



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



COVID-19 contagion forecasting framework based on curve decomposition and evolutionary artificial neural networks: A case study in Andalusia, Spain[☆]

Miguel Díaz-Lozano^{a,*}, David Guijo-Rubio^b, Pedro Antonio Gutiérrez^b, Antonio Manuel Gómez-Orellana^b, Isaac Túñez^d, Luis Ortigosa-Moreno^a, Armando Romanos-Rodríguez^a, Javier Padillo-Ruiz^c, César Hervás-Martínez^b

^a Fundación Pública Andaluza Progreso y Salud, 41092 Sevilla, Spain

^b Department of Computer Science and Numerical Analysis, University of Cordoba, 14071 Cordoba, Spain

^c University of Sevilla. University Hospital Virgen del Rocío, 41013 Sevilla, Spain

^d Maimonides Institute for Biomedical Research of Córdoba (IMIBIC), 14071 Córdoba, Spain

ARTICLE INFO

Keywords:

COVID-19 contagion forecasting
Curve decomposition
Evolutionary artificial neural networks
Time series

ABSTRACT

Many types of research have been carried out with the aim of combating the COVID-19 pandemic since the first outbreak was detected in Wuhan, China. Anticipating the evolution of an outbreak helps to devise suitable economic, social and health care strategies to mitigate the effects of the virus. For this reason, predicting the SARS-CoV-2 transmission rate has become one of the most important and challenging problems of the past months. In this paper, we apply a two-stage mid and long-term forecasting framework to the epidemic situation in eight districts of Andalusia, Spain. First, an analytical procedure is performed iteratively to fit polynomial curves to the cumulative curve of contagions. Then, the extracted information is used for estimating the parameters and structure of an evolutionary artificial neural network with hybrid architectures (i.e., with different basis functions for the hidden nodes) while considering single and simultaneous time horizon estimations. The results obtained demonstrate that including polynomial information extracted during the training stage significantly improves the mid- and long-term estimations in seven of the eight considered districts. The increase in average accuracy (for the joint mid- and long-term horizon forecasts) is 37.61% and 35.53% when considering the single and simultaneous forecast approaches, respectively.

1. Introduction

Since the first outbreak detected in Wuhan, China, the new coronavirus has widely and rapidly spread around the world due to its powerful human-to-human transmission capacity (Sanche et al., 2020), leading to an exponential growth in the number of infected people in all countries. This new coronavirus, which produces the disease known as coronavirus disease 2019 (COVID-19), has put human health at risk

by provoking fever, cough, and myalgia as common symptoms, and potentially leading to complications, such as acute respiratory distress syndrome in a significant percentage (Chen et al., 2020; Huang et al., 2020), among others. On the 11th of March 2020, when many countries were in an emergency situation, the World Health Organization (WHO) declared this virus a global pandemic thereby, forcing the countries to adopt prevention measures such as nationwide lockdowns, mandatory

[☆] This work was supported by “Ministerio de Economía y Competitividad del Gobierno de España y Fondos FEDER” (grant reference: TIN2017-85887-C2-1-P); the “Agencia Española de Investigación (España)” (grant reference: PID2020-115454GB-C22/AEI/10.13039/501100011033); the “Consejería de Salud y Familia (Junta de Andalucía)” (grant reference: PS-2020-780); and the “Consejería de Transformación Económica, Industria, Conocimiento y Universidades (Junta de Andalucía) y Programa Operativo FEDER 2014-2020” (grant references: UCO-1261651 and PY20_00074). David Guijo-Rubio’s research has been subsidized by the University of Córdoba through grants to Public Universities for the requalification of the Spanish university system of the Ministry of Universities, financed by the European Union - NextGenerationEU (grant reference: UCOR01MS). Antonio M. Gómez-Orellana’s research has been supported by “Consejería de Transformación Económica, Industria, Conocimiento y Universidades de la Junta de Andalucía” (Grant Ref. PREDOC-00489).

* Corresponding author.

E-mail addresses: i42dilom@uco.es (M. Díaz-Lozano), dguijo@uco.es (D. Guijo-Rubio), pagutierrez@uco.es (P.A. Gutiérrez), am.gomez@uco.es (A.M. Gómez-Orellana), isaac.tunez@juntadeandalucia.es (I. Túñez), luis.ortigosa@juntadeandalucia.es (L. Ortigosa-Moreno), armando.romanos@juntadeandalucia.es (A. Romanos-Rodríguez), fpadillo@us.es (J. Padillo-Ruiz), chervas@uco.es (C. Hervás-Martínez).

<https://doi.org/10.1016/j.eswa.2022.117977>

Received 31 January 2022; Received in revised form 17 June 2022; Accepted 22 June 2022

Available online 27 June 2022

0957-4174/© 2022 Elsevier Ltd. All rights reserved.

use of facial masks and human mobility controls to suppress virus transmission. Many studies have shown that reducing social interaction by applying control measures helps to mitigate the spread of contagions (Kharroubi & Saleh, 2020; Kraemer et al., 2020).

In each country, the number of outbreaks, their intensity and duration depend on a wide variety of parameters, such as the size of population, the control measures applied and even specific climatological features (Livadiotis, 2020; Malki et al., 2020). Nevertheless, the evolution of the outbreaks are similar regardless of the location where they take place. The dynamics of the pandemic are as follows: at the beginning, a small number of people became infected. Then, due to the high contagion capacity, if social interaction is not limited, the number of infections increases linearly in a short period of time. After that, when a sufficient number of people contract the virus, infections begin to rise exponentially, and the outbreak quickly spirals out of control. At this time, control measures are usually applied to bring the outbreak under control. Hence, it is of huge importance to predict the contagion rate to avoid the inflexion point when the spread velocity starts to increase exponentially. Forecasting will allow the anticipation of uncontrolled situations and thus can help in adopting health system preparedness measures to avoid hospital overcrowding and in devising social restrictions to minimize the number of infections and the economic impact (Maital & Barzani, 2020).

To devise suitable economic, social and health care strategies to mitigate the pandemic effects, anticipating the evolution of the outbreaks has thus become a crucial task. Multiple forecasting techniques have been proposed since data about COVID-19 impact (regarding infected, deceased and recovered people) started to be available. In this connection, epidemiological models (Kermack & McKendrick, 1927) are the most popular approach to estimating the evolution of infectious diseases. These models use mutually exclusive compartments or states and assign them to a population of individuals to describe the dynamics of the population. Individuals flow through the compartments according to the parameters of the model. Compartmental models have been mainly applied at the beginning of the pandemic period in different locations. For example, in Anastassopoulou, Russo, Tsakris, and Siettos (2020), a four-compartmental SIRD (Susceptible–Infectious–Recovered–Dead) model was used to estimate the most important epidemiological parameters, such as the basic reproduction number, R_0 , and the rates of infection, mortality and recovery, using the COVID-19 incidence data of Hubei (China) from January 11 to February 10, 2020. Moreover, in Hauser et al. (2020), also using data from Hubei, an age-stratified SEIR (Susceptible–Exposed–Infected–Removed) model was fitted to estimate the symptomatic case-fatality ratio (sCFR) and infection-fatality ratio (IFR) over six regions of Europe. An extension of this model, including predictions of the number of cases using different time series forecasting techniques, was presented in Katris (2021).

Alternatively, machine learning (ML) techniques have also been applied to model some threatening aspects of the COVID-19 pandemic. ML approaches have a long history in solving real-world problems in different fields, including health care, economy and natural language processing. Concerning the health care area, and particularly the new SARS-CoV-2 virus, several ML and artificial intelligence expert systems have been proposed for different purposes. On the one hand, with the aim of identifying screening and management of SARS-CoV-2 positive diagnoses, methods have been developed to augment traditional identification tools, using radio imaging technology to detect abnormalities associated to COVID-19 infections (Huang et al., 2020; Ng et al., 2020) as an alternative to conventional tests, or even as standalone methods when viral testing is not an option. In Song et al. (2021), a convolutional neural network (CNN) was applied to distinguish between COVID-19 infected patients, pneumonia infected patients and healthy patients using computed tomography (CT) images. A deep neural network called COVID-Net is presented in Wang, Lin and Wong (2020) for identifying positive SARS-CoV-2 diagnoses using chest X ray

(CXR) images. In Abbas, Abdelsamea, and Gaber (2021) a class decomposition (Abbas, Abdelsamea, & Gaber, 2020) technique is applied in combination with a pretrained CNN, improving the performance of the classifiers when the class decomposition layer is included as a preprocessing step. A combination of deep learning and classical ML algorithms is proposed in Sethy and Behera (2020), where deep features of CXR images are extracted from the fully connected layer of a CNN and used to feed a support vector machine for distinguishing between healthy, COVID positive and pneumonia images. In Tamal et al. (2021), a set of radiomics features from CXR images were selected and used to train three classical ML classifiers, providing an accurate, fast and automatic method that can be integrated with standard X-ray reporting systems. Most recently, in Garg, Salehi, Rocca, Garner, and Duncan (2022), the performances of 20 different CNNs trained for classifying patients into three and two classes using chest CT images achieved an accurate and very efficient classification model.

Apart from the new imaging diagnostic assistance mechanisms, multiple algorithms and statistical techniques have been applied to obtain an accurate prognosis of the pandemic rates, the peak of outbreaks in different countries or even estimations of specific pandemic wave scopes and to create analytical models that could act as decision support systems. In this sense, in Benvenuto, Giovanetti, Vassallo, Angeletti, and Ciccozzi (2020), a classical autoregressive integrated moving average (ARIMA) is applied to analyze the trend of COVID-19 prevalence and incidence and to perform short-term prediction about the spread of the virus. In Wang, Zheng, Li and Zhu (2020), three critical points of infections and recovered cases are estimated in different countries by considering a hybrid forecast model based on a logistic curve fit and Prophet (Taylor & Letham, 2018) application. A comparison of six time series forecasting techniques applied to active COVID-19 cases is performed in Papastefanopoulos, Linardatos, and Kotsiantis (2020). Following a similar methodology, in Ribeiro, da Silva, Mariani, and dos Santos Coelho (2020), the predictive capacity of different machine learning regression models is measured using data from 10 Brazilian states. More complex models are used in Verma, Mandal, and Gupta (2022) for forecasting purposes, where long short-term memory (LSTM) recurrent neural networks are designed for predicting the contagion rate in 4 Indian states. Multiple regressors were applied in An et al. (2020) to predict the risk of mortality using different characteristics of an infected person, concluding that variables such as advanced age or taking metformin are important predictors that influence the output probability. An assisted fuzzy case-based reasoning (FCBR) algorithm for determining patient attention priorities based on eight factors is presented in Geetha, Narayanamoorthy, Manirathinam, and Kang (2021) with the aim of improving medical assistance and reducing the mortality of COVID-19 patients. In Fidan and Yuksel (2022), an unsupervised study is carried out to analyze the importance of city-related parameters, such as population and environmental variables, in addition to the number of cases to establish restrictions to contain the spread of the virus. In Desai (2021), the effects of news sentiments are included in the forecast task, concluding that negative news sentiments could help to reduce the contagion rate. Finally, in Khan et al. (2021), a very comprehensive review of research regarding the use of diverse ML algorithms to combat the COVID-19 pandemic is presented, where applications such as diagnosis, detection or forecasting are carried out using different types of data in several countries.

Among all the state-of-the-art techniques in ML, there is one option that stands out: feed-forward neural networks (Bishop et al., 1995). This is a highly adaptive technique that can be used to model most nonlinear problems, mainly due to its universal approximation capability (Hornik, Stinchcombe, & White, 1989). These models are layer-divided structures, where each layer is composed of computational units that are connected to the nodes of the next layer by weighted connections. Artificial neural networks (ANNs) have been trained to solve multiple classification and regression problems (Abiodun et al., 2018). In this sense, backpropagation is the most popular

algorithm for adjusting the network parameters in the training stage. This method computes the gradient of the loss function with respect to the weights of the network connections. However, even though this approach can efficiently adjust the weight magnitudes, a global optimal solution is not guaranteed. The reason behind this is the large extent of the multidimensional error surface and the presence of local optima where the backpropagation can get stuck. To perform a more exhaustive search over this surface, evolutionary algorithms have been presented as an alternative to discover and explore different regions (trying to escape from local optima) where more accurate solutions could be located. For forecasting purposes, evolutionary artificial neural networks (EANNs) have been applied in different fields such as renewable energy (Gómez-Orellana, Guijo-Rubio, Gutiérrez, & Hervás-Martínez, 2022; Mason, Duggan, & Howley, 2018), economy (Au, Choi, & Yu, 2008; Chiroma, Abdulkareem, & Herawan, 2015; Yi-Hui, 2007) or environment (Lu, Fan, & Lo, 2003). When modeling a future value as a target variable, it is important to consider the existing trade-off between accuracy and anticipation. Short-term events are usually easier to model but do not offer enough anticipation time, while long-term horizon forecasts are very useful for anticipating important events, but their accuracy decreases when the forecasting horizon increases. It is worth mentioning that the concepts of short and long horizons are relative and depend on the field of study where the prediction is carried out. Thus, in Gómez-Orellana et al. (2022), a 48-h forecast is considered a long-term horizon, while in Au et al. (2008) they refer to weekly events as short-term events. In the field of epidemiology, since the parameters associated with an outbreak evolve rapidly and the information is reported daily, 1, 3, and 5 days can be considered to be short-, mid- and long-term horizons, respectively.

In this study, we make the following contributions. First, we propose a novel methodology for the analysis of the contagion curve by extracting evolution characteristics. The extraction procedure is performed through an iterative process of polynomial model fitting. The coefficients of the polynomial model fitting the cumulative contagion curve describe the shape and evolution of the curve, where points represent days since the beginning of each wave. This procedure allows the collection of inherent features of a process that can be used as related information, which is especially useful when process exogenous related information is scarce. Second, we apply this methodology to real data involving the number of infected people in different locations of Andalusia (Spain) to build a transfer function per zone that can predict the evolution of the contagion rate on the different stages of the pandemic. By means of this methodology, we build a multivariate time series dataset per sanitary district by considering the real cumulative contagion curve data as the dependent series and the polynomial coefficient series as independent correlated series. Third, we estimate the parameters of these transfer functions using EANNs, with data from different periods of the pandemic waves that took place in Andalusia. Given that every district presents different contagion rates in the considered outbreaks, different architectures of EANNs are considered using distinct basis functions in the nodes belonging to the hidden layer with the goal of finding the best network scheme for each district. By using distinct periods for the estimation of the parameters, we build a forecast model per sanitary district using a different order of lags that serves to obtain an accurate prediction about the contagion rate in different stages of an outbreak. This results in models that can identify the beginning and ending of new outbreaks and predict the evolution of the contagion rate during the most critical phase. The parameter estimation is carried out by considering mid- and long-term time forecast horizons (3 and 5 days, respectively). The reason behind the choice of these time horizons is that they offer an acceptable trade-off between anticipation and accuracy. To serve as decision support for adopting prevention measures, 3- and 5-day predictions of the number of contagions are enough to devise suitable containment schemes. Larger horizons lead to less accurate results. It is not appropriate to base sensitive decisions such as the application

of restriction measures on such results. Additionally, we consider a combined estimation where both time forecast horizons are involved in a single model using a multitask evolutionary artificial neural networks (MuEANNs) approach. The resulting models can act as a decision support system in new emergency situations caused by new viruses or strains with similar contagion rate behavior. The results obtained with the models that are trained with the extracted polynomial information significantly outperform those obtained with models trained with pure autoregressive data in almost all districts. Specially, the accuracy is increased by up to 73.55% in the case of the Almería district.

The remainder of this paper is organized as follows: in Section 2, we describe the data source and perform an analysis of the data used in this study. The proposed methodology is explained in Section 3 where the polynomial curve fit, the dataset building process, the autoregressive forecast models and the EANNs are presented. The experimental design and results obtained for each architecture, dataset and forecasting approach are presented in Section 4. Finally, we conclude the paper in Section 5.

