SARS-CoV-2. *Inflammopharmacology*, 2021, vol. 29, iss. 5, pp. 1357-1360. DOI: 10.1007/s10787-021-00847-2.
7. da Silva, J.G., Silva, C.S., Alexandre, B. and Morgado, P. Education as a Predictor Factor for

Knowledge of COVID-19 in Portugal. *Frontiers in Public Health*, 2021, vol. 9, article no. 680726. DOI: 10.3389/fpubh.2021.680726.

8. Meniailov, I., Krivtsov, S. and Chumachenko, T. Dimensionality Reduction of Diabetes Mellitus Patient Data Using the T-Distributed Stochastic Neighbor Embedding. *Lecture Notes in Network and Systems*, 2022, vol. 536, pp. 86-95. DOI: 10.1007/978-3-031-20141-7\_9.

9. Fedushko, S. and Ustyianovych, T. E-Commerce Customers Behavior Research Using Cohort Analysis: A Case Study of COVID-19. *Journal of Open Innovation: Technology, Market, and Complexity*, 2022, vol. 8, iss. 1, article no. 12. DOI: 10.3390/joitmc8010012.

10. Davidich, N., Chumachenko, I., Davidich, Y., Taisiia, H., Artsybasheva, N. and Tatiana, M. Advanced Traveller Information Systems to Optimizing Freight Driver Route Selection. *2020 13th International Conference on Developments in eSystems Engineering (DeSE)*, 2020, pp. 111-115. DOI: 10.1109/DeSE51703.2020.9450763.

11. Radutniy, R., Nechyporenko, A., Alekseeva, V., Titova, G., Bibik, D. and Gargin, V. V. Automated Measurement of Bone Thickness on SCT Sections and Other Images. *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*, 2020, pp. 222-226. DOI: 10.1109/DSMP47368.2020.9204289.

12. Bazilevych, K., Butkevych, M. and Padalko, H. Classification of Cardiovascular Disease Using AdaBoost Method. *Lecture Notes in Networks and Systems*, 2022, vol. 536, pp. 107-114. DOI: 10.1007/978-3-031-20141-7\_11.

13. Ju, X., Zhu, Y., Wang, Y., Li, J., Zhang, J., Gong, M., Ren, W., Li, S., Zhong, J., Zhang, L., Zhang, Q.C., Zhang, R. and Ding, Q. A novel cell culture system modeling the SARS-CoV-2 life cycle. *PLOS Pathogens*, 2021, vol. 17, iss. 3, article no. e1009439. DOI: 10.1371/journal.ppat.1009439.

14. Badi, H., Badi, I., Moutaouakil, K.E., Khamjane, A. and Bahri, A. Sentiment analysis and prediction of polarity vaccines based on Twitter data using deep NLP techniques. *Radioelectronic and Computer Systems*, 2022, no. 4, pp. 19-29. DOI: 10.32620/reks.2022.4.02.

15. Mohammadi, A., Meniailov, I., Bazilevych, K., Yakovlev, S. and Chumachenko, D. Comparative study of linear regression and SIR models of COVID-19 propagation in Ukraine before vaccination. *Radioelectronic and Computer Systems*, 2021, no. 3, pp. 5-18. DOI: 10.32620/reks.2021.3.01

16. Strilets, V., Donets, V., Ugryumov, M., Artiuch, S., Zelenskyi, R. and Goncharova, T. Agent-oriented data clustering for medical monitoring. *Radioelectronic and computer systems*, 2022, no. 1, pp.103-114. DOI: 10.32620/reks.2022.1.08.

17. Babaiev, V. M., Kadykova, I. M., Husieva, Yu. Yu. and Chumachenko, I. V. The method of adaptation of a project-oriented organization's strategy to exogenous changes. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 2017, no. 2, pp.134-140.

18. Ramírez, C. and Durón, R. M. The Russia-Ukraine war could bring catastrophic public-health challenges beyond COVID-19. *International Journal of Infectious Diseases*, 2022, vol. 120, pp. 44-45. DOI: 10.1016/j.ijid.2022.04.016.

