Journal of Business Analytics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjba20

Estimation and demographic analysis of COVID-19 infections with respect to weather factors in Europe

Reza Gharoie Ahangar, Robert Pavur, Mahdi Fathi & Abdulazeez Shaik

To cite this article: Reza Gharoie Ahangar, Robert Pavur, Mahdi Fathi & Abdulazeez Shaik (2020) Estimation and demographic analysis of COVID-19 infections with respect to weather factors in Europe, Journal of Business Analytics, 3:2, 93-106, DOI: 10.1080/2573234X.2020.1832866

To link to this article: https://doi.org/10.1080/2573234X.2020.1832866

Published online: 01 Nov 2020.

JOURNAL OF BUSINESS ANALYTICS
2020, VOL. 3, NO. 2, 93–106
https://doi.org/10.1080/2573234X.2020.1832866

THE OPERATIONAL RESEARCH SOCIETY
Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Estimation and demographic analysis of COVID-19 infections with respect to weather factors in Europe

Reza Gharoie Ahangar, Robert Pavur, Mahdi Fathi and Abdulazeez Shaik

Department of Information Technology and Decision Sciences, G. Brint Ryan College of Business, University of North Texas, 1155 Union Circle #305249, Denton, TX 76201, USA

ABSTRACT
The main objective of this study is to investigate the relationship between the COVID-19 and the weather factors of the most populated and industrialised countries in Europe and propose the best mathematical model to forecast the daily number of COVID-19 cases. To find the relationship between the COVID-19 and the weather factors of absolute humidity and temperature in Spain, France, Italy, Germany, and the United Kingdom, we conducted a Poisson analysis. We also used the General Linear Neural Network (GRNN) model to forecast the trend and number of daily COVID-19 cases in these European countries. The results reveal a statistically significant negative relationship between the number of COVID-19 infections and weather factors of temperature & absolute humidity. Furthermore, the results show a stronger negative relationship between COVID-19 and absolute humidity than temperature. In our proposed GRNN method, we find better compatibility for the COVID-19 cases in Italy relative to the other European countries in this study.

ARTICLE HISTORY
Received 19 July 2020
Accepted 30 September 2020

KEY WORDS
COVID-19; weather; poisson; GRNN; Europe

1. Introduction

A novel coronavirus originated from Wuhan, China, in December 2019 and can be traced back to a cluster of severe pneumonia cases (Li et al., 2020; World Health Organization, 2020). The World Health Organisation called this novel coronavirus as COVID-19 (World Health Organization, 2020). COVID-19 is a coronavirus known as SARS-CoV-2. The earliest cases have pointed to the zoonotic origin as they may have originated in a live seafood market.

This virus spreads primarily from person to person, mainly by respiratory droplets formed during coughs or sneezes by an infected individual. These droplets may land in surrounding people's mouths or noses and be inhaled into the lungs. The virus mostly spreads when people are in close contact with each other (about 6 feet or 2 metres). The COVID-19 is spreading rapidly and sustainably amongst humans. Research from the current COVID-19 pandemic shows that this virus is transmitting more rapidly than influenza (World Health Organization, 2020). Human-to-human contact has led to a rapid spread of the disease, which has totalled around 20 million cases globally, and about 3.5 million cases in Europe alone. Over 1.3 million cases have been reported in Spain, France, Italy, Germany, and the United Kingdom (U.K.) as of August 5 2020 (World Health Organization, 2020).

However, at the earliest stage, the transmission of coronavirus was similar to that of the severe acute respiratory syndrome (SARS) (Riou & Althaus, 2020), but the current total number of COVID-19 cases in the world has surpassed the total number of SARS since 2003 (J. T. Wu et al., 2020; World Health Organization, 2003; R. Zhang et al., 2020).

Several factors such as age (Zhao et al., 2011), genetics, blood types, ethnicity & race, diseases (Zietz & Tatonetti, 2020), social habits, and environmental factors (Eslami & Jalili, 2020) affect the rate of high infections and deaths by COVDI-19 among people in all around the world. Some environmental factors such as food, water, sanitation, and weather (World Health Organization, 2020) can be the most related reasons for the spread of infectious diseases. A study in Hong Kong stated that airflow and ventilation are likely factors of SARS transmission (Pica & Bouvier, 2012). Some other factors, such as healthcare availability (Ji et al., 2020), emigration (Chen et al., 2020), or travel restriction (Chinazzi et al., 2020), might affect the spread of infectious diseases such as COVID-19. Among these environmental factors, the weather is one of the factors that people cannot control. Recent studies show that weather factors such as temperature and humidity (Sajadi et al., 2020; Wang et al., 2020) affect the spread of COVID-19.