2. Data description

The data used in this study and some considerations are specified in this section. This article focuses on eight areas of Andalusia, the country's second largest and most populated autonomous region, located in southern Spain. The data were obtained from the official website of the Andalusia government,¹ where the number of COVID-19 diagnosed, cured and deceased people were reported daily and segregated into 34 sanitary districts. Sanitary districts are administrative divisions that have local health management competencies in specific zones of Andalusia. These divisions are represented in Fig. 1, where the eight districts that include the provincial capitals are highlighted in green and their populations are specified. For this study, the daily reported information in these eight sanitary districts from July 11, 2020, to August 30, 2021, is used, resulting in a total of 417 observations, each representing one day of an outbreak. During this period of time, four different waves took place in Andalusia. Although the data reported by the Andalusia government about positive diagnosed contagions are available since March 30, 2020, the data until July 11, 2020 correspond to the first wave of the pandemic. During this wave, the country wide lockdown resulted in a low intensity and short time wave. Hence we decided to exclude this period from the analysis.

The reported information presents a weekly pattern of fewer diagnosed contagions during weekends, making the positive diagnoses time series highly noisy. This effect is common to most countries (Ricon-Becker, Tarrasch, Blinder, & Ben-Eliyahu, 2020) and is due to the lower number of patients tested on Saturdays and Sundays, differences in testing timings and reporting delays. Missing data on weekends are usually included in the first days of the weekly reports thereby preserving the real number of positive diagnoses. Figs. 2(a) and 2(b) show, respectively, the daily and cumulative positive COVID-19 diagnosis time series belonging to the district of Córdoba, one of the districts considered in this study, from July 11, 2020 to August 30, 2021.

In Fig. 2(a), the noise produced by the weekly cycles is observable, resulting in a visible sawtooth effect in the contagion time series. The representation of the cumulative case time series is shown in Fig. 2(b) where the noise is considerably mitigated, leading to a soft and monotonically nondecreasing time series. The temporal trend of this district is repeated for the other 33 districts, while varying in magnitude according to specific district parameters, such as the geographic location or the total population. The temporal positive diagnoses returned four different waves, each one defined by an increment and followed by a drop in the contagion rate.

¹ The number of infections in Andalusia are reported daily in <https://www.juntadeandalucia.es/institutodeestadisticaycartografia>.

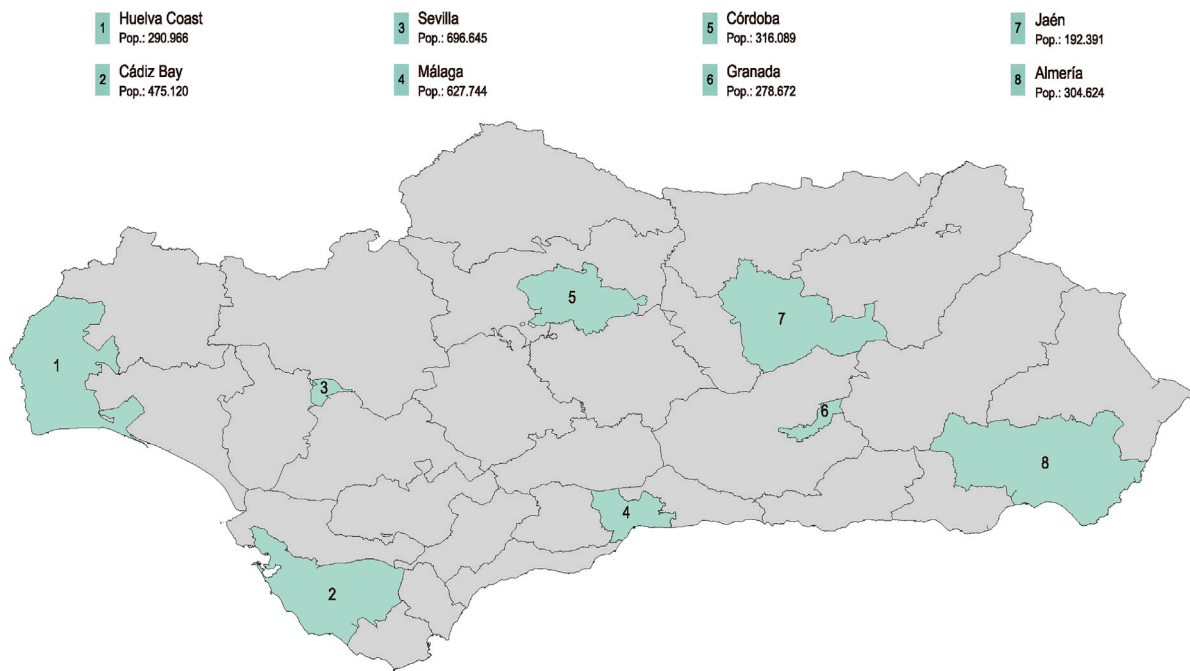
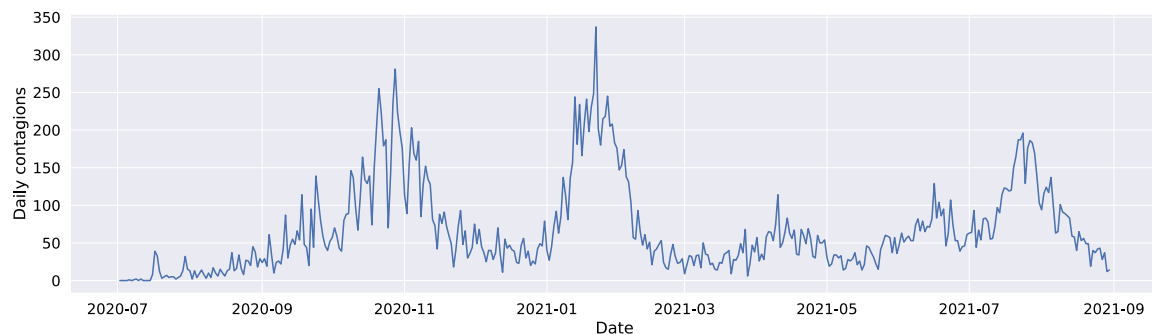
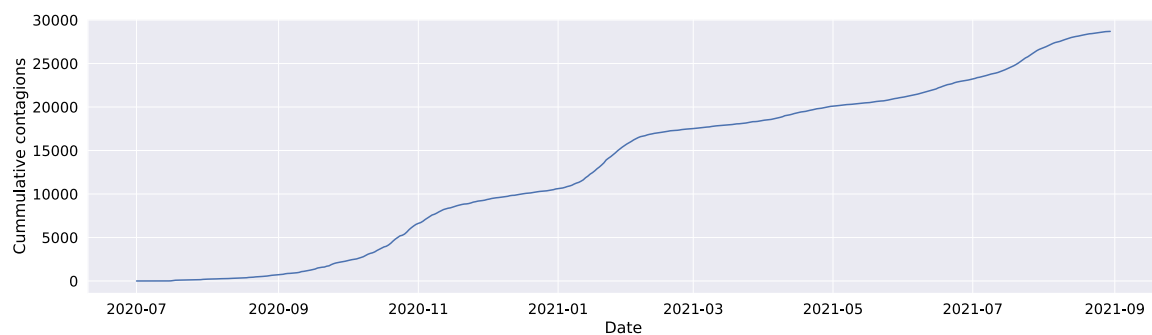


Fig. 1. Geography of the 34 sanitary districts of Andalusia. The 8 districts including provincial capitals are highlighted in green.



(a) Daily contagions in district of Córdoba



(b) Cumulative contagions in district of Córdoba

Fig. 2. Daily (a) and cumulative (b) reported positive COVID-19 diagnoses in the district of Córdoba from July 11, 2020, to August 30, 2021.

As mentioned above, the first wave has not been considered due to its low intensity and short wave time. Therefore, for this analysis, the following waves have been used: the second wave spans from July 11, 2020, to December 15, 2020, and its long-term evolution is characterized by a high peak of positive contagions and a gradual increment and drop in the contagion rate. During this wave, the average age of the people needing hospital care decreased compared to the first wave and the mean hospitalization time was also reduced (Iftimie et al., 2021). The

third wave spans from December 15, 2020, to March 13, 2021. The beginning of the contagion rate of this wave is highly accelerated because of the removal of control measures due to the Christmas holidays. During this period, mobility and social interaction increased, resulting in a quick increment in the number of infected people and the maximum number of COVID-19 cases was reached. At the end of the holiday period, when control measures were reset, the contagion rate slowed down. The fourth wave, lasting from March 13, 2021, to June 20, 2021, is defined by a soft initial diagnoses increment owing to the beginning

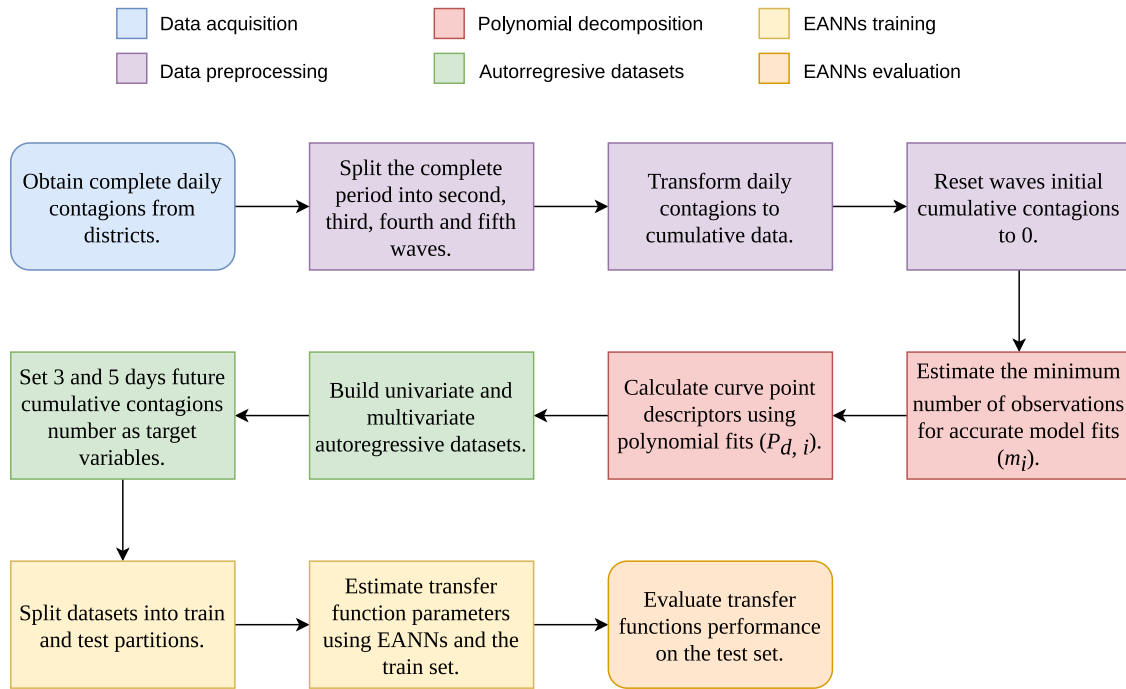


Fig. 3. Overview flowchart describing the applied methodology.

of the vaccination campaign. This wave presents two peaks provoked by the different changes in the containment measures policies. Finally, the fifth wave took place in summer, from June 20, 2021, to August 30, 2021, a period where control policies were loose. However, the high vaccination rate produced a wave with weak magnitude and short time period. Considering these partitions, the second, third, fourth and fifth waves consist of 157, 88, 99 and 71 observations, respectively. These dates have been chosen based on expert knowledge and the behavior observed in the daily contagion curves of the considered districts.

The nature of these four waves is very different in terms of evolution because of the distinct social policies, control measures and vaccination campaigns that took place along their respective developments, resulting in a heterogeneous set of rates for each wave. For this reason, with the main objective of modeling the complete period of the pandemic, including the most diverse information possible about all stages is crucial. In this study, 3 different periods for each wave are considered: beginning, growth and stabilization. Including information about these three stages in the parameter optimization phase results in models that can detect imminent increases in contagion rates that can potentially lead to local outbreaks, accurately predict the evolution of an initiated outbreak and recognize in advance the end of virus transmission.

3. Proposed methodology

In this section, the methodology proposed in this paper and the process to generate the datasets used in the experimentation are detailed. To provide an overview and facilitate understanding, Fig. 3 presents a flowchart containing all the steps detailed in this section.

3.1. Curve fitting

Many empirical data obtained from natural processes describe patterns that may be defined using algebraic expressions. Curve fitting techniques help to understand the inherent behavior of a process and may be useful in description, clustering or forecasting tasks (Motulsky & Ransnas, 1987). If the nature of the underlying process is known, an accurate parametric function working as an approximator results in a soft curve whose parameters constitute a characterization of the empirical data evolution. The information extracted from the fitted curve can

be used in ML tasks as important features, both in supervised (Hamidi, Ghasseman, & Imani, 2018) and unsupervised (Abraham, Cornillon, Matzner-Lber, & Molinari, 2003; Martin et al., 1998) contexts.

In this paper, we consider the approximation by a polynomial function of the cumulative curve of contagions of the 8 districts involved in this study by the least squares method. The shape of the curve described in each wave is defined by a soft initial growth. Then, when a sufficiently high number of infected people is reached, an exponential increment occurs due to the high contagion rate. The waves finish with a linear evolution when the outbreak is under control, but residual positive cases are still present. According to this shape, a third-order polynomial is the most interpretable and lowest-variance model that can accurately fit the cumulative curve of each wave individually. The polynomial regression is a type of fit where the dependent variable y_d linearly depends on the powers of a single independent variable, in this case, the number of days from the beginning of the wave. A k -degree polynomial model, which is composed of $k + 1$ parameters, is defined as:

$$y_d = \alpha_0 + \alpha_1 d + \alpha_2 d^2 + \alpha_3 d^3 + \dots + \alpha_k d^k + e_d, \quad (1)$$

where y_d is the cumulative number of contagions after d days from the beginning of an outbreak, $\alpha^k = \{\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_k\}$ are the $k + 1$ parameters to be estimated and e_d is the error term. For this study, the chosen 3-degree polynomial model is composed of 4 parameters to be estimated during fit and has a single inflexion point along with its domain, where the function changes its concavity.

Following an iterative procedure, for each time period between the start of the outbreak i (Day 0) and a day d , a 3-degree polynomial model is fitted. Once the polynomial model has been fitted, the estimated parameters are used to describe a point of the curve, which is represented as: $P_{d,i} = \{\alpha_{d,i}, y_{d,i}\}$, where $\alpha_{d,i} = \{\alpha_{1,d,i}, \alpha_{2,d,i}, \alpha_{3,d,i}, \alpha_{4,d,i}\}$ includes the four parameters of the third order polynomial. Note that each point of the curve represents a single day belonging to the i th outbreak, where $i \in \{2, 3, 4, 5\}$. Moreover, it is worth mentioning that the sanitary districts are considered independently. A more detailed description of the variables representing each point $P_{d,i}$ is given in Table 1.

Given that for building accurate models a minimum number of days is needed, d can only take values in the range $[m_i, D_i]$. In this way, m_i is

Table 1

Descriptors of a point of the cumulative curve of contagions.

Variable	Description
$\alpha_{1,d,i}$	Estimated α_1 after d days of the i th wave
$\alpha_{2,d,i}$	Estimated α_2 after d days of the i th wave
$\alpha_{3,d,i}$	Estimated α_3 after d days of the i th wave
$\alpha_{4,d,i}$	Estimated α_4 after d days of the i th wave
$y_{d,i}$	Cum. contagions after d days of the i th wave

Table 2Mean, and standard deviation, (*SD*), of the R^2 statistics for the fitted polynomial models.

District	R^2	
	Mean	SD
Almería	0.9954	0.0043
Cádiz Bay	0.9961	0.0023
Córdoba	0.9944	0.0075
Granada	0.9947	0.0047
Huelva Coast	0.9932	0.0044
Jaén	0.9929	0.0166
Málaga	0.9959	0.0029
Sevilla	0.9967	0.0028

the minimum number of observations needed to build accurate models, and D_i is the total duration, in days, of the i th outbreak. Therefore, the number of polynomial models to be fitted for outbreak i is $(D_i - m_i + 1)$.

For the estimation of m_i , the evolution of the quality of the fitted polynomial models has been measured for the 8 sanitary districts under study, using the coefficient of determination R^2 as an evaluation metric, defined in Eq. (2) for the i th wave. R^2 statistic represents the proportion of the variance of the dependent variable explained by the independent variables:

$$R_i^2 = 1 - \frac{\sum_{d=1}^{D_i} (y_{d,i} - \hat{y}_{d,i})^2}{\sum_{d=1}^{D_i} (y_{d,i} - \bar{y}_i)^2}, \quad (2)$$

where $y_{d,i}$, a real number, is the cumulative contagions after d days, $\hat{y}_{d,i}$ is the number of cumulative contagions predicted by the model, and \bar{y}_i is the mean of observed cumulative contagions. This metric ranges from 0 to 1, reaching its maximum value when the model is able to entirely explain the variance present in the real data, e.g., fitting perfectly to the empirical nature of the observed contagion rate.