19. *Ukraine Emergency: The United Nations Refugee Agency*, 2022. Available at: <https://www.unhcr.org/ukraine-emergency.html> (accessed at 10.01.2023).

20. *UNHCR. Operational Data Portal. Ukraine Refugee Station: The United Nations Refugee Agency*, 2022. Available at: <https://data.unhcr.org/en/situations/ukraine> (accessed: 10.01.2023).

21. McBryde, E. S., Meehan, M. T., Adegboye, O. A., Adekunle, A. I., Caldwell, J. M., Pak, A., Rojas, D. P., Williams, B. M. and Trauer, J. M. Role of modelling in COVID-19 policy development. *Paediatric Respiratory Reviews*, 2020, vol. 35, pp. 57-60. DOI: 10.1016/j.prrv.2020.06.013.

22. Martínez, V. A Modified SIRD Model to Study the Evolution of the COVID-19 Pandemic in Spain. *Symmetry*, 2021, vol. 13, iss. 4, article no. 723. DOI: 10.3390/sym13040723.

23. Kumar, P., Erturk, V. S. and Murillo-Arcila, M. A new fractional mathematical modelling of COVID-19 with the availability of vaccine. *Results in Physics*, 2021, vol. 24, article no. 104213. DOI: 10.1016/j.rinp.2021.104213.

24. Ghosh, S., Samanta, G. and Nieto, J. J. Application of non-parametric models for analyzing survival data of COVID-19 patients. *Journal of Infection and Public Health*, 2021, vol. 14, iss. 10, pp. 1328-1333. DOI: 10.1016/j.jiph.2021.08.025.

25. Morillas, F. and Valero, J. Random resampling numerical simulations applied to a SEIR compartmental model. *The European Physical Journal Plus*, 2021, vol. 136, iss. 10, article no. 1067. DOI: 10.1140/epjp/s13360-021-02003-9.

26. Matabuena, M., Rodríguez-Mier, P., García-Meixide, C. and Leborán, V. COVID-19: Estimation of the transmission dynamics in Spain using a stochastic simulator and black-box optimization techniques. *Computer Methods and Programs in Biomedicine*, 2021, vol. 211, article no. 106399. DOI: 10.1016/j.cmpb.2021.106399.

27. Murphy, C., Laurence, E. and Allard, A. Deep learning of contagion dynamics on complex networks. *Nature Communications*, 2021, vol. 12, iss. 1, article no. 4720. DOI: 10.1038/s41467-021-24732-2.