COVID-19 has a global footprint with "hot" clusters occurring in well-defined climate conditions (Araujo & Naimi, 2020). The virus originated in Wuhan, China (Munster et al., 2020), then spread in epidemic proportions throughout the northern

CONTACT Mahdi Fathi mahdi.fathi@unt.edu Department of Information Technology and Decision Sciences, G. Brint Ryan College of Business, University of North Texas, 1155 Union Circle #305249, Denton, TX 76201, USA

hemisphere, namely from China through the Middle East, then to Europe and the United States. Some studies have illustrated a relationship between SARS-CoV-1 cases and the temperature of some cities in China (Tan et al., 2005). This evidence was further displayed in a different study in Hong Kong (Lin et al., 2006), which indicated a decrease in cases as the temperature rose from 15ºC to 29ºC degrees. A possible explanation is that SARS-CoV-1 can survive on smooth surfaces (Casanova et al., 2010; Chan et al., 2011) for over five days in relative humidity of 40–50% (Araujo & Naimi, 2020; Doremalen et al., 2013), and it drastically loses its viability as temperatures increased (Casanova et al., 2010; Chan et al., 2011). SARS-CoV-1 and SARS-CoV-2 are in the same category and are shown to have similarities (Van Doremalen et al., 2020).

A connection exists between the COVID-19 virus and the respiratory syncytial virus (RSV) (Bloom-Feshbach et al., 2013) since both exhibit distinct winter seasonality in temperate latitudes (Lofgren et al., 2007; Tamerius et al., 2011). Winter conditions characterised by low absolute humidity may increase the spread and survival of the influenza virus (Lowen et al., 2006, 2008; Shaman & Kohn, 2009). Other studies (Bukhari & Jameel, 2020; Gupta et al., 2020; Oliveiros et al., 2020; Sajadi et al., 2020) have investigated the effect of average monthly temperatures, relative and absolute humidity, precipitation, and wind speed, on COVID-19 and the results support that there is a relationship between the number of infection and weather.

Studies show that until the last week of February 2020, a majority of COVID-19 cases occurred in the northern hemisphere (Caramelo et al., 2020). The northern hemisphere with the highest median age population and low temperatures during the winter has seen COVID-19 spread at epidemic proportions. In contrast, the southern hemisphere, including most African counties, has not seen this phenomenon. Therefore, a principal research question is, how do meteorological characteristics such as temperature and humidity modulate COVID-19's duplication time in their countries?

This study investigates the effect of temperature and humidity on the number of daily cases in highly infected and populated countries in Europe: Spain, France, Italy, Germany, and the U.K., and make a comparison according to their level of humidity, temperature, and predict future cases by Artificial Neural Networks (ANNs). This framework can help health officials in these countries gain insights into efficient ways to control the spread of COVID-19.

The organisation of this study is as follows. In Section 2, the data collection, which contains temperature and humidity data and the daily number of COVID-19 cases, are discussed. In Section 3, the methods and ANNs are discussed. Descriptive statistics and results from Poisson regression and ANNs analyses are provided in Sections 4, 5, and 6, respectively. Finally, Section 7 presents the discussion and conclusion.

2. Data collection

2.1. Temperature and humidity data

The weather data collected in this study include the average daily temperature in Fahrenheit and the relative humidity. The weather data were collected from Visual Crossing Weather's website in Germany at https://www.visualcrossing.com, which is part of the library for the National Centres for Environmental Information (NCEI). Data were collected from March 1 2020, through August 5 2020. European countries use Celsius as a measure for temperature, but the units of collected data in this study are Fahrenheit, that is, Fahrenheit units were converted to Celsius units.