Analyzing the evolution of the mean R^2 magnitude of the fitted polynomial models of the eight districts, the minimum feasible m_i choice for each outbreak is based on the instant from which it begins to stabilize. At the beginning of the waves, the cumulative number of contagions is reset to 0, considering the waves individually, e.g., ignoring the cumulative number of infected people of past waves. Fig. 4 shows the average, maximum and minimum evolution of the coefficient of determination for the eight sanitary districts, and the chosen m_i is indicated with a vertical line. The minimum number of days m_i where the determination coefficient variance begins to stabilize corresponds to 21, 13, 12 and 15 for the second, third, fourth and fifth waves, respectively. Considering these initial thresholds for fitting polynomial curves, from the 415 observations of each district, a total of 61 days are used as the minimum number of observations to build accurate regressions, resulting in a total of 358 polynomial fits on the entire pandemic period for each district. In Table 2, the Mean R^2 , and standard deviation, *SD*, of all approximated models for each district are presented. As shown in Table 2, the selected polynomial models fit analytically well for all districts and waves.

3.2. Autoregressive models

Formally, a time series $T = (v_1, v_2, \dots, v_l)$ of length l is an ordered sequence of values v_j . The timestamps $j = \{1, 2, \dots, l\}$ conform to a sequence of positive and ascending discrete values, which, in most

cases, are equally spaced. Time series are produced by any kind of sequenced phenomenon, whether natural, such as the climate in a particular location, or artificial, such as stock prices of a given enterprise over time. Due to its intrinsic temporal nature, time series have been the object of investigation in multiple fields of data mining, such as forecasting, classification, clustering, and outlier analysis (Fu, 2011). A time series may be composed of more than one time-dependent variable measured at the same time instant. In this case, time series are called multivariate, and its analysis involves modeling the inherent relationship between the dimensions that compose it. Mathematically, a multivariate time series $T = (v_1, v_2, \dots, v_l)$ of length l , is composed of N dimensions; thus, the value is a vector $v_j^T = (v_{1,j}, v_{2,j}, \dots, v_{N,j})$ of length N . In this sense, the point descriptors, obtained from the polynomial fit iterative process described in Section 3.1 and detailed in Table 1, are used to build a multivariate time series dataset per district.

Within the scope of time series forecasting, autoregressive (AR) models are a useful technique for modeling future values by using the information of the p past instants. A p -order AR model, denoted as $AR(p)$, can be applied to the target variable as follows:

$$v_d = c + \sum_{j=1}^p \gamma_j v_{d-j} + e_d, \quad (3)$$

where v_d is the target value obtained from past lagged values (v_{d-j} for $j = \{1, 2, \dots, p\}$), c is a constant that serves as the intercept of the model, γ_j is the j th coefficient of the model, and e_d is the error term. AR models represent any value of the series as a combination of its p past values. Specifically, in this study, AR models will be used for the prediction of the number of cumulative contagions of the i th outbreak of d ; hence, $v_d = y_{d,i}$ (note that d is the number of days since the beginning of the i th outbreak). Furthermore, when the forecasting of the number of cumulative contagions is carried out by using multivariate time series, a vector autoregressive model (VAR) (Zivot & Wang, 2003) is used. These VAR models generalize the AR model by including information of the different dimensions of the time series (independent terms) to model the target time series (dependent term). $VAR(p)$ models include p lags of all the dimensions of a given time series, in our case, lags of the point descriptors presented in Table 1. Consequently, the $VAR(p)$ model composed of N dimensions is defined as follows:

$$v_d = c + \sum_{j=1}^p \gamma_j v_{d-j} + e_d, \quad (4)$$

where γ_j is the vector of model coefficients and v_{d-j} is the vector containing the j th lagged representation of the curve point belonging to the i th outbreak, i.e., the j th lagged values of the variables described in Table 1. Note that the v vector has length N . In the same way that happens with AR models, VAR models will be used for the prediction of the number of cumulative contagions of the i th outbreak of d ; hence, $v_d = y_{d,i}$.

3.3. Artificial neural networks

Artificial neural network (ANN) models (Bishop et al., 1995) have been applied to a large number of regression and classification tasks, demonstrating great performance due to their capability of being universal approximators (Hornik et al., 1989). In this sense, multiple ANN models have been proposed in recent years: one of the first approaches was the multilayer perceptron (MLP) (Bishop et al., 1995) with sigmoid units (SUs) as basis functions for the nodes in the hidden layers. Other alternatives include the radial basis function (RBF) (Broomhead & Lowe, 1988; Poggio & Girosi, 1989) neural networks, which usually make use of Gaussian transfer functions in the nodes of the hidden layers, and multiplicative networks, which use product units (PUs) (Durbin & Rumelhart, 1989), computing a weighted product instead of a weighted sum for a node input.

In this paper, given that mid- and long-term forecasts are considered, the use of single and simultaneous forecasts is proposed. In

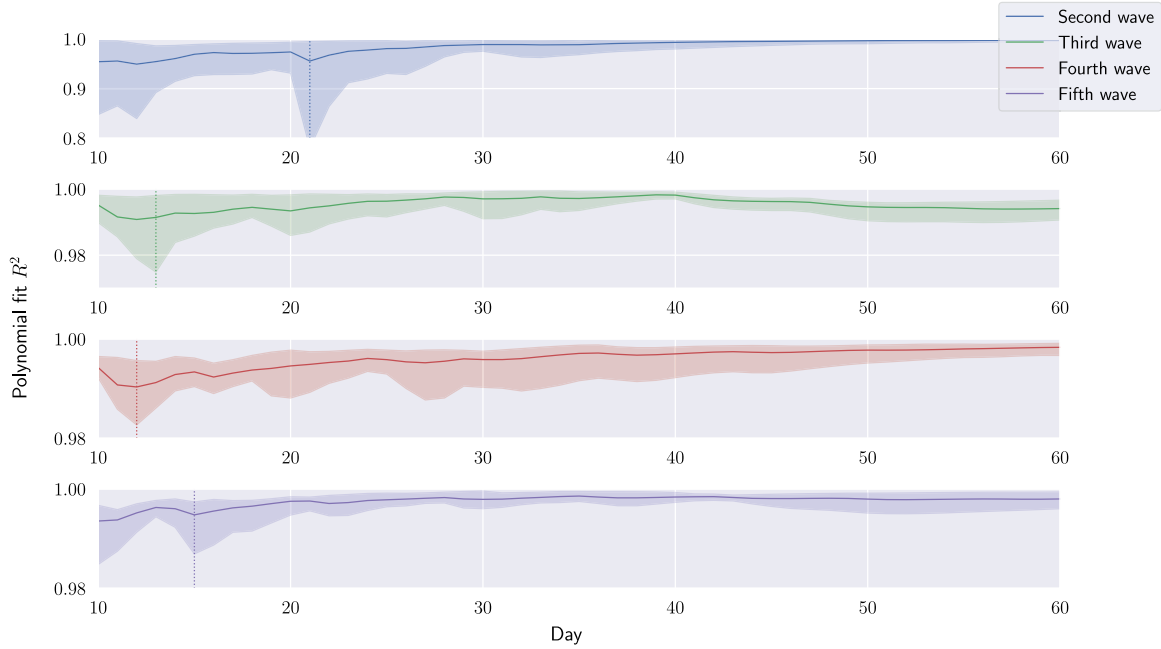


Fig. 4. Mean, maximum and minimum R^2 evolution of the eight sanitary districts for each wave.

this sense, two different models of ANNs are developed. On the one hand, ANNs aim to perform each forecast task separately (monotask), and on the other hand, ANNs achieve both forecast tasks simultaneously (multitask). Accordingly, monotask models only have one output, whereas multitask models have two outputs. Considering that the aim of this work is to forecast the cumulative number of contagions, two different horizons have been considered: $y_{d+3,i}$ and $y_{d+5,i}$. These two forecast horizons are focused on predicting the cumulative number of contagions at mid- and long-term, i.e., 3 and 5 days ahead, respectively.

Therefore, the ANNs models applied to perform the forecast tasks can be defined as follows:

$$f_q(\mathbf{x}, \boldsymbol{\theta}) = \beta_{q0} + \sum_{j=1}^m \beta_{qj} B_j(\mathbf{x}, \mathbf{w}_j), \quad q = \{1, 2\}, \quad (5)$$

where f_q is the model output for the forecast task considering the output q , $\boldsymbol{\theta} = \{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_q, \mathbf{w}_1^1, \dots, \mathbf{w}_{m_1}^1, \mathbf{w}_1^2, \dots, \mathbf{w}_{m_2}^2\}$ contains the coefficients of the ANN model, $\boldsymbol{\beta}_q = \{\beta_{q0}, \beta_{q1}, \dots, \beta_{qm}\}$ including the weights from the hidden layer to the q th output node, and $\mathbf{w}_j = \{w_{j0}, w_{j1}, \dots, w_{jn}\}$, containing the weights of the connections from the n input nodes to the j th hidden node, m being the number of neurons in the hidden layer. $B_j(\mathbf{x}, \mathbf{w}_j)$ represents the basis function of the j th hidden neuron, \mathbf{x} being the set of inputs of the ANN model (Table 3 shows a complete description of the different variables used).

With respect to the type of neurons of the hidden layer, in this study, two different basis functions (BFs) are considered:

1. Radial Basis Function (RBF), as a kernel BF, with a Gaussian Transfer Function (GTF), defined as follows:

$$B_j(\mathbf{x}_{i,d}, \mathbf{c}_j | \mathbf{r}_j) = \exp\left(-\frac{\|\mathbf{x}_{i,d} - \mathbf{c}_j\|^2}{2r_j^2}\right), \quad (6)$$

where $\mathbf{c}_j = (w_{j1}, w_{j2}, \dots, w_{jn})$ and $r_j = w_{j0}$ being the centroid and radius of the GTF, respectively.

2. Product Unit (PU), as a projection BF, defined as follows:

$$B_j(\mathbf{x}_{i,d}, \mathbf{w}_j) = \prod_{k=1}^n x_{i,d}^{w_{jk}}, \quad (7)$$

where $\mathbf{w}_j = \{w_{j1}, w_{j2}, \dots, w_{jn}\}$ are the weights from the input layer to the hidden layer.

Furthermore, combining BFs in the hidden layer has some advantages, such as providing flexibility to the decision rules. Any continuous function can be decomposed into two different types of functions: one belonging to the projection group and the other belonging to the kernel group (Donoho & Johnstone, 1989). In this sense, hybrid ANNs (both monotask and multitask) have also been considered. Such hybrid models combine PU (projection) and RBF (kernel) BFs in the hidden layer, i.e., they perform a linear combination of both types of BFs. These hybrid ANN models are defined as follows:

$$f_q(\mathbf{x}, \boldsymbol{\theta}) = \beta_{q0} + \sum_{j=1}^{m_1} \beta_{qj}^1 B_j^1(\mathbf{x}, \mathbf{w}_j^1) + \sum_{j=1}^{m_2} \beta_{qj}^2 B_j^2(\mathbf{x}, \mathbf{w}_j^2), \quad q = \{1, 2\}, \quad (8)$$

where m_1 and m_2 are the numbers of hidden neurons of the first and second types, respectively. $\boldsymbol{\theta} = \{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_q, \mathbf{w}_1^1, \dots, \mathbf{w}_{m_1}^1, \mathbf{w}_1^2, \dots, \mathbf{w}_{m_2}^2\}$ contains the coefficients of the ANN model, $\boldsymbol{\beta}_q = \{\beta_{q0}, \beta_{q1}^1, \dots, \beta_{qm_1}^1, \beta_{q1}^2, \dots, \beta_{qm_2}^2\}$ being the coefficients between the hidden layer and the q th output node, and \mathbf{w}_j^1 and \mathbf{w}_j^2 being the weights connecting the input layer to the j th hidden neuron of the first and the second type, respectively. $B_j^1(\mathbf{x}, \mathbf{w}_j^1)$ and $B_j^2(\mathbf{x}, \mathbf{w}_j^2)$ represent the basis functions of each type, defined in Eqs. (6) and (7).

The general forecast task proposed in this study considers the real number of cumulative contagions as the continuous variable to be forecasted. Consequently, for the evaluation of the ANNs, the mean square error (MSE) of the network output layer with respect to the real values is used as the loss function, defined as follows for the monotask models:

$$MSE(\mathbf{x}, \boldsymbol{\theta}) = \frac{\sum_{i=2}^5 \sum_{d=m_i}^{D_i-h} (y_{d+h,i} - \hat{y}_{d+h,i})^2}{\sum_{i=2}^5 D_i - h - m_i}, \quad (9)$$

where $y_{d+h,i}$ and $\hat{y}_{d+h,i}$ are the real and forecasted cumulative numbers of contagions, respectively. Note that $\hat{y}_{d+h,i}$ specifies the general output $f_q(\mathbf{x}, \boldsymbol{\theta})$ of the monotask model ($q = 1$) forecasting the cumulative number of contagions for the day $d + h$ of the i th outbreak.

Moreover, as mentioned above, in this study, two forecast horizons are considered: $h = 3$ or $h = 5$. Accordingly, the loss function to

evaluate the multitask models is defined as:

$$MSE(\mathbf{x}, \boldsymbol{\theta}) = \frac{\sum_{i=2}^5 \sum_{d=m_i}^{D_i-5} [(y_{d+3,i} - \hat{y}_{d+3,i})^2 + (y_{d+5,i} - \hat{y}_{d+5,i})^2]}{2 \sum_{i=2}^5 D_i - 5 - m_i}, \quad (10)$$

where, in this case, $\hat{y}_{d+3,i} = f_1(\mathbf{x}, \boldsymbol{\theta})$ and $\hat{y}_{d+5,i} = f_2(\mathbf{x}, \boldsymbol{\theta})$.

3.4. Evolutionary artificial neural networks

Even though ANNs are considered to be universal approximators, for the different BFs, the required training time and final complexities of the networks may vary significantly. The speed of convergence may be a problem for those cases involving a sufficiently high number of patterns in the training set or when the complexity of the network increases, becoming the main challenge to be addressed (Livni, Shalev-Shwartz, & Shamir, 2014). In most studies, gradient-based methods have been applied to optimize ANN parameters. However, due to the nonlinear nature of ANNs, the optimization process may converge to one of the many local optima existing on the error surface of these models (Sutton, 1986).

To perform a more exhaustive exploration process over the error surface, Yao (1993) presented a thorough review of evolutionary algorithms (EAs) applied to the optimization of ANN parameters. EAs use a set of different candidate solutions (population) to: (1) simulate the evolutionary process and natural selection and (2) provide the algorithm with a global search capability with the objective of discovering regions on the error surface where good performance solutions are located. Therefore, EAs represent an efficient method to optimize both the structure and weight connections of ANNs.

In this work, the EA proposed in Martínez-Estudillo, Martínez-Estudillo, Hervás-Martínez, and García-Pedrajas (2006) is applied to optimize the ANN models described in Section 3.3. In Algorithm 1 the pseudocode of the EA is shown. As the goal of the EA is to minimize the MSE of the ANNs, the fitness function $A(\mathbf{x}, \boldsymbol{\theta})$ guiding the evolutionary process is defined as a strictly decreasing transformation of the MSE:

$$A(\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{1 + MSE(\mathbf{x}, \boldsymbol{\theta})}. \quad (11)$$

Consequently, the fitness for the monotask or multitask ANNs is calculated using the MSE defined in Eq. (9) or Eq. (10), respectively.

The EA begins creating an initial population of N ANNs. Each ANN (individual) of the population is randomly generated, i.e., the number of neurons in the hidden layer, the number of connections of each hidden neuron (to link both the input layer to the hidden layer and the hidden layer to the output layer), and the weights of those connections are randomly generated considering the parameters described in Table 4 of Section 4.2. Next, the ANNs are evaluated and ranked according to their fitness.

After that, ANN optimization (evolutionary process) is performed as long as the stopping criteria are not met. Specifically, two stopping criteria are considered, as shown in Table 4 of the following Section 4.2. Hence, if either of them is reached, then the EA stops. In each generation of the evolutionary process, the worst 10% of ANNs are replaced with a copy of the best 10% of ANNs, which represents an elitist pressure since the best 10% of individuals (after being cloned, they constitute the 20% of the population) are optimized in a different way, as described below. Next, all individuals are evolved by simultaneously applying two types of mutations: parametric and structural.