28. Guzmán-Merino, M., Durán, C., Marinescu, M.-C., Delgado-Sanz, C., Gomez-Barroso, D., Carretero, J. and Singh, D. E. Assessing population-sampling strategies for reducing the COVID-19 incidence. *Computers in Biology and Medicine*, 2021, vol. 139, article no. 104938. DOI: 10.1016/j.compbiomed.2021.104938.
29. Carballosa, A., Balsa-Barreiro, J., Garea, A., García-Selfa, D., Miramontes, Á. and Muñuzuri, A. P. Risk evaluation at municipality level of a COVID-19 outbreak incorporating relevant geographic data: the study case of Galicia. *Scientific Reports*, 2021, vol. 11, iss. 1, article no. 21248. DOI: 10.1038/s41598-021-00342-2.
30. Canga, A. and Bidegain, G. Modelling the Effect of the Interaction between Vaccination and Nonpharmaceutical Measures on COVID-19 Incidence. *Global Health*, 2022, vol. 2022, pp. 1-15. DOI: 10.1155/2022/9244953.
31. Merino, M. G., Marinescu, M. C. and Singh, D. E. Evaluating the spread of Omicron COVID-19 variant in Spain. *2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2022, pp. 999-1006. DOI: 10.1109/CCGrid54584.2022.00122.
32. Kumar, P., Erturk, V. S., Nisar, K. S., Jamshed, W. and Mohamed, M. S. Fractional dynamics of 2019-nCoV in Spain at different transmission rate with an idea of optimal control problem formulation. *Alexandria Engineering Journal*, 2022, vol. 61, iss. 3, pp. 2204–2219. DOI: 10.1016/j.aej.2021.07.005.
33. Garcia-Vicuña, D., Esparza, L. and Mallor, F. Hospital preparedness during epidemics using simulation: the case of COVID-19. *Central European Journal of Operations Research*, 2021, vol. 30, iss. 1, pp. 213–249. DOI: 10.1007/s10100-021-00779-w.
34. Carpio, A. and Pierret, E. Uncertainty quantification in Covid-19 spread: Lockdown effects. *Results in Physics*, 2022, vol. 35, article no. 105375. DOI: 10.1016/j.rinp.2022.105375.
35. Carballosa, A., Balsa-Barreiro, J. et al. Assessing the risk of pandemic outbreaks across municipalities with mathematical descriptors based on age and mobility restrictions. *Chaos, Solitons & Fractals*, 2022, vol. 160, article no. 112156. DOI: 10.1016/j.chaos.2022.112156.
36. Pájaro, M., Fajar, N. M., Alonso, A. A. and Otero-Muras, I. Stochastic SIR model predicts the evolution of COVID-19 epidemics from public health and wastewater data in small and medium-sized municipalities: A one year study. *Chaos, Solitons & Fractals*, 2022, vol. 164, article no. 112671. DOI: 10.1016/j.chaos.2022.112671.
37. Clement, J. C., Ponnusamy, V., Sriharipriya, K. C. and Nandakumar, R. A Survey on Mathematical, Machine Learning and Deep Learning Models for COVID-19 Transmission and Diagnosis. *IEEE Reviews in Biomedical Engineering*, 2022, vol. 15, pp. 325-340. DOI: 10.1109/RBME.2021.3069213.
38. *Worldometer. Worldometers – real time world statistics*, 2022. Available at: <https://www.worldometers.info/> (accessed: 10.01.2023).
39. The Lancet Public Health. COVID-19 in Spain: a predictable storm? *The Lancet Public Health*, 2020, vol. 5, iss. 11, article no. e568. DOI: 10.1016/S2468-2667(20)30239-5.
40. Mars, L., Arroyo, R. and Ruiz, T. Mobility and wellbeing during the covid-19 lockdown. evidence from spain. *Transportation Research Part A: Policy and Practice*, 2022, vol. 161, pp. 107-129. DOI: 10.1016/j.tra.2022.05.004.
41. Carriedo, Dr. A., Cecchini, P. Dr. J. A., Fernandez-Rio, Dr. J. and Méndez-Giménez, Dr. A. COVID-19, Psychological Well-being and Physical Activity Levels in Older Adults during the Nationwide Lockdown in Spain. *The American Journal of Geriatric Psychiatry*, 2020, vol. 28, iss. 11, pp. 1146-1155. DOI: 10.1016/j.jagp.2020.08.007.
42. Hoffmann, C. and Wolf, E. SARS-CoV-2 seroprevalence in Spain. *The Lancet*, 2020, vol. 396, iss. 10261, article no. 1484. DOI: 10.1016/S0140-6736(20)32272-8.
43. Martínez-de-Quel, Ó., Suárez-Iglesias, D., López-Flores, M. and Pérez, C.A. Physical Activity, Dietary Habits and Sleep Quality Before and During COVID-19 Lockdown: A Longitudinal Study. *Appetite*, 2020, vol. 158, article no. 105019. DOI: 10.1016/j.appet.2020.105019.
44. White, T.M., Cash-Gibson, L., Martin-Moreno, J. M., Matesanz, R., Crespo, J., Alfonso-Sanchez, J. L., Villapol, S., El-Mohandes, A. and Lazarus, J. V. COVID-SCORE Spain: public perceptions of key government COVID-19 control measures. *European Journal of Public Health*, 2021, vol. 31, iss. 5, pp. 1095-1102. DOI: 10.1093/eurpub/ckab066.
45. Bielecki, M., Patel, D., Hinkelbein, J., Komorowski, M., Kester, J., Ebrahim, S., Rodriguez-Morales, A.J., Memish, Z.A. and Schlagenhauf, P. Air travel and COVID-19 prevention in the pandemic and peri-pandemic period: a narrative review. *Travel Medicine and Infectious Disease*, 2020, vol. 39, article no. 101915. DOI: 10.1016/j.tmaid.2020.101915.
46. Roel, E., Raventós, B., Burn, E., Pistillo, A., Prieto-Alhambra, D. and Duarte-Salles, T. Socioeconomic Inequalities in COVID-19 Vaccination and Infection in Adults, Catalonia, Spain. *Emerging Infectious Diseases*, 2022, vol. 28, iss. 11, pp. 2243–2252. DOI: 10.3201/eid2811.220614.