Moreover, some studies (Bukhari & Jameel, 2020; Gupta et al., 2020) have revealed that absolute humidity may be a more compelling factor than relative humidity in predicting the number of COVID-19 infection. Therefore we calculated the daily Absolute Humidity (g/m3) in an excel sheet using Clausius Clapeyron's formula (Herrmann & Bucksch, 2014) as follows.

$$AH = \frac{6.112 * RH * 2.1674 * e^{\frac{(17.67*T)}{(T+243.5)}}}{(273.15 + T)} \quad (1)$$

where AH is Absolute Humidity, RH is Relative Humidity and T is the average of daily temperature

2.2. Daily number of COVID-19 data

The number of daily new COVID-19 cases in Spain, France, Italy, Germany, and the United Kingdom was taken from John Hopkins University Coronavirus Resource Centre repository (Johns Hopkins, 2020). The collected data are the daily number of COVID-19 cases from March 1 2020, through August 5 2020.

3. Methods

In this study, first, descriptive analysis is used to demonstrate a possible relationship between daily new cases and weather factors to draw the hypothesis of our research graphically. Then, the Poisson regression method is implemented to find the statistical relationship between the COVID-19 and the weather factors of temperature and absolute humidity. Next, ANNs[1] forecasting method is used to predict the number and the trend of new daily cases of COVID-19 in Spain, France, Italy, Germany, and the United Kingdom, which were the most infected countries in Europe in the first half of the year 2020. The statistical software and the programming language used in this study are the SAS 9.4 and Python 3.8.

3.1. Artificial neural network

ANNs are adopted from the human brain's biological neural systems and consist of some characteristics that a large number of neurons are interconnected to each other, and operate in parallel (Gurney, 1997).

Recent studies show that the ANNs models can help the health care systems to investigate the spread of COVID-19 (Bandyopadhyay & Dutta, 2020; Huang et al., 2020; Pal et al., 2020).

Several evaluations on the previous SARS-CoV -1 virus show that data corresponding to new cases may follow a Gaussian or Exponential distribution (Bai & Jin, 2005; Hsieh et al., 2004; Lai, 2005; Wang & Ruan, 2004). The GRNN is a type of ANNs that uses a Gaussian activation function in the hidden layer (Specht, 1991). Figure (1) shows the structure of a GRNN, which consists of input units, pattern (hidden) units, summation units, and output units, and it depicts an estimated vector of Y from a measurement of vector X with a feedforward structure network (Tou & Gonzalez, 1974).

The computation of the most probable value of Y for each value of X is based on a set number of noisy measurements of X, and their Y values that can be considered a regression of dependent, Y, and independent X. The included variables of X and Y are considered vectors. Within the system, X is input, and Y is viewed as an output (Specht, 1991).

3.2. Mathematical structure of a general regression neural network

The mathematical structure of regression used in GRNN is as following. Let the f(x, y) is the joint continuous probability density function, and x, and y are its vector and scalar random variables, respectively. If we call the measured value of the vector random variable x as X, then equation (2) shows the conditional mean of y given X (Specht, 1991).

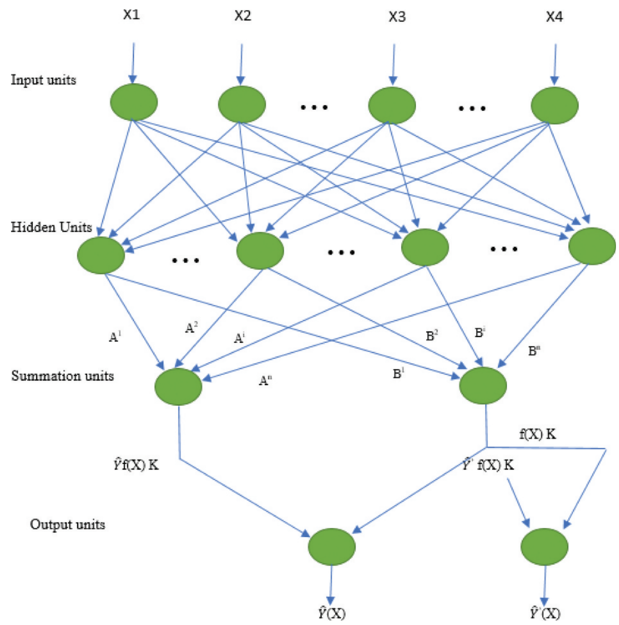

Figure 1. General regression neural network block diagram.

$$E[y|X] = \frac{\int_{-\infty}^{\infty} y f(X,y)dy}{\int_{-\infty}^{\infty} f(X,y)dy} \tag{2}$$

The density f(x, y) is usually estimated by a sample of observations x and y. We use some class of consistent estimators in nonparametric statistics estimation for different cases (Cacoullos, 1966; Parzen, 1962), which helps to estimate the PDF. The equation (3), shows the probability estimator f(X, Y), which n is the sample of observation, p indicates the dimension of the vector variable x, and σ shows the width of the sample probability for each sample of X_i and Y_i. (Specht, 1991).