Parametric mutation is applied to the best 10% of ANNs and updates their weights by adding Gaussian noise, whose variance decreases throughout the evolutionary process, i.e., the strength of the changes decreases as the individuals become better. In that way, this adaptive variance (particularly to each individual) dynamically modifies the intensity of the exploration of the error surface, favoring exploitation as the performance of the individual increases.

On the other hand, structural mutation (applied to the remaining 90% of ANNs) alters the structure of individuals. More specifically,

this type of mutation modifies the number of hidden neurons and their number of connections between input and output layers, i.e., it explores a wider area of the search space while trying to maintain the diversity of individuals. In particular, the EA applies five types of structural mutations: Add Neuron, Delete Neuron, Add Link, Delete Link and Neuron Fusion (Gutiérrez, Hervás, Carbonero, & Fernández, 2009), which are applied sequentially to each individual.

Thus, the performance of individuals is improved throughout the evolutionary process, generation after generation, maximizing their fitness (and hence minimizing their MSE). Finally, the ANN with the best fitness is selected as the final solution when the EA stops. Since the population is randomly initialized, the EA has a stochastic component. Therefore, different runs will lead to distinct populations and, consequently, to different final solutions.

In this way, and regardless of the BFs used in the hidden layer, the developed evolutionary artificial neural networks (EANNs) will be denominated as monotask EANNs (MoEANNs) and multitask EANNs (MuEANNs) depending on whether they perform the forecast tasks separately or simultaneously, respectively.

Algorithm 1 Evolutionary Algorithm

- 1: Randomly create a population of ANNs with size N
 - 2: Evaluate and rank ANNs with respect to their fitness using Eq. (11)
 - 3: **while** stopping criteria are not met **do**
 - 4: Replace the worst 10% of ANNs with a copy of the best 10% of ANNs
 - 5: Apply parametric mutation to the best 10% of ANNs
 - 6: Apply structural mutation to the remaining 90% of ANNs
 - 7: Evaluate and rank ANNs with respect to their fitness using Eq. (11)
 - 8: **end while**
 - 9: Select the best fitness individual as the final solution
-

4. Experiment settings and results

The settings used in the experimentation and the results obtained using EANN models are presented in this section. In addition, a statistical analysis is performed to evaluate significant differences between the models and datasets used, with respect to the distinct sanitary districts considered.

4.1. Datasets

To analyze the effect of the polynomial information in the forecast task, for each district, different multivariate time series datasets are constructed by combining different sets of input variables and autoregressive orders. Experimental results have empirically demonstrated that using the four polynomial coefficients simultaneously generates overfitted models, performing poorly on the generalization set. For this reason, only the coefficients defining the critical point of the 3-degree polynomial model are considered in the dataset building process. The critical point is produced when the function changes its concavity which is related to the instant where the outbreak gets out of control, and, therefore, the number of positive cases begins to rise exponentially. This moment is reached when the second derivative of the polynomial is equal to zero, i.e., $d_{\text{inflection}} = -\frac{\alpha_3}{3\alpha_4}$. Consequently, only these two coefficients will be considered as polynomial information in the datasets because they characterize the inflexion moment with α_4 having more importance.

Table 3 shows the inputs of the EANN models. Regarding the coefficients extracted from the polynomial approximation, except for the coefficient y , which is always included in the AR/VAR models, the other two coefficients, α_3 and α_4 , can be included or not. Moreover, apart from these coefficients, the number of days d_i from the beginning of the i th wave is always included as an input variable with the aim

Table 3

Inputs ($\mathbf{x}_{d,i}$) included in the EANNs for day d belonging to the i th outbreak. I is the total number of inputs for the generated datasets. Note that the different datasets are named according to the AR models used (either AR or VAR).

Lags	Dataset name	$\mathbf{x}_{d,i}$	I
$p = 1$	$AR(1)_y$	$\{y_{d-1,i}, d_i\}$	2
	$VAR(1)_{y,\alpha_4}$	$\{y_{d-1,i}, \alpha_{4,d-1,i}, d_i\}$	3
	$VAR(1)_{y,\alpha_3,\alpha_4}$	$\{y_{d-1,i}, \alpha_{3,d-1,i}, \alpha_{4,d-1,i}, d_i\}$	4
$p = 2$	$AR(2)_y$	$\{y_{d-1,i}, y_{d-2,i}, d_i\}$	3
	$VAR(2)_{y,\alpha_4}$	$\{y_{d-1,i}, \alpha_{4,d-1,i}, y_{d-2,i}, \alpha_{4,d-2,i}, d_i\}$	5
	$VAR(2)_{y,\alpha_3,\alpha_4}$	$\{y_{d-1,i}, \alpha_{4,d-1,i}, \alpha_{3,d-1,i}, y_{d-2,i}, \alpha_{4,d-2,i}, \alpha_{3,d-2,i}, d_i\}$	7

of temporally locating the fitted evolution. Notably, the $AR(p)$ and $VAR(p)$ models are applied with $p = 1, 2$. Therefore, the simplest dataset is built with only two input variables (y lagged 1 time and d_i), being identified as $AR(1)_y$. On the other hand, the most complex dataset is built with seven input variables (being y , α_3 and α_4 lagged two times and d_i), which is identified as $VAR(2)_{y,\alpha_3,\alpha_4}$. Note that the total number of input variables is specified in column I of Table 3.

Regarding the target or output variables, as mentioned in the previous Section 3.3, the cumulative number of contagions at two different time horizons is employed, $y_{d+3,i}$ and $y_{d+5,i}$. For the MoEANN approaches, two different datasets are built as $S_m = \{(\mathbf{x}_{d,i}, y_{d+h,i})\}$, depending on the forecasting horizon, $h = 3$ or $h = 5$. On the other hand, for the simultaneous forecasting with $h = 3$ and $h = 5$ carried out by the MuEANN approach, the dataset is built as $S_M = \{(\mathbf{x}_{d,i}, y_{d+3,i}, y_{d+5,i})\}$. Note that $\mathbf{x}_{d,i}$ is one of the six combinations presented in Table 3.

The train and test partitions have been generated while taking into account the three stages previously mentioned in Section 2. From each wave, a specific period has been included in the generalization test, aiming to include as many different cases as possible to build robust models. The rest of the stages are included in the training set. Specifically, the testing partition is composed of the beginning stage belonging to the second wave, the growth stage of the fourth wave and the stabilization stage from the third wave and fifth wave. This splitting methodology is shown in Fig. 5, where periods for training and testing sets are represented over the cumulative contagion curve and the evolution of polynomial coefficients (α_4 and α_3) for the district of Sevilla. This hold-out splitting methodology responds to the need to involve as much information as possible about the different stages of all the waves in the training set. With these nonrandom partitions, we aim to create a heterogeneous and representative set of observations for training and testing sets, given that the application of cross-validation would have increased unaffordably the needed runs of the algorithm. Therefore, from the 358 patterns that were involved in the datasets for each district, the final datasets used in the experimentation contain 350 and 354 observations for the two considered autoregressive orders $p = 1$ and $p = 2$, respectively. Note that for $p = 1$, the first observation from each of the 4 waves is needed to build the final datasets, whereas in the case of $p = 2$, the initial observations needed are 2. From these final datasets, 283 and 71 for $p = 1$, (and 280 and 70 for $p = 2$) patterns are used for building the training and testing sets, respectively (representing 80% and 20% of the total data).

4.2. Experimental settings

The contagion curve is specific for every district, varying in magnitude and transmission rate according to geographical and population parameters. To find the best model and dataset for each specific district, the experimentation carried out in this paper makes use of all generated options detailed in Table 3 for all districts. In this way, the considered BFs for the ANNs detailed in Section 3.3 have been applied to estimate the weights from the input layer to the hidden layer of the models, resulting in 18 different combinations per district: 3 ANNs (using each one a different BF combination: PU, RBF or the

Table 4

Parameter values that have been used in the EA for all models (PU, RBF and RBF+PU for both MoEANNs and MuEANNs).

Parameter	Value
Independent runs	40
Stopping criteria:	
(1) maximum number of generations	1500
(2) consecutive generations without improving individuals	10
Population size	1000
Number of hidden layers of each individual	1
Minimum number of hidden neurons (initialisation)	2
Maximum number of hidden neurons (initialisation)	3
Maximum number of hidden neurons (whole process)	4
Range of hidden neurons to be added or deleted	[1, 3]
Range of links to be added or deleted	[1, 5]
Range for weights between input and hidden layer	[0.1, 0.9]
Range for weights between hidden and output layer	[-10, 10]

hybrid one RBF+PU) applied to the 6 different datasets generated from the AR/VAR models. In addition, depending on whether the two forecast tasks are approached simultaneously or not, the two different approaches proposed in this work are compared: (1) MoEANN, which considers the two tasks separately (3 or 5 days ahead as different tasks), and (2) MuEANN, which groups both tasks carrying out them simultaneously. Consequently, 3 different models (2 MoEANNs and 1 MuEANN) were considered for the 18 combinations of each sanitary district, resulting in a total of 54 experiments being executed for each district.

Regardless of the district, dataset and computing error perspective, the evolutionary algorithm used to optimize the EANN models is applied with the same configuration. More specifically, in Table 4, the parameters of the EA are shown, together with their values.

The weight values of the connections from the hidden layer to the output layer are randomly generated regardless of the type of BF of the model. However, the weight values of the connections from the input layer to the hidden layer are randomly generated for PU units, whereas for RBF, units the weight values are initialized using K-means (Lloyd, 1982) as a clustering method to determine the Gaussian function centroids. For hybrid EANN models combining RBF+PU units, both types of BFs have the same probability of being selected in the initialization process of the individuals, and the final number of both types of units will depend on the evolutionary process.

Input features for all datasets are scaled in the interval [0.1, 0.9]. The output layer is composed of 1 and 2 linear units for the two MoEANN and the MuEANN approaches, respectively, and their output values are scaled in the interval [-10, 10].

The described configuration for performing the optimization of the models follows the guidelines published in Gutiérrez et al. (2009) and Martínez-Estudillo et al. (2006).

The EANN models containing SU basis functions in the hidden layer, either pure or combined with PU or RBF, were affected by a high overfitting and, consequently, performed poorly in the test set. Thus, this basis function has been excluded from this study.

4.3. Results

The results presented in this section aim to compare the model performance differences produced when polynomial coefficients are included as training features and the extent to which multitask models are able to take advantage of using the same structure and parameters for tackling several tasks simultaneously. The results are expressed in terms of root mean square error (RMSE, applying the square root to Eqs. (9) and (10)). As explained in Section 3.4 the EANN models are stochastic. Thus, 40 runs are performed and the results are expressed as the mean and standard deviation of the RMSE obtained for the generalization sets, i.e., patterns unseen during the training stage. Given the interest in knowing which BF is better and the best value for p in the AR/VAR

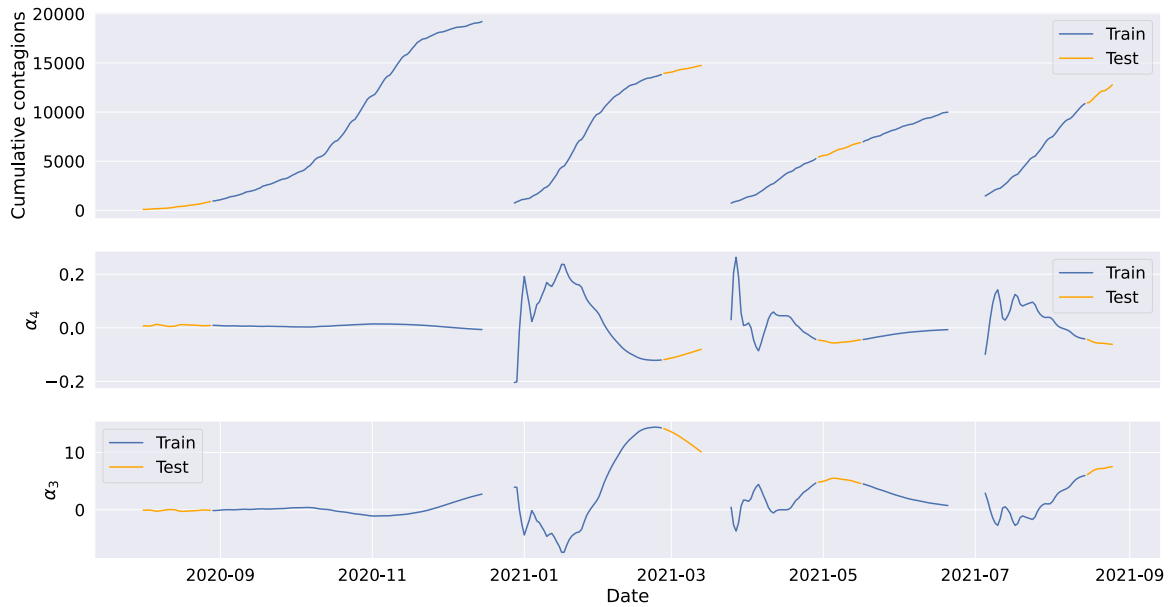


Fig. 5. Train and test partitions of the district of Sevilla.

models, the results are segregated considering the different options, as well as for the individual districts.

In addition, the performances of the models for the two forecasting horizons are also compared. Note that for single horizon forecasts, only the corresponding output neuron error of the MuEANN model is considered. Regarding individual forecasting horizons, Tables 5 and 6 show the results obtained for $y_{d+3,i}$ and $y_{d+5,i}$, respectively, in the eight districts considered. On the other hand, worth mentioning that when simultaneous time horizon forecasting tasks are approached, the results achieved by the MoEANN approaches (one tackling $y_{d+3,i}$ and the other tackling $y_{d+5,i}$) are the average of the errors of the two single MoEANN models. Table 7 shows the results obtained for all zones and datasets considering these two horizons simultaneously. In Tables 5–7, best results per forecasting approach and district are highlighted in bold, whereas the second best result is highlighted in italics.

Apart from the model performances, the complexities of the two MoEANN and the MuEANN models are an important aspect to be compared. Since almost all experiments resulted in the maximum limited number of nodes in the hidden layer (4 neurons), complexities are measured in terms of the total number of connections. Table 8 shows the average number of connections for each model, forecasting approach and dataset. For the sake of a fair comparison, model complexities are measured from a simultaneous forecast horizon perspective. For MoEANN models, the average of connections is computed for each input feature combination and autoregressive order. Then, the resulting averages for the two horizons are summed. Bold and italics represent the least complex and the second least complex architectures by forecasting approach, basis function and dataset, respectively. As expected, models become more complex when the number of autoregressive input features increases. Considering the average number of connections of the models, it can be said that the use of MuEANN models is justified. Even though the results obtained with this approach are slightly worse, the models are much simpler in all cases while their performances remain competitive.

In most cases, the results produced by the models involving one single output neuron (MoEANN) are slightly better than those obtained using the MuEANN approach, as seen in the results shown in Tables 5 and 6. This is to be expected since MoEANN model training is guided by considering only the error made for a single output variable, focusing the model training on a single objective. Nevertheless, it is important to point out the special case of the district of Málaga, where the results of

the single and simultaneous forecast improve if the models are trained considering the errors of the two horizons in the training set. This phenomenon is due to the capability of the MuEANN model to infer relationships between the related tasks used to optimize itself.

Considering the results achieved, it can be concluded that there is no single combination of BF and dataset that performs accurately for all districts. However, some highlights can be extracted from the experimental results. Pure RBF models always produce poor performances when modeling from the MuEANN perspective, and except for Cádiz Bay, they also underperform using the MoEANN approach. In the particular case of simultaneous forecasting (see MuEANN columns of Table 7), PU architecture is the best choice when modeling the cumulative curve of contagion of the districts of Almería, Granada, Huelva and Sevilla, while an RBF+PU architecture is needed for the rest of the locations. In this respect, a different number of input variable lags are needed to achieve the best results depending on the district to be modeled. Overall, considering Tables 5–7, it can be seen that hybridization outperforms pure model architectures in the districts of Málaga, Jaén and Cádiz when modeling from a MuEANN approach. On the other hand, MoEANN approaches achieve better results in Sevilla, Jaén and Granada when modeling the considered horizons, either individually or jointly.