47. Statista. *Ukraine War Casualties 2022*, 2022. Available at: <https://www.statista.com/statistics/1293492/ukraine-war-casualties/> (accessed: 10.01.2023).
48. *Ukraine: civilian casualty update 6 February 2023: Office of the high commissioner for human rights*, 2023. Available at: <https://www.ohchr.org/en/news/2023/02/ukraine-civilian-casualty-update-6-february-2023> (accessed: 10.02.2023).
49. The Lancet. Russia's invasion of Ukraine: an attack on health. *The Lancet*, 2023, vol. 401, iss. 10377, article no. 617. DOI: 10.1016/S0140-6736(23)00387-2.
50. *WHO Ukraine crisis response: January 2023 bulletin*, 2023. Available at: <https://www.who.int> (accessed: 10.02.2023).
51. Chumachenko, D., Pyrohov, P., Meniailov, I. and Chumachenko, T. Impact of war on COVID-19 pandemic in Ukraine: the simulation study. *Radioelectronic and computer systems*, 2022, vol. 2, pp. 6–23. DOI: 10.32620/reks.2022.2.01
52. *Ukraine. Internal Displacement Report. General Population Survey Round 12 (16 - 23 January 2023). Displacement Tracking Matrix: International Organization for Migration*, 2023. Available at: <https://dtm.iom.int> (accessed: 10.02.2023).
53. *Coronavirus Resource Center. Johns Hopkins University & Medicine*, 2020. Available at: <https://coronavirus.jhu.edu/> (accessed 10.01.2023).
54. *WHO COVID-19 Dashboard. Geneva: World Health Organization*, 2020. Available at: <https://covid19.who.int/> (accessed 10.01.2023).
55. Izonin, I., Tkachenko, R., Vitynskyi, P., Zub, K., Tkachenko, P. and Dronyuk, I. Stacking-based GRNN-SGTM Ensemble Model for Prediction Tasks. *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 326-330. DOI: 10.1109/DASA51403.2020.9317124.
56. Rahman, Md. S. and Chowdhury, A. H. A data-driven eXtreme gradient boosting machine learning model to predict COVID-19 transmission with meteorological drivers. *PLOS ONE*, 2022, vol. 17, iss. 9, article no. e0273319. DOI: 10.1371/journal.pone.0273319.
57. Chen, T. and Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 2016, pp. 785–794. DOI: 10.1145/2939672.2939785.
58. *Spain: Over 130 000 refugees from Ukraine granted temporary protection: European Website on Integration*, 2022. Available at: [https://ec.europa.eu/migrant-integration/news/spain-over-130-000-refugees-ukraine-granted-temporary-protection\\_en](https://ec.europa.eu/migrant-integration/news/spain-over-130-000-refugees-ukraine-granted-temporary-protection_en) (accessed: 10.01.2023).
59. *Strategy for early detection, surveillance and monitoring of Covid-19. Updated 1 December 2021 (Ministry of Health, Spain): Resources to support community and institutional Long-Term Care responses to COVID-19*, 2021. Available at: <https://ltccovid.org/resources/strategy-for-early-detection-surveillance-and-monitoring-of-covid-19-updated-1-december-2021-ministry-of-health-spain/> (accessed: 10.01.2023).