Algorithm: GRNN

Procedure
begin
__initialize__ the COVID-19 Cases (Y), variables (X) and algorithm schemes (ID);
input:
X_1: Input temperature
X_2: Input absolute humidity
X_3: Input relative humidity
output:
Y: Output number of daily COVID-19 cases
function:
f = GRNN (inputs = $X_1 X_2 X_3$), Y=output, weight = w^n, ε: Threshold parameter, $d_i = |f(x_i) - y_i|$ for i= 1 to n
*Apply robust fitting and find unit vector * size*
Update the f_{best}
Repeat *the steps for the iteration n from 0*
$\boldsymbol{if |w_i^n - w_i^{n+1}| < \varepsilon}$ ***then***
end *of the iteration and* ***then***
break
else
Apply all the above processes
until *the end of the iterations*
end

Figure 2. The overall structure of proposed GRNN algorithm.

$$\hat{f}(X,Y) = \frac{1}{(2\pi)^{(p+1)/2}\sigma^{(p+1)}} \cdot \frac{1}{n}\sum_{i=1}^{n}\exp\left[-\frac{(X-X^i)^T(X-X^i)}{2\sigma^2}\right] \cdot \exp\left[-\frac{(Y-Y^i)^2}{2\sigma^2}\right] \tag{3}$$

After combing and reordering the equations (2) and (3), we get the desired conditional mean estimator, which it shows in equation (4).

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n}\exp\left[-\frac{(X-X^i)^T(X-X^i)}{2\sigma^2}\right]\int_{-\infty}^{\infty} y\exp\left[-\frac{(y-Y^i)^2}{2\sigma^2}\right]dy}{\sum_{i=1}^{n}\exp\left[-\frac{(X-X^i)^T(X-X^i)}{2\sigma^2}\right]\int_{-\infty}^{\infty}\exp\left[-\frac{(y-Y^i)^2}{2\sigma^2}\right]dy} \tag{4}$$

and the scalar function is as follows,

$$D_i^2 = \left(X-X^i\right)^T\left(X-X^i\right) \tag{5}$$

and by substituting equation (5) into equation (4) yields the following equation (6), which indicates the summations over the observations.

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y^i\exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n}\exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \tag{6}$$

The density estimator function in equation (3), is consistent, continuous estimators at all points (x, y) of function n (Cacoullos, 1966; Parzen, 1962) such that.

$$\lim_{n\to\infty}\sigma(n) = 0 \tag{7}$$

and

$$\lim_{n\to\infty} n\sigma(n) = \infty \tag{8}$$

In the next step, we need to scale all input variables to get the same range. To do this, we need to estimate the probability density function with a kernel with the same dimensional width, which we call this process normalisation of variables. Then for computational simplicity, the Gaussian kernel in equation (3), replace with some functions to express them as neural network nodes, which result in the estimator in equation (9) and is called the city block distance (Specht, 1991).

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y^i\exp\left(-\frac{C_i}{\sigma}\right)}{\sum_{i=1}^{n}\exp\left(-\frac{C_i}{\sigma}\right)}$$

where

$$C_i = \sum_{j=1}^{p}\left|X_j - X_j^i\right| \tag{9}$$

We can use the estimator in equation (9) if the number of observations (X, Y) is small. Still, if the number of observations is large, we need to use the clustering technique to group the samples for each neuron to measure the distance of input vectors from the cluster centre (Burrascano, 1991). Then the equation (6), after adding a new variable, N, which indicates the number of samples of i_{th} cluster centre, can be rewritten as equation (10).

$$\hat{Y}(X) = \frac{\sum_{i=1}^{m} A^i\exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{m} B_i\exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \tag{10}$$