Moreover, concerning the complexities presented in Table 8, the average number of links for the eight considered districts is higher for the hybrid architectures, regardless of the autoregressive order p and the number of model outputs. Pure PU models trained with $AR(p)_y$ are less complex than RBF models trained with the same dataset. Nevertheless, it is noteworthy that the opposite effect occurs when pure models are trained including polynomial information, i.e. when they are trained with the $VAR(p)_{y,\alpha_4}$ and $VAR(p)_{y,\alpha_4,\alpha_3}$ datasets.

The results show that the performances obtained for both approaches (the MoEANN and MuEANN models) improve when some sort of autoregressive polynomial information derived from curve fitting is included in the model training set, except for the district of Jaén. This fact demonstrates that adding lags of polynomial coefficients has a positive effect on the forecasting task, not only in terms of accuracy but also reducing, in some cases, the variance of the results. The results obtained in Jaén may be because it is the least populated district, and the fitness of the polynomial models along its outbreaks has the higher standard deviation of the eight districts, as shown in Table 2. Fig. 6 shows a boxplot diagram where the RMSE achieved with the MuEANN

Table 5

Performances of the MoEANN and MuEANN models trained with the different combinations of input features and autoregressive orders, p , evaluating the errors of the $\mathcal{V}_{d+3,d}$ forecast horizon. The results are expressed as $Mean_{SD}$ RMSE of the generalization set, and SD stands out for Standard Deviation.

District	BF	p	MoEANN			MuEANN		
			$AR(p)_y$	$VAR(p)_{y,a_4}$	$VAR(p)_{y,a_4,a_3}$	$AR(p)_y$	$VAR(p)_{y,a_4}$	$VAR(p)_{y,a_4,a_3}$
Almería	PU	1	312.61 _{13.72}	95.62 _{13.24}	85.42 _{9.52}	275.97 _{18.92}	119.67 _{23.75}	108.78 _{16.77}
		2	300.24 _{12.58}	103.39 _{15.85}	90.72 _{12.90}	275.81 _{22.80}	136.72 _{30.20}	126.75 _{26.41}
	RBF	1	312.11 _{13.72}	176.52 _{29.38}	181.87 _{49.91}	332.09 _{18.92}	257.04 _{67.98}	304.01 _{80.76}
		2	293.53 _{12.58}	184.22 _{46.05}	203.34 _{61.35}	353.01 _{22.80}	255.39 _{84.67}	301.07 _{72.12}
	RBF+PU	1	131.90 _{34.75}	107.12 _{26.05}	98.69 _{28.41}	147.53 _{34.31}	142.82 _{34.69}	146.51 _{42.91}
		2	136.14 _{54.19}	97.71 _{25.18}	99.61 _{46.13}	152.75 _{34.91}	164.12 _{45.26}	171.78 _{45.95}
Cádiz Bay	PU	1	358.68 _{4.41}	223.44 _{7.76}	222.20 _{10.90}	397.65 _{20.54}	237.84 _{32.91}	247.07 _{24.69}
		2	345.91 _{12.52}	234.57 _{15.14}	235.53 _{22.00}	354.00 _{33.69}	228.06 _{36.10}	229.83 _{41.18}
	RBF	1	292.56 _{4.41}	168.07 _{35.62}	173.03 _{33.94}	295.57 _{20.54}	268.97 _{91.29}	281.15 _{68.42}
		2	299.97 _{12.52}	153.89 _{37.15}	195.45 _{63.40}	303.96 _{33.69}	333.86 _{382.05}	271.46 _{78.10}
	RBF+PU	1	306.97 _{42.29}	201.65 _{27.82}	194.85 _{39.42}	226.17 _{38.12}	185.51 _{57.84}	174.02 _{59.10}
		2	295.82 _{58.44}	210.14 _{38.15}	191.68 _{40.64}	217.75 _{49.77}	184.24 _{67.41}	175.12 _{65.75}
Córdoba	PU	1	195.73 _{4.50}	82.58 _{5.87}	76.00 _{10.42}	194.04 _{8.90}	95.40 _{9.57}	93.70 _{8.34}
		2	182.71 _{13.97}	80.43 _{7.57}	74.41 _{12.31}	186.74 _{9.00}	89.40 _{9.87}	86.68 _{11.13}
	RBF	1	189.78 _{4.50}	155.67 _{32.39}	166.90 _{33.23}	226.67 _{8.90}	196.20 _{33.28}	209.91 _{40.78}
		2	186.41 _{13.97}	161.03 _{29.05}	176.94 _{30.49}	235.62 _{9.00}	197.97 _{36.07}	220.06 _{30.83}
	RBF+PU	1	116.65 _{20.27}	82.28 _{28.21}	71.96 _{16.16}	120.36 _{16.08}	74.58 _{10.96}	77.15 _{16.61}
		2	117.15 _{15.00}	77.14 _{13.39}	70.52 _{13.79}	120.75 _{15.61}	78.10 _{12.98}	78.60 _{15.63}
Granada	PU	1	198.09 _{5.62}	112.25 _{9.39}	118.49 _{13.14}	213.20 _{14.01}	121.39 _{11.73}	127.45 _{12.34}
		2	157.57 _{40.79}	114.70 _{19.86}	113.14 _{12.95}	210.26 _{17.62}	113.15 _{12.98}	128.26 _{17.07}
	RBF	1	203.49 _{5.62}	182.22 _{45.85}	169.50 _{42.52}	318.22 _{14.01}	227.12 _{43.82}	229.42 _{46.94}
		2	186.01 _{40.79}	168.01 _{31.91}	181.62 _{41.68}	316.98 _{17.62}	246.96 _{46.30}	270.13 _{51.13}
	RBF+PU	1	148.56 _{45.33}	130.61 _{104.41}	108.39 _{69.97}	173.50 _{36.95}	105.69 _{21.87}	109.03 _{26.16}
		2	152.28 _{59.65}	112.08 _{52.14}	115.38 _{103.97}	165.85 _{35.86}	104.71 _{22.91}	123.64 _{64.24}
Huelva Cost	PU	1	147.36 _{18.14}	61.30 _{10.94}	68.92 _{10.70}	139.29 _{9.39}	62.53 _{4.05}	67.97 _{8.01}
		2	147.93 _{18.95}	69.72 _{9.51}	76.85 _{14.70}	133.15 _{7.50}	60.94 _{4.69}	70.00 _{11.91}
	RBF	1	135.65 _{18.14}	87.27 _{18.56}	109.96 _{46.24}	158.09 _{9.39}	114.14 _{38.05}	141.25 _{41.63}
		2	133.62 _{18.95}	83.45 _{19.20}	121.65 _{45.67}	154.99 _{9.50}	108.66 _{35.66}	142.12 _{57.78}
	RBF+PU	1	89.77 _{33.70}	63.54 _{17.01}	71.23 _{29.77}	77.04 _{15.55}	69.55 _{20.88}	79.07 _{20.38}
		2	75.00 _{23.14}	67.65 _{21.77}	74.14 _{23.85}	75.46 _{17.85}	68.14 _{19.17}	69.24 _{16.99}
Jaén	PU	1	100.27 _{12.54}	78.89 _{6.54}	81.39 _{6.44}	117.33 _{4.59}	98.80 _{12.75}	105.99 _{8.80}
		2	86.57 _{16.14}	78.30 _{6.98}	77.30 _{8.38}	113.97 _{12.73}	96.97 _{10.78}	103.66 _{11.09}
	RBF	1	109.06 _{12.54}	96.61 _{12.49}	88.38 _{13.97}	112.81 _{4.59}	99.44 _{13.78}	112.94 _{32.96}
		2	108.08 _{16.14}	106.57 _{12.88}	96.79 _{19.70}	114.19 _{12.73}	115.94 _{25.38}	119.28 _{29.18}
	RBF+PU	1	74.03 _{35.94}	70.64 _{61.72}	94.16 _{49.77}	70.83 _{18.46}	84.69 _{43.50}	78.41 _{21.37}
		2	80.66 _{47.86}	74.21 _{34.16}	74.11 _{20.98}	70.64 _{23.42}	85.71 _{54.24}	77.53 _{19.70}
Málaga	PU	1	428.79 _{14.34}	312.39 _{20.39}	305.80 _{23.28}	391.10 _{28.60}	296.15 _{27.29}	293.25 _{33.05}
		2	423.88 _{43.40}	300.56 _{29.19}	293.06 _{39.57}	370.87 _{41.41}	281.45 _{40.26}	276.09 _{43.11}
	RBF	1	365.13 _{14.34}	467.07 _{85.02}	499.11 _{84.02}	451.47 _{28.60}	559.18 _{150.48}	559.78 _{105.88}
		2	395.36 _{43.40}	499.41 _{68.21}	528.72 _{68.64}	448.60 _{41.41}	564.19 _{107.76}	531.75 _{107.13}
	RBF+PU	1	358.95 _{33.49}	316.25 _{38.65}	332.67 _{148.19}	288.31 _{60.17}	255.68 _{56.73}	259.13 _{46.60}
		2	352.63 _{26.08}	312.10 _{48.37}	298.15 _{72.09}	274.45 _{65.26}	243.80 _{65.70}	267.21 _{127.55}
Sevilla	PU	1	268.46 _{18.32}	155.93 _{19.72}	156.78 _{27.77}	292.42 _{14.57}	147.59 _{19.40}	162.60 _{16.60}
		2	226.27 _{22.92}	174.82 _{23.72}	187.60 _{37.87}	270.86 _{16.85}	141.32 _{22.04}	158.98 _{19.26}
	RBF	1	388.28 _{18.32}	320.38 _{86.43}	348.93 _{123.80}	473.52 _{14.57}	446.58 _{102.65}	459.12 _{96.61}
		2	365.89 _{22.92}	331.61 _{91.40}	371.61 _{82.34}	483.75 _{16.85}	442.58 _{112.32}	518.32 _{71.35}
	RBF+PU	1	263.98 _{47.13}	171.73 _{39.08}	146.11 _{34.11}	273.38 _{58.92}	150.53 _{23.62}	152.44 _{21.21}
		2	256.68 _{35.30}	163.81 _{34.53}	150.71 _{20.47}	274.94 _{52.97}	159.38 _{30.53}	156.38 _{31.06}

model and the best $VAR(p)$ dataset is compared with the results of the analogous model trained with no polynomial data considering the dual and simultaneous horizon forecast. The improvement in the results when autoregressive polynomial coefficients are included in the training data is graphically noticeable in the districts except Jaén, especially in Córdoba, Huelva Coast, Almería, Sevilla and Granada, where the median obtained with the best combination is distinctly lower than its analogous model. There is no distinguished difference in the dispersion of the results obtained for the Cádiz and Málaga districts. However, in these districts, the worst result obtained with the $VAR(p)$ model is always better than the worst result obtained with the $AR(p)$ model. Additionally, to quantitatively summarize the

mentioned improvement, Table 9 presents the percentage accuracy gain of the double forecast results obtained with models trained with best $VAR(p)$ with respect to training with best $AR(p)$ datasets. In terms of percentage, the results for all the districts except Jaén are notably improved when modeled with polynomial information, with Almería being the district that benefits the most from this additional information, increasing the accuracy by 73.55% and 59.46% for the MoEANN and MuEANN approaches, respectively.

To graphically illustrate the improvement produced by the inclusion of the polynomial curve point descriptors in the training phase of the EANNs, in Figs. 7 and 8 the best performances obtained for the district of Sevilla with MuEANN models trained with $AR(p)$ and $VAR(p)$

Table 6

Performances of MoEANN and MuEANN models trained with the different combinations of input features and autoregressive orders, p , evaluating the errors of the $\mathcal{V}_{d+5,d}$ forecast horizon. The results are expressed as $Mean_{SD}$ RMSE of the generalization set, SD stands out for Standard Deviation.

District	BF	p	MoEANN			MuEANN		
			$AR(p)_y$	$VAR(p)_{y,a_4}$	$VAR(p)_{y,a_4,a_3}$	$AR(p)_y$	$VAR(p)_{y,a_4}$	$VAR(p)_{y,a_4,a_3}$
Almería	PU	1	426.60 _{20.90}	137.24 _{15.40}	124.94 _{12.66}	322.39 _{15.91}	140.40 _{22.01}	131.47 _{15.30}
		2	415.68 _{27.30}	147.07 _{21.20}	138.45 _{18.57}	326.93 _{33.21}	169.14 _{35.54}	146.16 _{25.96}
	RBF	1	421.46 _{20.90}	254.38 _{57.33}	261.43 _{76.61}	463.63 _{15.91}	272.98 _{60.76}	310.54 _{84.80}
		2	426.25 _{27.30}	248.08 _{70.96}	247.53 _{71.40}	466.01 _{33.21}	288.35 _{75.45}	307.70 _{77.09}
	RBF+PU	1	182.97 _{59.55}	147.01 _{31.56}	152.29 _{43.97}	215.80 _{37.53}	156.12 _{32.20}	154.41 _{27.30}
		2	180.70 _{55.17}	148.54 _{34.73}	137.52 _{42.93}	228.23 _{36.45}	174.99 _{42.88}	176.23 _{39.83}
Cádiz Bay	PU	1	501.23 _{25.16}	325.03 _{11.60}	322.42 _{24.88}	409.40 _{18.17}	276.86 _{29.40}	269.10 _{22.70}
		2	498.95 _{19.77}	312.92 _{13.03}	312.07 _{26.27}	429.54 _{34.05}	278.18 _{25.54}	257.23 _{43.86}
	RBF	1	460.89 _{25.16}	219.52 _{40.75}	242.40 _{67.64}	396.46 _{18.17}	324.42 _{93.69}	353.20 _{86.84}
		2	454.45 _{19.77}	217.76 _{48.41}	258.11 _{57.20}	410.00 _{34.05}	381.32 _{401.25}	326.98 _{77.12}
	RBF+PU	1	502.89 _{47.86}	285.99 _{47.07}	275.16 _{50.38}	293.54 _{37.96}	230.03 _{55.89}	216.91 _{53.80}
		2	442.87 _{107.47}	286.29 _{52.97}	267.10 _{67.05}	311.51 _{112.80}	220.24 _{58.14}	212.68 _{60.66}
Córdoba	PU	1	264.52 _{6.69}	132.21 _{5.69}	110.73 _{18.68}	252.26 _{4.47}	135.56 _{11.81}	124.95 _{15.13}
		2	251.62 _{11.44}	128.85 _{8.86}	115.51 _{18.41}	255.80 _{6.53}	133.14 _{10.32}	128.61 _{12.48}
	RBF	1	263.82 _{6.69}	207.74 _{37.88}	216.63 _{46.81}	328.34 _{4.47}	249.11 _{38.39}	260.34 _{43.14}
		2	258.87 _{11.44}	223.00 _{26.99}	228.65 _{28.77}	338.04 _{6.53}	258.78 _{41.17}	265.93 _{30.29}
	RBF+PU	1	174.55 _{27.13}	111.63 _{17.62}	111.06 _{18.58}	184.47 _{15.81}	125.03 _{13.78}	128.51 _{20.28}
		2	172.91 _{30.25}	121.88 _{19.65}	110.95 _{16.27}	190.21 _{18.62}	135.81 _{20.29}	131.58 _{20.06}
Granada	PU	1	279.01 _{21.82}	152.27 _{9.72}	164.82 _{18.47}	301.14 _{12.14}	139.00 _{7.85}	158.66 _{12.52}
		2	192.38 _{45.54}	152.24 _{14.59}	154.67 _{23.55}	311.78 _{21.74}	143.17 _{12.28}	165.91 _{15.85}
	RBF	1	294.82 _{21.82}	220.14 _{37.12}	212.94 _{52.10}	412.32 _{12.14}	284.42 _{41.23}	300.08 _{64.77}
		2	281.19 _{45.54}	242.33 _{38.38}	237.80 _{46.39}	407.60 _{21.74}	314.26 _{34.37}	324.77 _{54.32}
	RBF+PU	1	216.08 _{44.87}	176.55 _{118.66}	126.36 _{31.26}	240.34 _{61.51}	168.62 _{24.40}	170.27 _{36.28}
		2	207.33 _{44.49}	196.73 _{123.74}	158.56 _{75.74}	241.67 _{53.77}	169.84 _{27.97}	196.56 _{69.05}
Huelva Cost	PU	1	202.60 _{19.02}	86.37 _{6.78}	93.89 _{10.47}	198.09 _{10.09}	90.97 _{9.85}	92.13 _{9.49}
		2	203.57 _{15.97}	79.06 _{7.22}	89.88 _{8.48}	196.18 _{9.77}	89.03 _{9.64}	96.41 _{10.44}
	RBF	1	204.09 _{19.02}	114.35 _{23.67}	135.86 _{40.22}	226.78 _{10.09}	140.19 _{33.66}	164.75 _{36.37}
		2	189.28 _{15.97}	105.75 _{22.14}	137.26 _{39.83}	227.02 _{9.77}	134.19 _{30.46}	169.98 _{67.43}
	RBF+PU	1	127.05 _{66.98}	84.60 _{25.29}	95.66 _{39.06}	116.44 _{22.26}	97.03 _{21.84}	101.24 _{21.02}
		2	108.34 _{39.04}	79.92 _{20.69}	96.17 _{32.21}	114.94 _{25.31}	96.38 _{22.83}	94.87 _{28.81}
Jaén	PU	1	155.56 _{8.00}	117.36 _{10.04}	122.96 _{7.68}	149.90 _{5.03}	129.34 _{14.86}	117.03 _{10.98}
		2	128.23 _{28.05}	111.72 _{38.35}	115.08 _{12.26}	148.61 _{15.77}	127.90 _{13.89}	123.06 _{15.93}
	RBF	1	140.76 _{8.00}	121.14 _{13.64}	121.52 _{14.38}	181.86 _{5.03}	160.27 _{20.03}	166.23 _{27.05}
		2	138.02 _{22.05}	128.71 _{16.94}	124.17 _{24.20}	188.39 _{15.77}	169.52 _{24.90}	173.19 _{33.89}
	RBF+PU	1	106.77 _{54.71}	102.83 _{34.25}	109.76 _{36.07}	108.39 _{41.12}	124.17 _{91.61}	113.36 _{36.17}
		2	98.82 _{31.06}	115.55 _{103.98}	116.29 _{33.37}	102.82 _{30.77}	115.75 _{45.87}	108.93 _{24.21}
Málaga	PU	1	511.47 _{33.48}	410.57 _{22.16}	395.97 _{34.33}	308.50 _{26.07}	326.36 _{27.25}	325.70 _{30.92}
		2	528.86 _{76.06}	385.95 _{36.31}	365.23 _{46.20}	351.93 _{96.41}	306.17 _{37.17}	308.01 _{44.44}
	RBF	1	542.03 _{33.48}	608.60 _{67.13}	639.00 _{70.48}	528.34 _{26.07}	572.49 _{69.65}	581.85 _{87.50}
		2	570.64 _{76.06}	628.99 _{46.10}	687.12 _{61.91}	552.15 _{96.41}	594.26 _{80.60}	586.21 _{92.53}
	RBF+PU	1	492.83 _{55.81}	401.03 _{37.86}	407.21 _{57.16}	342.61 _{82.43}	288.63 _{58.38}	289.45 _{45.87}
		2	472.38 _{41.33}	419.93 _{73.59}	385.02 _{62.96}	352.97 _{78.50}	268.53 _{62.05}	303.64 _{111.44}
Sevilla	PU	1	378.64 _{22.45}	229.26 _{25.04}	209.98 _{38.14}	374.89 _{11.51}	206.52 _{23.67}	218.54 _{18.82}
		2	318.66 _{27.24}	260.88 _{34.39}	233.25 _{28.07}	385.27 _{28.82}	216.98 _{26.45}	226.48 _{25.73}
	RBF	1	553.19 _{22.45}	436.69 _{104.15}	427.74 _{91.88}	654.71 _{11.51}	556.51 _{109.41}	563.96 _{121.52}
		2	365.89 _{27.24}	470.94 _{89.03}	494.17 _{73.37}	665.62 _{28.82}	550.15 _{126.87}	613.40 _{78.47}
	RBF+PU	1	394.14 _{61.60}	227.19 _{66.71}	196.44 _{56.24}	449.79 _{92.50}	252.38 _{41.93}	251.01 _{30.25}
		2	394.38 _{79.74}	251.60 _{74.07}	226.51 _{44.38}	435.62 _{88.04}	271.02 _{39.57}	258.87 _{42.70}