Надійшла до редакції 11.01.2023, розглянута на редколегії 20.02.2023

## ОЦІНКА ВПЛИВУ РОСІЙСЬКОЇ ВІЙНИ В УКРАЇНІ НА ПЕРЕДАЧУ COVID-19 В ІСПАНІЇ: ДОСЛІДЖЕННЯ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Дмитро Чумаченко, Тетяна Дудкіна,  
Тетяна Чумаченко

Пандемія COVID-19 справила значний вплив на світ, викликавши мільйони інфікувань та смертей, переваження систем охорони здоров'я, руйнування економіки та зміну повсякденного життя. Моделювання було визнано цінним інструментом у боротьбі з пандемією, що допомагає імітувати поширення вірусу, оцінювати вплив втручань та розробляти процеси прийняття рішень. Точність та ефективність моделювання залежать від якості вихідних даних, припущень та методів моделювання. Постійні зусилля щодо покращення та вдосконалення підходів до моделювання можуть підвищити їхню цінність при вирішенні майбутніх надзвичайних ситуацій у галузі охорони здоров'я. Повномасштабне військове вторгнення Росії в Україну 24 лютого 2022 р. призвело до серйозної гуманітарної кризи та кризи в галузі охорони здоров'я, що призвело до перебоїв у наданні медичних послуг, нестачі медикаментів та збільшення попиту на невідкладну допомогу. Війна, що триває, призвела до переміщення мільйонів людей, при цьому Іспанія посідає 5-е місце у світі за кількістю зареєстрованих біженців з України. **Метою** дослідження є оцінка впливу війни Росії в Україні на поширення COVID-19 в Іспанії з використанням засобів машинного навчання. **Об'єктом дослідження** є вивчення епідемічного процесу COVID-19 під час війни. **Предметом дослідження** є моделі та методи моделювання епідемічного процесу на основі машинного навчання. Для досягнення мети дослідження ми використали **методи**

прогнозування та побудували модель епідемічного процесу COVID-19 на основі методу XGBoost. В **результаті** експериментів точність прогнозування нових випадків COVID-19 в Іспанії на 30 днів становила 99,79 %, а летальних випадків від COVID-19 в Іспанії – 99,86 %. Модель застосована до даних про захворюваність на COVID-19 в Іспанії за перші 30 днів після початку ескалації війни (24.02.2022 – 25.03.2022). Розраховані прогнозні значення показали, що вимушена міграція населення України до Іспанії, спричинена повномасштабним вторгненням Росії, не є вирішальним фактором, що впливає на динаміку епідемічного процесу COVID-19 в Іспанії. **Висновки.** У статті описані результати експериментального дослідження щодо оцінки впливу повномасштабного вторгнення Росії в Україну на динаміку поширення COVID-19 в Іспанії. Розроблена модель показала хороші результати для використання у практиці охорони здоров'я. Аналіз отриманих результатів експериментального дослідження показав, що вимушена міграція населення України до Іспанії, спричинена повномасштабним вторгненням Росії, не є вирішальним фактором, що впливає на динаміку епідемічного процесу COVID-19 в Іспанії.

**Ключові слова:** епідемічна модель; епідемічний процес; моделювання епідемії; моделювання; COVID-19; XGBoost, війна.

**Чумаченко Дмитро Ігорович** – канд. техн. наук, доц., доц. каф. математичного моделювання та штучного інтелекту, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Дудкіна Тетяна Василівна** – здобувачка каф. математичного моделювання та штучного інтелекту, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

**Чумаченко Тетяна Олександрівна** – д-р мед. наук, проф., зав. каф. епідеміології, Харківський національний медичний університет, Харків, Україна.

**Dmytro Chumachenko** – PhD in Artificial Intelligence, Associate Professor of Department of Mathematical Modelling and Artificial Intelligence, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine, e-mail: dichumachenko@gmail.com, ORCID: 0000-0003-2623-3294.

**Tetiana Dudkina** – applicant of Department of Mathematical Modelling and Artificial Intelligence, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine, e-mail: dudkinatetiana@gmail.com, ORCID: 0000-0001-6309-2836.

**Tetyana Chumachenko** – Doctor of Medical Sciences, Professor, Head of Department of Epidemiology, Kharkiv National Medical University, Kharkiv, Ukraine, e-mail: tatalchum@gmail.com, ORCID: 0000-0002-4175-2941.