In equation (11), the values of the coefficients for cluster i after k observations are $A^i(k)$ and $B^i(k)$, which $A^i(k)$ is the sum of the Y values, and $B^i(k)$ is the number of samples assigned to cluster i (Specht, 1991).

$$\begin{aligned} A^i(k) &= A^i(k-1) + Y^j \\ B^i(k) &= B^i(k-1) + 1 \end{aligned} \tag{11}$$

The above equations change to equation (12) after adding forgetting function to the coefficients A and B. It is appropriate for a system with varying characteristics. The first two equations are new sample assigned to cluster i, and the second two equations in equation number (12), represent a new sample that does not belong to cluster i, and τ measures the time constant of the exponential decay function (Specht, 1991).

$$\begin{aligned} A^i(k) &= \frac{\tau-1}{\tau}A^i(k-1) + \frac{1}{\tau}Y^j \\ B^i(k) &= \frac{\tau-1}{\tau}B^i(k-1) + \frac{1}{\tau} \\ A^i(k) &= \frac{\tau-1}{\tau}A^i(k-1) \\ B^i(k) &= \frac{\tau-1}{\tau}B^i(k-1) \end{aligned} \tag{12}$$

The equation (13), which has the equivalent estimator using city block distances, uses the same coefficient of equations (11) and (12) and is a more straightforward estimator for the model (Specht, 1991).

$$\hat{Y}(X) = \frac{\sum_{i=1}^{m} A^i\exp\left(-\frac{C_i}{\sigma}\right)}{\sum_{i=1}^{m} B^i\exp\left(-\frac{C_i}{\sigma}\right)} \tag{13}$$

Therefore, we use the estimator in equation (9), because the COVID-19 observations (X, Y) are almost small in this study. The algorithm of the proposed GRNN model in this study is as follows (Figure 2).

Algorithm: GRNN

Procedure
begin
initialise the COVID-19 Cases (Y), variables (X) and algorithm schemes (ID);
input:

X_1: Input temperature
X_2: Input absolute humidity
X_3: Input relative humidity
output:
Y: Output number of daily COVID-19 cases
function:
f = GRNN (inputs = X_1 X_2X_3), Y = output, weight = w^n, ε: Threshold parameter, $di = |f(x_i) - y_i|$ for $i = 1$ to n
*Apply robust fitting and find unit vector * size*
Update the f_{best}
Repeat the steps for the iteration n from 0
if $\left|w_i^n - w_i^{n+1}\right| < \varepsilon$ then
end of the iteration and then
break
else
Apply all the above processes
until the end of the iterations
end

4. Descriptive statistics analysis

The Tables 1–3, illustrate the descriptive analysis of the number of daily new cases, temperature, and absolute humidity, respectively, in the five selected European countries (Spain, France, Italy, Germany, and the U.K.) from March 1st to August 5th for 158 days. The data in the below tables includes mean, standard deviation, minimum, first quartile, median, third quartile, and maximum.

Table 1 shows a descriptive analysis of daily new cases. According to this table, France has the lowest mean "1312", and Spain has the highest mean "2085" of the daily number of new cases among these countries. Italy has the highest minimum "113," and the U. K. has the lowest minimum "5" number of the average number of daily infections among these countries. The highest maximum number of daily new cases belongs to Spain with "9181" cases, and on the other hand, the U.K. has the lowest maximum "5487" number of daily cases per day. We can observe that Italy has the lowest median number, "618" of daily new cases, which means half of the number of daily-infected people in Italy is less than this number.

Table 2 displays the statistics of temperature in those selected countries. According to the above table, Germany has the lowest maximum "24.39°C"; temperature, and Spain has the highest maximum "32.06°C" temperature. The lowest minimum "0.22° C" is recorded by Germany, and the highest minimum "5.44°C" is for Italy. Germany has the lowest mean "13.64°C", and Spain has the highest mean temperature "19.27°C". Thus Spain is the warmest country and Germany is the coldest one. In addition, we can observe that Germany has more cold days than the other European countries in this study because its median temperature is the lowest one, "14.28°C"

Table 3 depicts the statistics of absolute humidity in these selected European countries. Italy has the highest maximum "18.10," and on the other hand, Germany has the lowest minimum "2.35" absolute humidity among the countries in this study. Moreover, Germany has the lowest median "7.40", and Italy has the highest median "10.73" among the countries.