datasets are compared. More specifically, in Figs. 7(a) and 8(a), the model test predictions are represented over time with the real curve of cumulative contagions of the four waves. For both forecast horizons, the model trained with the best $VAR(p)$ dataset fits better than the one trained with the best $AR(p)$. This improvement is especially noteworthy in the second and fourth waves for both horizons and in the third wave for the mid-term forecast. In Figs. 7(b) and 8(b), where the real test values for mid- and long-term horizons are scattered with the values predicted by both models, the improvements are visibly easier to appreciate. Points located in the gray equality line represent a perfect prediction. Predictions resulting from models trained with $VAR(p)$ models of the two considered forecast horizons are closer to

this line, representing a more accurate prognosis of the cumulative number of contagions. Note that in Figs. 7(b) and 8(b) the plotted data correspond, from left to right, to the second, fourth, fifth and third waves.

4.4. Statistical analysis

Once all the results for the 18 models and datasets combinations for each district have been descriptively analyzed, a statistical analysis has been carried out to obtain robust conclusions from several aspects of the experimentation. For simplicity purposes, this analysis is mainly focused on the results obtained in simultaneous horizon forecasting with

Table 7

Performances of MoEANN and MuEANN models trained with the different combinations of input features and autoregressive orders, p , simultaneously evaluating the errors of the $y_{d+3,i}$ and $y_{d+5,i}$ forecast horizons. The results are expressed as $Mean_{SD}$ RMSE of the generalization set, SD stands out for Standard Deviation.

District	BF	p	MoEANN			MuEANN		
			$AR(p)_y$	$VAR(p)_{y,a_i}$	$VAR(p)_{y,a_i,a_3}$	$AR(p)_y$	$VAR(p)_{y,a_i}$	$VAR(p)_{y,a_i,a_3}$
Almería	PU	1	369.61 _{57.15}	116.71 _{21.15}	104.92 _{11.46}	299.18 _{29.05}	130.04 _{25.13}	120.13 _{11.46}
		2	357.96 _{57.89}	125.23 _{22.26}	114.58 _{24.19}	301.37 _{38.27}	152.93 _{36.74}	136.45 _{27.93}
	RBF	1	366.09 _{57.15}	214.94 _{39.49}	221.15 _{40.57}	397.86 _{29.05}	265.01 _{64.96}	307.28 _{82.87}
		2	359.89 _{57.89}	216.15 _{32.84}	225.44 _{23.55}	409.51 _{38.27}	271.87 _{81.87}	304.38 _{74.72}
	RBF+PU	1	157.43 _{26.44}	127.65 _{20.66}	125.89 _{27.47}	181.67 _{46.53}	149.47 _{34.12}	150.46 _{36.18}
		2	158.42 _{23.48}	123.13 _{25.99}	118.56 _{20.09}	190.49 _{48.97}	169.55 _{44.42}	174.01 _{43.06}
Cádiz Bay	PU	1	441.32 _{72.28}	274.24 _{50.89}	271.64 _{50.29}	403.52 _{20.26}	257.35 _{36.81}	258.09 _{26.15}
		2	422.43 _{76.62}	273.75 _{39.36}	273.80 _{38.58}	391.77 _{50.73}	253.12 _{35.98}	243.53 _{44.69}
	RBF	1	375.60 _{72.28}	193.79 _{26.46}	207.71 _{35.41}	346.01 _{20.26}	296.70 _{36.56}	317.17 _{86.08}
		2	378.18 _{76.62}	185.42 _{11.46}	226.38 _{32.28}	356.98 _{50.73}	357.59 _{392.49}	299.22 _{82.43}
	RBF+PU	1	412.70 _{98.75}	245.51 _{42.65}	237.12 _{40.77}	259.86 _{60.81}	207.77 _{61.08}	195.47 _{60.45}
		2	369.35 _{74.09}	248.22 _{38.67}	229.87 _{38.42}	264.63 _{98.98}	202.24 _{65.47}	193.90 _{11.46}
Córdoba	PU	1	230.13 _{34.48}	108.05 _{24.94}	93.58 _{17.78}	223.15 _{29.95}	115.48 _{22.78}	109.32 _{19.84}
		2	217.16 _{34.64}	104.64 _{24.38}	95.22 _{20.92}	221.27 _{35.41}	111.27 _{24.09}	107.65 _{24.07}
	RBF	1	226.80 _{34.48}	182.39 _{26.71}	192.40 _{25.66}	277.50 _{29.95}	222.66 _{44.62}	235.13 _{48.97}
		2	222.64 _{34.64}	192.41 _{31.43}	202.79 _{26.42}	286.83 _{35.41}	228.37 _{49.22}	243.00 _{38.21}
	RBF+PU	1	145.60 _{29.36}	96.96 _{15.44}	91.51 _{19.99}	152.42 _{35.80}	99.80 _{11.46}	102.83 _{31.67}
		2	145.03 _{28.29}	99.79 _{22.74}	90.74 _{11.46}	155.48 _{38.75}	106.96 _{33.51}	105.09 _{32.01}
Granada	PU	1	238.55 _{40.63}	132.52 _{20.25}	141.66 _{23.50}	257.17 _{45.88}	130.19 _{13.31}	143.06 _{19.95}
		2	174.97 _{18.60}	133.47 _{19.22}	133.91 _{21.20}	261.02 _{54.48}	128.16 _{11.46}	147.09 _{25.01}
	RBF	1	249.73 _{40.63}	201.18 _{20.02}	191.22 _{22.79}	365.27 _{45.88}	255.77 _{51.29}	264.75 _{66.69}
		2	233.60 _{18.60}	205.17 _{37.63}	209.71 _{28.87}	362.29 _{54.48}	280.61 _{52.86}	297.45 _{59.41}
	RBF+PU	1	182.75 _{34.42}	153.58 _{25.28}	117.49 _{11.46}	206.92 _{60.76}	137.16 _{39.07}	139.65 _{44.02}
		2	179.81 _{28.46}	155.49 _{43.36}	136.97 _{23.58}	203.76 _{59.38}	137.27 _{41.40}	160.10 _{76.00}
Huelva Cost	PU	1	174.98 _{27.96}	73.83 _{2.88}	81.40 _{12.90}	168.69 _{30.97}	76.75 _{16.09}	80.05 _{14.94}
		2	175.75 _{28.13}	74.39 _{5.49}	83.36 _{7.35}	164.67 _{32.70}	74.99 _{11.46}	83.21 _{17.32}
	RBF	1	169.87 _{27.96}	100.81 _{14.29}	122.91 _{14.53}	192.44 _{30.97}	127.17 _{38.21}	153.00 _{40.81}
		2	161.45 _{28.13}	94.74 _{12.04}	129.46 _{10.18}	191.00 _{32.70}	121.43 _{35.53}	156.05 _{64.32}
	RBF+PU	1	108.41 _{19.94}	74.07 _{11.49}	83.44 _{13.55}	96.74 _{27.51}	83.29 _{25.40}	90.15 _{23.49}
		2	91.88 _{17.58}	73.79 _{11.46}	85.15 _{12.22}	95.20 _{29.48}	82.26 _{25.37}	82.05 _{26.90}
Jaén	PU	1	127.92 _{27.83}	98.13 _{19.45}	102.44 _{20.96}	133.62 _{16.98}	114.07 _{20.61}	111.51 _{11.38}
		2	107.40 _{21.28}	95.01 _{16.94}	96.43 _{19.16}	131.29 _{22.48}	112.43 _{19.84}	113.36 _{16.81}
	RBF	1	124.91 _{27.83}	109.04 _{12.78}	104.95 _{16.99}	147.34 _{16.98}	129.86 _{34.94}	139.58 _{40.24}
		2	123.05 _{21.28}	117.36 _{11.73}	110.48 _{14.47}	151.29 _{22.48}	142.73 _{36.74}	146.24 _{41.56}
	RBF+PU	1	90.61 _{17.70}	86.73 _{11.46}	101.96 _{10.18}	89.61 _{36.99}	104.43 _{34.37}	95.89 _{34.46}
		2	89.74 _{11.04}	94.88 _{22.28}	95.20 _{21.72}	86.73 _{11.46}	100.73 _{52.43}	93.23 _{27.09}
Málaga	PU	1	472.89 _{41.72}	361.48 _{49.30}	351.46 _{45.40}	349.80 _{49.54}	311.25 _{31.17}	309.48 _{35.88}
		2	476.37 _{53.05}	343.80 _{43.08}	329.15 _{11.46}	361.40 _{44.79}	293.81 _{40.67}	292.05 _{46.60}
	RBF	1	453.58 _{41.72}	537.83 _{71.30}	569.05 _{70.50}	489.90 _{49.54}	565.83 _{117.44}	570.81 _{97.75}
		2	483.00 _{53.05}	564.20 _{65.23}	607.92 _{79.61}	500.37 _{74.79}	579.22 _{96.33}	558.98 _{103.73}
	RBF+PU	1	431.63 _{67.52}	359.18 _{42.84}	370.41 _{38.63}	315.46 _{77.10}	272.16 _{59.87}	274.29 _{48.66}
		2	412.51 _{60.16}	366.02 _{54.48}	342.14 _{44.21}	313.71 _{82.17}	256.17 _{11.46}	285.42 _{121.15}
Sevilla	PU	1	323.55 _{55.28}	193.53 _{36.98}	183.04 _{27.22}	333.65 _{43.27}	177.05 _{11.46}	190.57 _{33.13}
		2	273.05 _{46.47}	217.85 _{43.37}	210.43 _{23.54}	328.06 _{61.88}	179.15 _{44.99}	192.73 _{40.69}
	RBF	1	470.73 _{55.28}	378.54 _{58.97}	387.83 _{40.75}	564.12 _{43.27}	501.54 _{19.48}	511.54 _{121.64}
		2	365.89 _{46.47}	401.27 _{70.31}	432.89 _{61.91}	574.68 _{61.88}	496.37 _{131.34}	565.86 _{88.79}
	RBF+PU	1	332.48 _{65.59}	199.81 _{28.67}	171.59 _{11.46}	361.59 _{117.44}	201.45 _{61.25}	201.72 _{55.78}
		2	326.40 _{69.27}	208.26 _{44.51}	188.61 _{38.33}	355.28 _{108.32}	215.20 _{66.06}	207.62 _{63.40}

the MoEANN and MuEANN approaches (see Table 7). As mentioned in Section 4.3, these models are less complex than using both MoEANNs (with one forecasting horizon each), while performances remain competitive. These tests assume normality in the data being compared. In this respect, a previous Kolmogorov–Smirnov test (Massey, 1951) was applied using the results from the 40 different seeds before performing the comparison tests. In all cases, the p -values obtained from the Kolmogorov–Smirnov test show that the null hypothesis of normality is accepted for the different sets of results used.

For reasons of clarity, several comparisons of mean tests have been performed to conclude significant differences within several aspects of the results obtained. First, the results of the best MuEANN model and

training data combination for all districts are contrasted with the best results obtained from the MoEANN model, which performs better in most cases. Table 10 shows the p -values obtained in a paired t-test for the mean results of the best models in both approaches. Considering a level of significance of $\alpha = 0.05$, significant differences can only be found in the results of the districts of Córdoba, Almería and Málaga (i.e., 3 out of 8 sanitary districts). In this latter case, the MuEANN method is more accurate than the monotask method. For this specific sanitary district, in addition to being less complex in terms of total links, models that simultaneously predict mid- and long-term time forecasting horizons perform significantly better. Notwithstanding the additional difficulty of modeling two different output time horizons,

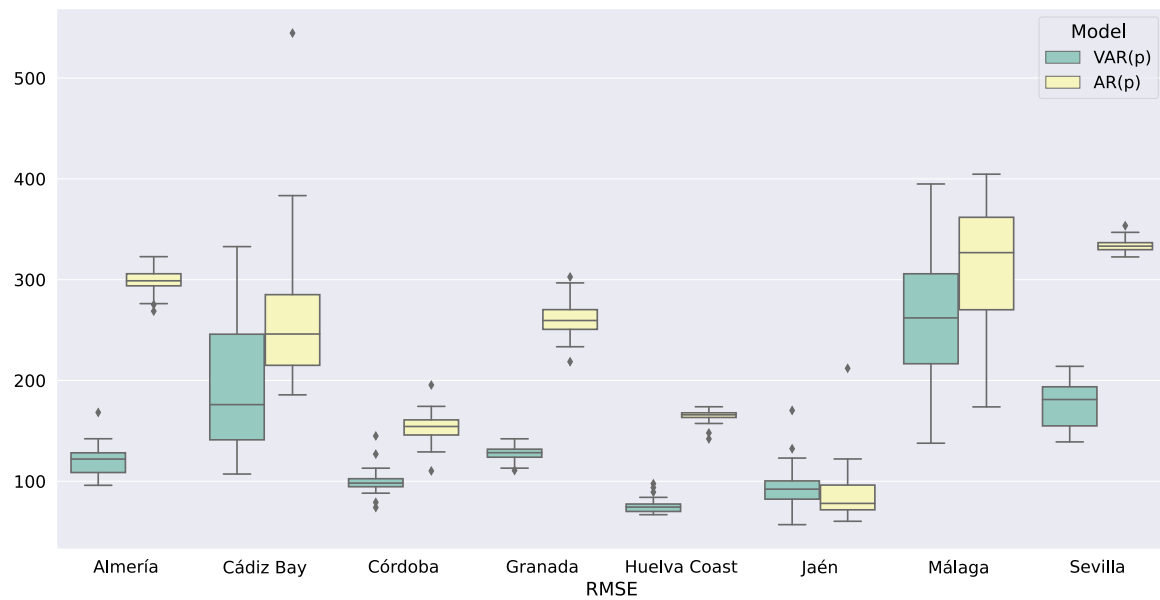


Fig. 6. Comparison boxplots of simultaneous forecast results using MuEANN models including ($VAR(p)$) and not including ($AR(p)$) polynomial information in the training set. The results are expressed as the RMSE obtained for the test set of the 40 executions for each district.