Figure 3, represents the bar graph related to the comparison of the number of daily infections among the most five affected European countries by COVID-19. Overall, these countries experienced an upward trend in daily infections in March. Spain has recorded the highest number of 9181 daily cases on the 27th of March.

We also observe that France, Italy, Germany, and the U.K. are positively skewed, which means the number of daily new cases in their country is decreasing, but the distribution of Spain is not like the other four countries in the study and daily new cases are increasing again.

Spain shows an increasing trend in the daily number of new cases in July, whereas France, Italy, Germany, and the U.K. have almost a decreasing curve trend for daily infections. The maximum number of daily cases for Spain, France, Italy, and Germany is around late March, but the peak for the U.K. is almost mid-April, which shows the trend of the spread of COVID-19 in Europe is from East to West.

Table 2. Descriptive analysis of the average daily temperature.

Temperature	Spain	France	Italy	Germany	United Kingdom
Mean	19.27	15.16	19.04	13.64	14.08
Standard Deviation	7.26	5.18	6.07	5.77	4.53
Minimum	4.22	3.78	5.44	0.22	5.33
First quartile	13	11.04	13.61	9.42	10
Median	19.08	16.33	19.72	14.28	14.83
Third quartile	26.25	18.81	24.43	18.22	17.39
Maximum	32.06	27.78	30.11	24.39	25.56

Table 1. Descriptive analysis of the number of daily new cases.

New cases	Spain	France	Italy	Germany	United Kingdom
Mean	2085	1312	1568	1342	1942
Standard Deviation	22,289	1332	1753	1608	1660
Minimum	28	30	113	18	5
First quartile	405	459	250	354	675
Median	1007	760	618	636	1111
Third quartile	3011	1623	2663	1435	3288
Maximum	9181	7578	6557	6294	5487

Table 3. Descriptive analysis of the daily absolute humidity.

Humidity	Spain	France	Italy	Germany	United Kingdom
Mean	8.05	8.22	10.44	7.72	8.03
Standard Deviation	1.57	2.36	2.97	2.93	2.23
Minimum	3.95	2.68	3.32	2.35	3.86
First quartile	7.13	6.71	8.35	5.54	6.49
Median	8.12	8.09	10.73	7.40	7.78
Third quartile	9.01	10.07	12.51	9.39	9.54
Maximum	11.93	14.12	18.10	15.69	13.01