Table 8

Comparison of MoEANN and MuEANN model complexities considering double and simultaneous forecasting, $y_{d+3,d+5,i}$. The results are expressed in terms of the average number of links involved in the EANNs of the eight considered districts.

	BF	p	Input dataset		
			$AR(p)_y$	$VAR(p)_{y,\alpha_4}$	$VAR(p)_{y,\alpha_4,\alpha_3}$
MoEANN	PU	1	18.70	23.26	24.80
		2	20.76	26.74	29.42
	RBF	1	19.46	20.45	21.84
		2	21.14	25.50	27.74
	RBF+PU	1	20.81	24.49	26.80
		2	22.13	29.02	33.09
MuEANN	PU	1	13.04	16.20	17.03
		2	13.73	18.59	19.53
	RBF	1	14.33	14.91	15.48
		2	15.26	16.92	18.16
	RBF+PU	1	15.41	16.85	17.67
		2	15.82	19.17	20.59

Table 9

Percentage accuracy gain obtained in simultaneous double-horizon forecasting, $y_{d+3,i}$ and $y_{d+5,i}$, training the models with the best $VAR(p)$ datasets with respect to the best $AR(p)$ in all districts.

District	$y_{d+3,d+5}$	
	MoEANN	MuEANN
Córdoba	37.43%	34.52%
Huelva Cost	19.68%	54.46%
Almería	73.55%	59.84%
Cádiz Bay	50.97%	26.72%
Sevilla	48.39%	46.93%
Granada	35.71%	50.90%
Jaén	4.28%	-7.49%
Málaga	30.90%	18.34%
Mean	37.61%	35.53%

this approach takes advantage of the inherent relationship between the two tasks addressed achieving similar results to the models with a single time horizon target. For the rest of the districts, the mean results obtained in both approaches are not significantly different, although the MuEANN approach is simpler in terms of the number of connections, and, thus, easier to implement.

Table 10

Statistical differences between the average mean RMSE test results of the MoEANN and MuEANN best models.

District	t	p-value
Córdoba	-3.26	2.31E-03 ^a
Huelva Cost	-0.42	6.78E-01
Almería	-6.21	2.63E-07 ^a
Cádiz Bay	-0.85	3.98E-01
Sevilla	-0.92	3.61E-01
Granada	-1.69	9.82E-02
Jaén	-1.18	2.46E-01
Málaga	6.75	4.69E-08 ^b

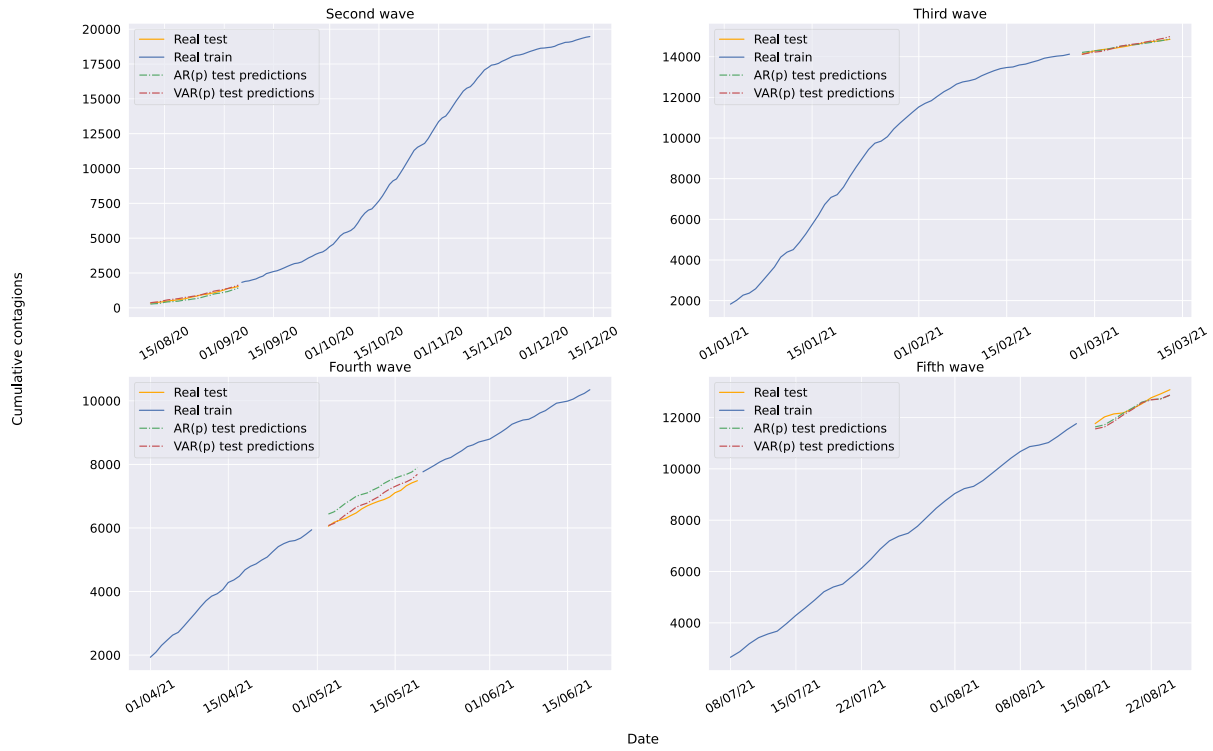
^aStatistically significant differences favoring MoEANN method.

^bStatistically significant differences favoring MuEANN method.

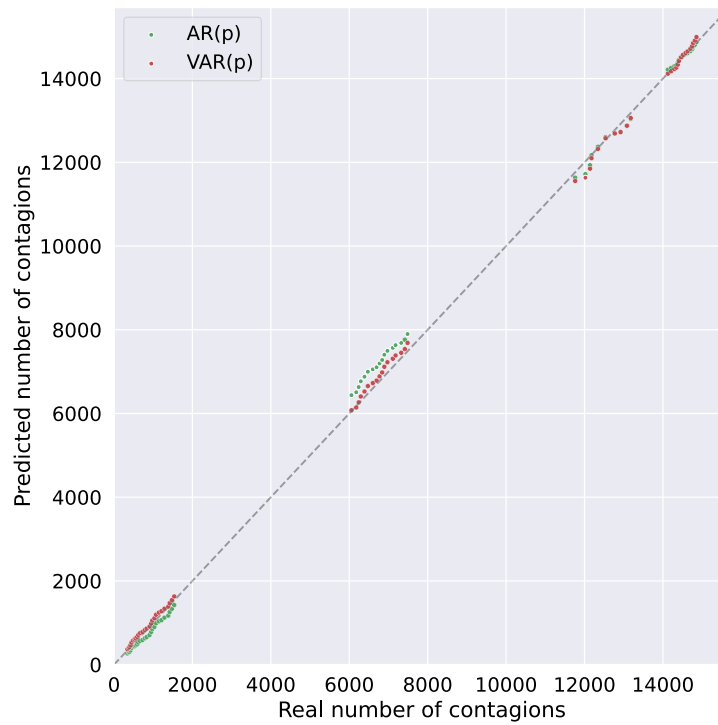
To analyze the effect of the polynomial information on the performance of the models, a paired t-test was applied to the results obtained with the best $VAR(p)$ and $AR(p)$ datasets for every horizon and model approach. The results, expressed as the p-value obtained in the test, are presented in Table 11 for all districts, forecast horizon approaches and methodologies (MoEANN or MuEANN). With the exception of Jaén, where no significant differences can be found between the best results obtained using the $AR(p)$ or $VAR(p)$ datasets, the rest of the models benefit significantly from the additional data derived from the polynomial curve fitting.

5. Conclusions

In this paper, a novel approach for extracting inherent information of the cumulative curve of contagions produced by the COVID-19 transmission rate has been proposed. By iteratively fitting a 3-degree polynomial, a 4 coefficient representation of the curve is generated for every day of a specific outbreak. The experimentation carried out, using the information of eight specific locations in Andalusia, Spain, demonstrates that using the coefficients that define the polynomial curve inflexion instant improves model performances when forecasting the mid- and long-term values of the original curve. In the experimentation carried out in this paper, evolutionary artificial neural networks (EANNs) are used to model the cumulative contagion curve

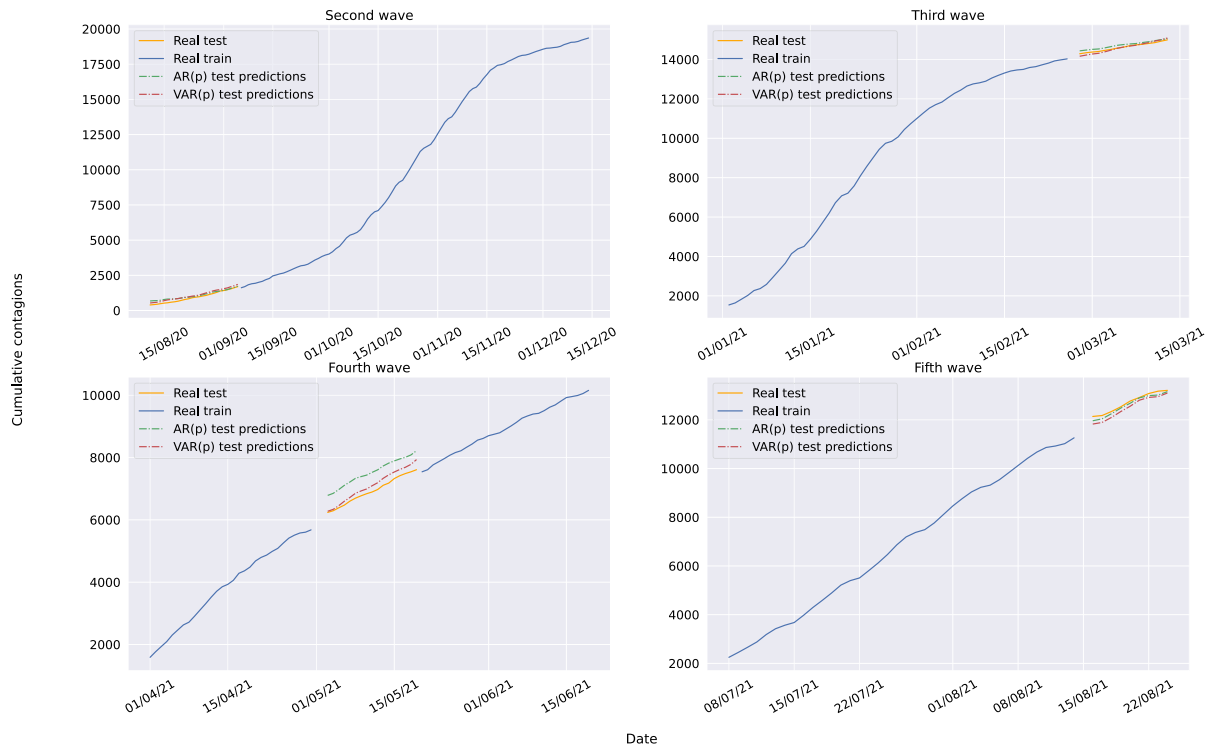


(a) $y_{d+3,i}$ test predictions over time on the cumulative curve of contagions of Sevilla using the best $VAR(p)$ and $AR(p)$ datasets.

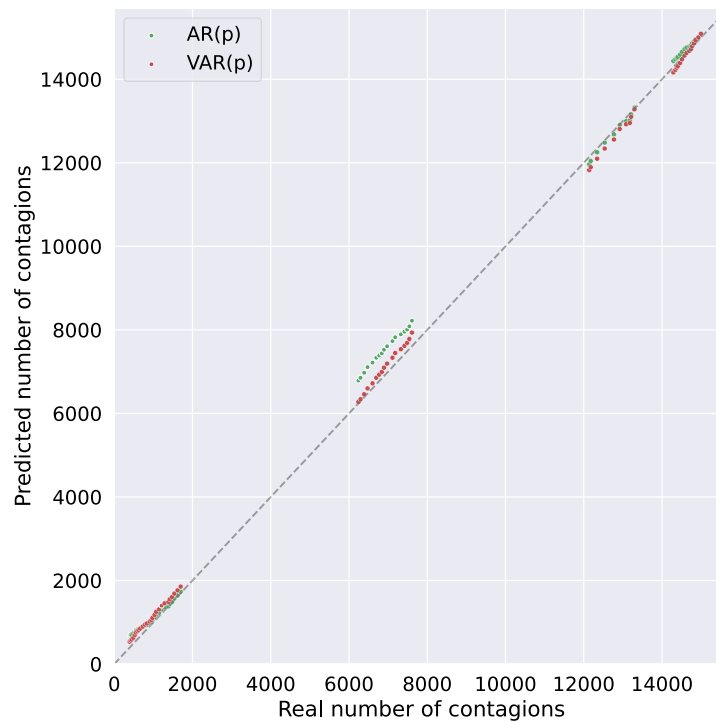


(b) Scatter plot of real and predicted $y_{d+3,i}$ test values of Sevilla using the best $VAR(p)$ and $AR(p)$ datasets.

Fig. 7. $y_{d+3,i}$ test predictions of the cumulative number of Sevilla over time (a) and scattered with real values (b) using the best $VAR(p)$ and $AR(p)$ datasets.



(a) $y_{d+5,i}$ test predictions over time on the cumulative curve of contagions of Sevilla using the best MuEANN model trained with $VAR(p)$ and $AR(p)$ datasets.



(b) Scatter plot of real and predicted $y_{d+5,i}$ test values of Sevilla using the best $VAR(p)$ and $AR(p)$ datasets.

Fig. 8. $y_{d+5,i}$ test predictions of the cumulative number of Sevilla over time (a) and scattered with real values (b) using the best MuEANN model trained with $VAR(p)$ and $AR(p)$ datasets.

of different districts using two different basis functions and considering the hybridization of both. For each location, different datasets are used to estimate the transfer function parameters using combinations of

autoregressive information of the original curve and extracted polynomial features. Due to the unique and specific characteristics of the sanitary districts considered, there is no single combination of network

Table 11

Statistical differences between the best results obtained with the different models trained with the $VAR(p)$ and $AR(p)$ datasets in the eight districts. The results are expressed as the resulting p -value of a paired t -test.