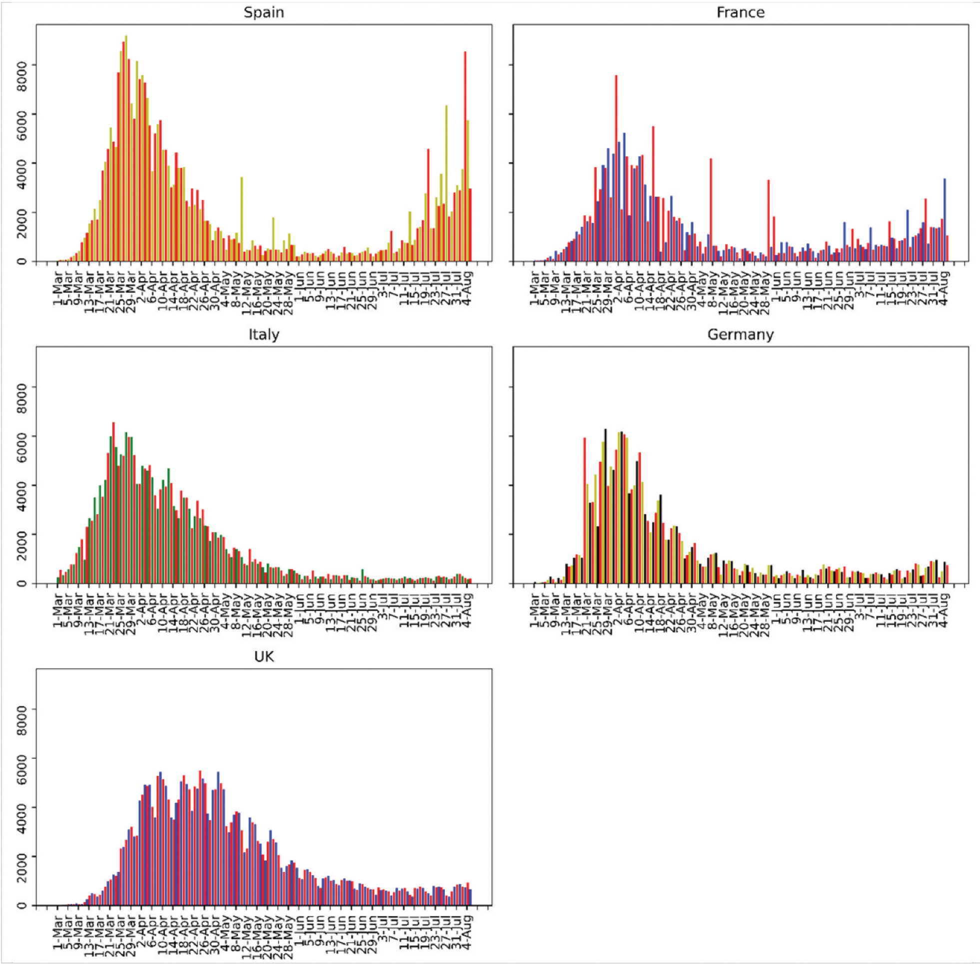

Figure 3. Bar graph comparison of daily new cases by country.

The bar charts in Figure 4 illustrate a comparison of temperature information among these European countries from March to the first week of August. By comparing these charts, we find that their average daily temperature is all negative skewed, which means that the temperature is gradually increasing in all these European countries in mid-April. The above charts show that Spain and Italy have almost a similar trend in daily temperature, and among these countries, Germany has a lower daily temperature on average.

We observe that the temperature is gradually increasing from March to August in these five countries, and by comparing Figures 4 and 3, we find that the overall number of daily infections started decreasing in this period. This observation leads us to hypothesise that by increasing the temperature, the number of daily infections would decrease. Nevertheless, we cannot verify this hypothesis based on the chart, and we need to test this relationship statistically, and in the next section, we will test this hypothesis statistically.

The bar charts Figure 5, provides some information about the comparison of absolute humidity among these European countries from March to the first week of August. The charts show that Italy has the highest level of absolute humidity, whereas Germany, and then the U.K., experiences the lowest level of absolute humidity among these European countries. In addition, the graphs show that the absolute humidity in all these five European countries is gradually increasing at the beginning of April.

On the other hand, if we look at Figure 3, we can find that the daily number of new COVID-19 started to decrease at this period. This observation leads us to hypothesise that by increasing the absolute humidity, the number of daily infections would decrease. Again, we cannot accept this hypothesis by looking at the graphs. We need to statistically test this hypothesis in the next section to find a possible relationship between them.

5. Poisson analysis

When a variable is countable, unimodal, and skewed, the Poisson distribution may be an acceptable fit in an analysis. The daily number of COVID-19 cases is a discrete and countable number with a daily interval. Therefore, the Poisson analysis may be appropriate for this type of data. In this study, we use the Poisson analysis to test the relationship between the temperature & absolute humidity and the daily number of new cases in five most affected counties by COVID-19 in