District	$y_{d+3,i}$ MoEANN	MuEANN	$y_{d+5,i}$ MoEANN	MuEANN	$y_{d+3,d+5,i}$ MoEANN	MuEANN
Córdoba	1.63E-15 ^a	2.11E-18 ^a	2.27E-12 ^a	2.58E-21 ^a	4.17E-15 ^a	3.63E-21 ^a
Huelva Cost	1.79E-03 ^a	3.84E-06 ^a	5.95E-05 ^a	5.86E-07 ^a	1.45E-05 ^a	4.36E-07 ^a
Almería	5.84E-10 ^a	1.52E-09 ^a	2.57E-07 ^a	4.75E-16 ^a	1.64E-09 ^a	5.89E-14 ^a
Cádiz Bay	1.91E-21 ^a	6.66E-04 ^a	3.59E-15 ^a	2.54E-08 ^a	1.56E-14 ^a	3.67E-07 ^a
Sevilla	6.57E-15 ^a	3.64E-28 ^a	5.26E-16 ^a	2.01E-33 ^a	6.29E-24 ^a	1.39E-32 ^a
Granada	4.90E-03 ^a	8.82E-12 ^a	4.35E-08 ^a	7.73E-15 ^a	1.84E-07 ^a	1.26E-13 ^a
Jaén	7.51E-01	1.56E-01	5.89E-01	3.22E-01	9.45E-01	2.23E-01
Málaga	2.00E-09 ^a	3.93E-02 ^a	8.02E-12 ^a	3.06E-04 ^a	3.76E-14 ^a	1.98E-04 ^a

^aStatistically significant differences favoring models trained with $VAR(p)$ datasets.

architecture and data that perform well for all cases. However, on the whole, the methodology applied achieves excellent results.

However, from the results obtained, it can be concluded that including lags of the polynomial features to the forecaster significantly improves the results, except for the district of Jaén. In most cases, the best performances obtained in Jaén are achieved using only autoregressive data of the number of cumulative contagion time series. In this district, the methods that slightly benefit from the additional information of the polynomial coefficients (MoEANNs approaches modeling $y_{d+3,i}$ and $y_{d+3,d+5,i}$) return results that are not significantly different from the same approaches using only autoregressive information, as seen in the p -values presented in Table 11.

Moreover, two different EANN approaches have been applied. On the one hand, monotask EANNs (MoEANN models) have been considered for modeling one target value at a time, i.e., two different models are needed to forecast on two different forecasting horizons. On the other hand, a multitask EANN (MuEANN model) has been used for forecasting both horizons simultaneously, resulting in a much simpler model. It is expected that models performing only one task are more accurate, as their optimization procedure is less complex. Nevertheless, there are no significant differences between the performances achieved by the MoEANN approaches and the MuEANN one for the sanitary districts of Huelva, Cádiz, Sevilla, Granada and Jaén, although the complexity of MuEANN models is lower in terms of average number links. Furthermore, in the district of Málaga, the MuEANN models perform significantly better than the average results of the two best MoEANN models. In this case, the MuEANN approach benefits from the intrinsic relationship between the two output targets.

Finally, it has been demonstrated that considering the features extracted from the inherent behavior of a process in the modeling stage can significantly improve the performance of the tasks carried out. Applying a similar methodology for any kind of problem, if the intrinsic problem nature is known, adapting a low-variance model as a description step may provide high quality descriptors that are able to support several supervised and unsupervised tasks.

CRedit authorship contribution statement

Miguel Díaz-Lozano: Formal analysis, Visualization, Methodology, Writing – original draft. **David Guijo-Rubio:** Validation, Writing – original draft, Writing – review & editing. **Pedro Antonio Gutiérrez:** Software, Methodology, Writing – review & editing. **Antonio Manuel Gómez-Orellana:** Software. **Isaac Túniz:** Project administration, Resources. **Luis Ortigosa-Moreno:** Project administration, Resources. **Armando Romanos-Rodríguez:** Project administration, Resources. **Javier Padillo-Ruiz:** Project administration, Resources. **César Hervás-Martínez:** Funding acquisition, Conceptualization, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2020). DeTraC: Transfer learning of class decomposed medical images in convolutional neural networks. *IEEE Access*, 8, 74901–74913. <http://dx.doi.org/10.1109/ACCESS.2020.2989273>.
- Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 51(2), 854–864. <http://dx.doi.org/10.1007/s10489-020-01829-7>.
- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11), Article e00938. <http://dx.doi.org/10.1016/j.heliyon.2018.e00938>.
- Abraham, C., Cornillon, P. A., Matzner-Lber, E., & Molinari, N. (2003). Unsupervised curve clustering using B-splines. *Scandinavian Journal of Statistics*, 30(3), 581–595. <http://dx.doi.org/10.1111/1467-9469.00350>.
- An, C., Lim, H., Kim, D.-W., Chang, J. H., Choi, Y. J., & Kim, S. W. (2020). Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Scientific Reports*, 10(1), 1–11. <http://dx.doi.org/10.1038/s41598-020-75767-2>.
- Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One*, 15(3), 1–21. <http://dx.doi.org/10.1371/journal.pone.0230405>.
- Au, K.-F., Choi, T.-M., & Yu, Y. (2008). Fashion retail forecasting by evolutionary neural networks. *International Journal of Production Economics*, 114(2), 615–630. <http://dx.doi.org/10.1016/j.ijpe.2007.06.013>.
- Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data Brief*, 29, Article 105340. <http://dx.doi.org/10.1016/j.dib.2020.105340>.
- Bishop, C. M., et al. (1995). *Neural networks for pattern recognition*. Oxford University Press.
- Broomhead, D. S., & Lowe, D. (1988). *Radial basis functions, multi-variable functional interpolation and adaptive networks: Technical Report*, Royal Signals and Radar Establishment Malvern (United Kingdom).
- Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*, 395(10223), 507–513. [http://dx.doi.org/10.1016/S0140-6736\(20\)30211-7](http://dx.doi.org/10.1016/S0140-6736(20)30211-7).
- Chiroma, H., Abdulkareem, S., & Herawan, T. (2015). Evolutionary neural network model for west texas intermediate crude oil price prediction. *Applied Energy*, 142, 266–273. <http://dx.doi.org/10.1016/j.apenergy.2014.12.045>.
- Desai, P. S. (2021). News sentiment informed time-series analyzing AI (SITALA) to curb the spread of COVID-19 in houston. *Expert Systems with Applications*, 180, Article 115104. <http://dx.doi.org/10.1016/j.eswa.2021.115104>.
- Donoho, D. L., & Johnstone, I. M. (1989). Projection-based approximation and a duality with kernel methods. *The Annals of Statistics*, 17(1), 58–106. <http://dx.doi.org/10.1214/aos/1176347004>.
- Durbin, R., & Rumelhart, D. E. (1989). Product units: A computationally powerful and biologically plausible extension to backpropagation networks. *Neural Computation*, 1(1), 133–142. <http://dx.doi.org/10.1162/neco.1989.1.1.133>.
- Fidan, H., & Yuksel, M. E. (2022). A comparative study for determining Covid-19 risk levels by unsupervised machine learning methods. *Expert Systems with Applications*, 190, Article 116243. <http://dx.doi.org/10.1016/j.eswa.2021.116243>.
- Fu, T.-c. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24(1), 164–181. <http://dx.doi.org/10.1016/j.engappai.2010.09.007>.

- Garg, A., Salehi, S., Rocca, M. L., Garner, R., & Duncan, D. (2022). Efficient and visualizable convolutional neural networks for COVID-19 classification using chest CT. *Expert Systems with Applications*, 195, Article 116540. <http://dx.doi.org/10.1016/j.eswa.2022.116540>.
- Geetha, S., Narayanamoorthy, S., Manirathinam, T., & Kang, D. (2021). Fuzzy case-based reasoning approach for finding COVID-19 patients priority in hospitals at source shortage period. *Expert Systems with Applications*, 178, Article 114997. <http://dx.doi.org/10.1016/j.eswa.2021.114997>.
- Gómez-Orellana, A., Guijo-Rubio, D., Gutiérrez, P., & Hervás-Martínez, C. (2022). Simultaneous short-term significant wave height and energy flux prediction using zonal multi-task evolutionary artificial neural networks. *Renewable Energy*, 184, 975–989. <http://dx.doi.org/10.1016/j.renene.2021.11.122>.
- Gutiérrez, P. A., Hervás, C., Carbonero, M., & Fernández, J. C. (2009). Combined projection and kernel basis functions for classification in evolutionary neural networks. *Neurocomputing*, 72(13–15), 2731–2742. <http://dx.doi.org/10.1016/j.neucom.2008.09.020>.
- Hamidi, M., Ghassemian, H., & Imani, M. (2018). Classification of heart sound signal using curve fitting and fractal dimension. *Biomedical Signal Processing and Control*, 39, 351–359. <http://dx.doi.org/10.1016/j.bspc.2017.08.002>.
- Hauser, A., Counotte, M. J., Margossian, C. C., Konstantinoudis, G., Low, N., Althaus, C. L., et al. (2020). Estimation of SARS-CoV-2 mortality during the early stages of an epidemic: A modeling study in Hubei, China, and six regions in Europe. *PLoS Medicine*, 17(7), 1–17. <http://dx.doi.org/10.1371/journal.pmed.1003189>.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. [http://dx.doi.org/10.1016/0893-6080\(89\)90020-8](http://dx.doi.org/10.1016/0893-6080(89)90020-8).
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 395(10223), 497–506. [http://dx.doi.org/10.1016/S0140-6736\(20\)30183-5](http://dx.doi.org/10.1016/S0140-6736(20)30183-5).
- Iftimie, S., López-Azcona, A. F., Vallverdú, I., Hernández-Flix, S., de Febrer, G., Parra, S., et al. (2021). First and second waves of coronavirus disease-19: A comparative study in hospitalized patients in reus, Spain. *PLoS One*, 16, 1–13. <http://dx.doi.org/10.1371/journal.pone.0248029>.
- Katris, C. (2021). A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece. *Expert Systems with Applications*, 166, Article 114077. <http://dx.doi.org/10.1016/j.eswa.2020.114077>.
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of A Mathematical and Physical Character*, 115(772), 700–721. <http://dx.doi.org/10.1098/rspa.1927.0118>.
- Khan, M., Mehran, M. T., Haq, Z. U., Ullah, Z., Naqvi, S. R., Ihsan, M., et al. (2021). Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review. *Expert Systems with Applications*, 185, Article 115695. <http://dx.doi.org/10.1016/j.eswa.2021.115695>.
- Kharroubi, S., & Saleh, F. (2020). Are lockdown measures effective against COVID-19? *Frontiers in Public Health*, 8, 610. <http://dx.doi.org/10.3389/fpubh.2020.549692>.
- Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., et al. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490), 493–497. <http://dx.doi.org/10.1126/science.abb4218>.
- Livadiotis, G. (2020). Statistical analysis of the impact of environmental temperature on the exponential growth rate of cases infected by COVID-19. *PLoS One*, 15(5), Article e0233875. <http://dx.doi.org/10.1371/journal.pone.0233875>.
- Livni, R., Shalev-Shwartz, S., & Shamir, O. (2014). On the computational efficiency of training neural networks. In *Advances in neural information processing systems 27: Annual conference on neural information processing systems 2014, December 8-13 2014, Montreal, Quebec, Canada* (pp. 855–863).
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137. <http://dx.doi.org/10.1109/TIT.1982.1056489>.
- Lu, W. Z., Fan, H., & Lo, S. (2003). Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. *Neurocomputing*, 51, 387–400. [http://dx.doi.org/10.1016/S0925-2312\(02\)00623-9](http://dx.doi.org/10.1016/S0925-2312(02)00623-9).
- Maital, S., & Barzani, E. (2020). The global economic impact of COVID-19: A summary of research. *Samuel Neaman Institute for National Policy Research*, 2020, 1–12.
- Malki, Z., Atlam, E.-S., Hassani, A. E., Dagnew, G., Elhosseini, M. A., & Gad, I. (2020). Association between weather data and COVID-19 pandemic predicting mortality rate: Machine learning approaches. *Chaos, Solitons & Fractals*, 138, Article 1110137. <http://dx.doi.org/10.1016/j.chaos.2020.1110137>.
- Martin, H., Gustafsson, F., Hörnsten, E. G., Winquist, F., Nilsson, L. E., Ljung, L., et al. (1998). Bacteria classification based on feature extraction from sensor data. *Biotechnology Techniques*, 12, 319–324. <http://dx.doi.org/10.1023/A:1008814701152>.
- Martínez-Estudillo, A., Martínez-Estudillo, F., Hervás-Martínez, C., & García-Pedrajas, N. (2006). Evolutionary product unit based neural networks for regression. *Neural Networks*, 19(4), 477–486. <http://dx.doi.org/10.1016/j.neunet.2005.11.001>.
- Mason, K., Duggan, J., & Howley, E. (2018). Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks. *Energy*, 155, 705–720. <http://dx.doi.org/10.1016/j.energy.2018.04.192>.
- Massey, F. J., Jr. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253), 68–78. <http://dx.doi.org/10.2307/2280095>.
- Motulsky, H. J., & Ransnas, L. A. (1987). Fitting curves to data using nonlinear regression: a practical and nonmathematical review. *The FASEB Journal*, 1(5), 365–374. <http://dx.doi.org/10.1096/fasebj.1.5.3315805>.
- Ng, M.-Y., Lee, E. Y., Yang, J., Yang, F., Li, X., Wang, H., et al. (2020). Imaging profile of the COVID-19 infection: radiologic findings and literature review. *Radiology: Cardiothoracic Imaging*, 2(1), Article e200034. <http://dx.doi.org/10.1148/ryct.2020200034>.
- Papastefanopoulos, V., Linardatos, P., & Kotsiantis, S. (2020). COVID-19: A comparison of time series methods to forecast percentage of active cases per population. *Applied Sciences*, 10(11), <http://dx.doi.org/10.3390/app10113880>.
- Poggio, T., & Girosi, F. (1989). *A theory of networks for approximation and learning: Technical report*, Massachusetts INST of TECH Cambridge Artificial Intelligence Lab.
- Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., & dos Santos Coelho, L. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals*, 135, Article 109853. <http://dx.doi.org/10.1016/j.chaos.2020.109853>.
- Ricon-Becker, I., Tarrasch, R., Blinder, P., & Ben-Eliyahu, S. (2020). A seven-day cycle in COVID-19 infection, hospitalization, and mortality rates: Do weekend social interactions kill susceptible people? <http://dx.doi.org/10.1101/2020.05.03.20089508>, MedRxiv.
- Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases*, 26(7), 1470. <http://dx.doi.org/10.3201/eid2607.200282>.
- Sethy, P. K., & Behera, S. K. (2020). Detection of coronavirus disease (covid-19) based on deep features. preprint on webpage at pdfs.semanticscholar.org/9da0/35f1d7372cfe52167ff301bc12d5f415caf1.pdf.
- Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., et al. (2021). Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18(6), 2775–2780. <http://dx.doi.org/10.1101/2020.02.23.20026930>.
- Sutton, R. (1986). Two problems with back propagation and other steepest descent learning procedures for networks. In *Proceedings of the eighth annual conference of the cognitive science society*, 1986 (pp. 823–832).
- Tamal, M., Alshammari, M., Alabdullah, M., Hourani, R., Alola, H. A., & Hegazi, T. M. (2021). An integrated framework with machine learning and radiomics for accurate and rapid early diagnosis of COVID-19 from chest X-ray. *Expert Systems with Applications*, 180, Article 115152. <http://dx.doi.org/10.1016/j.eswa.2021.115152>.
- Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45. <http://dx.doi.org/10.7287/peerj.preprints.3190v2>.
- Verma, H., Mandal, S., & Gupta, A. (2022). Temporal deep learning architecture for prediction of COVID-19 cases in India. *Expert Systems with Applications*, 195, Article 116611. <http://dx.doi.org/10.1016/j.eswa.2022.116611>.
- Wang, L., Lin, Z. Q., & Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, 10(1), 1–12. <http://dx.doi.org/10.1038/s41598-020-76550-z>.
- Wang, P., Zheng, X., Li, J., & Zhu, B. (2020). Prediction of epidemic trends in COVID-19 with logistic model and machine learning techniques. *Chaos, Solitons & Fractals*, 139, Article 110058. <http://dx.doi.org/10.1016/j.chaos.2020.110058>.
- Yao, X. (1993). A review of evolutionary artificial neural networks. *International Journal of Intelligent Systems*, 8(4), 539–567. <http://dx.doi.org/10.1007/s10462-011-9270-6>.
- Yi-Hui, L. (2007). Evolutionary neural network modeling for forecasting the field failure data of repairable systems. *Expert Systems with Applications*, 33(4), 1090–1096. <http://dx.doi.org/10.1016/j.eswa.2006.08.032>.
- Zivot, E., & Wang, J. (2003). Vector autoregressive models for multivariate time series. In *Modeling financial time series with s-plus®* (pp. 369–413). New York, NY: Springer New York, http://dx.doi.org/10.1007/978-0-387-21763-5_11, (Chapter 12).