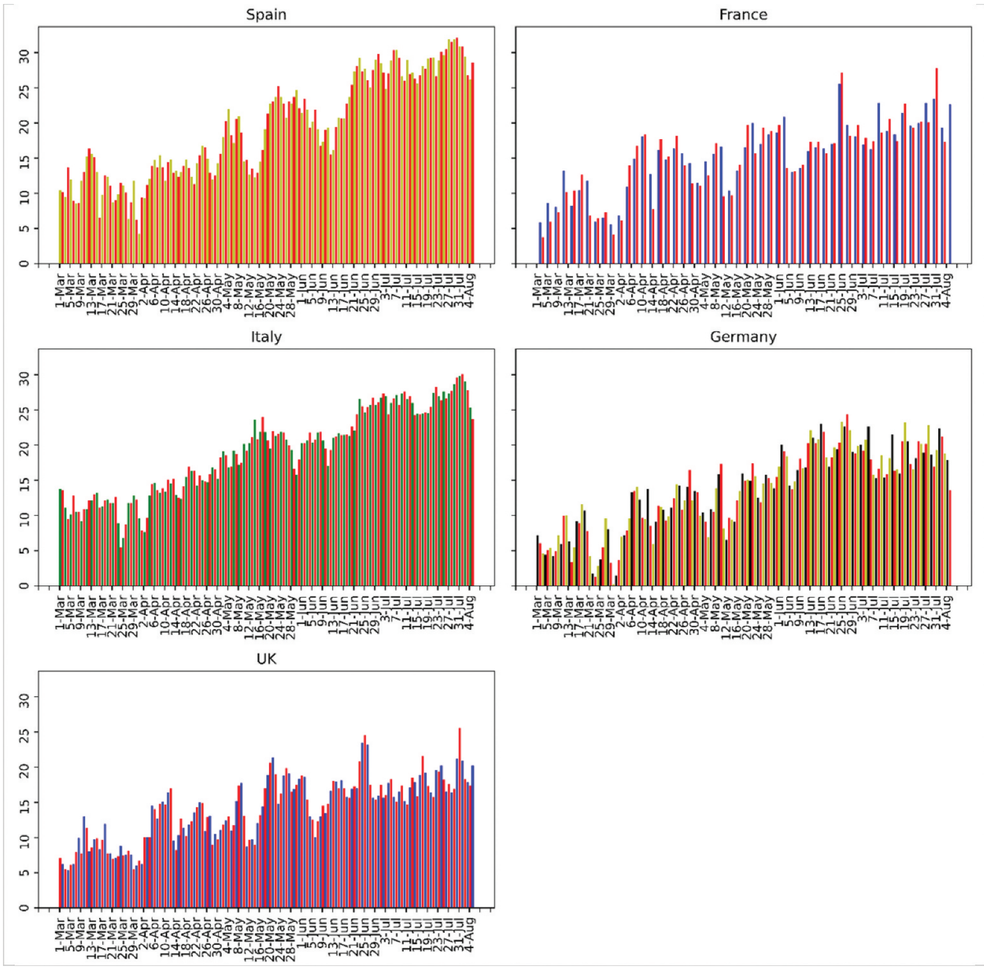

Figure 4. Bar graph comparison of temperature by country.

Europe: Spain, France, Italy, Germany, and the U.K. for 158 days from March 1st to August 5th.

Table 4 shows the Poisson results for Spain, France, Italy, Germany, and the U.K. According to this table, both independent variables, temperature, and absolute humidity for all these five European countries have a P-value of less than 0.0001. Therefore, we strongly reject the H0 and conclude that there is a relationship between the number of daily new cases in these European countries with the temperature and absolute humidity weather in their countries.

The results indicate that for every one-unit increase in temperature, the predicted log count of the number of cases in Spain decreases by 0.0394 units, in France increases by 0.065 units, in Italy decreases by 0.1705 units, in Germany increases by 0.0321, and in the U.K. increases by 0.092. Moreover, at the same time, for every one-unit increase in absolute humidity, the predicted log count of the number of new cases in Spain decreases by 0.1145 units, in France decreases by 0.3123 units, in Italy increases by 0.0407 units, in Germany decreases by 0.3617 units, and in the U.K., decreases by 0.3398 units.

In the above table, we find a negative relationship between the number of daily new cases and the temperature & absolute humidity in Spain. There is a weak positive relationship between the number of daily new cases & temperature and a negative relationship between the number of daily new cases & absolute humidity in France. There is a negative relationship between the number of daily new cases & temperature; on the other hand, there is a weak positive relationship between daily new cases & absolute humidity in Italy. There is a weak positive relationship between the number of daily new cases & temperature and a negative relationship between the number of daily new cases & absolute humidity in Germany. There is a weak positive relationship between the number of daily new cases & temperature, and there is a negative relationship between the number of daily new cases & absolute humidity in the U.K. Table 5 demonstrates these summaries.

Figure (6) shows the Poisson's goodness of fit lines that they fit with the original data in Spain, France, Italy, Germany, and the U.K. We can observe that these Poisson's fitted values align with the underlying data and follow the same trend. For Poisson's match metrics, the original data's relationship with temperature and absolute humidity also replicate.

After showing statistical analysis and finding a meaningful relationship between the daily new cases and the weather factors, in the next step, we are interested in analysing the trend of COVID-19 in these

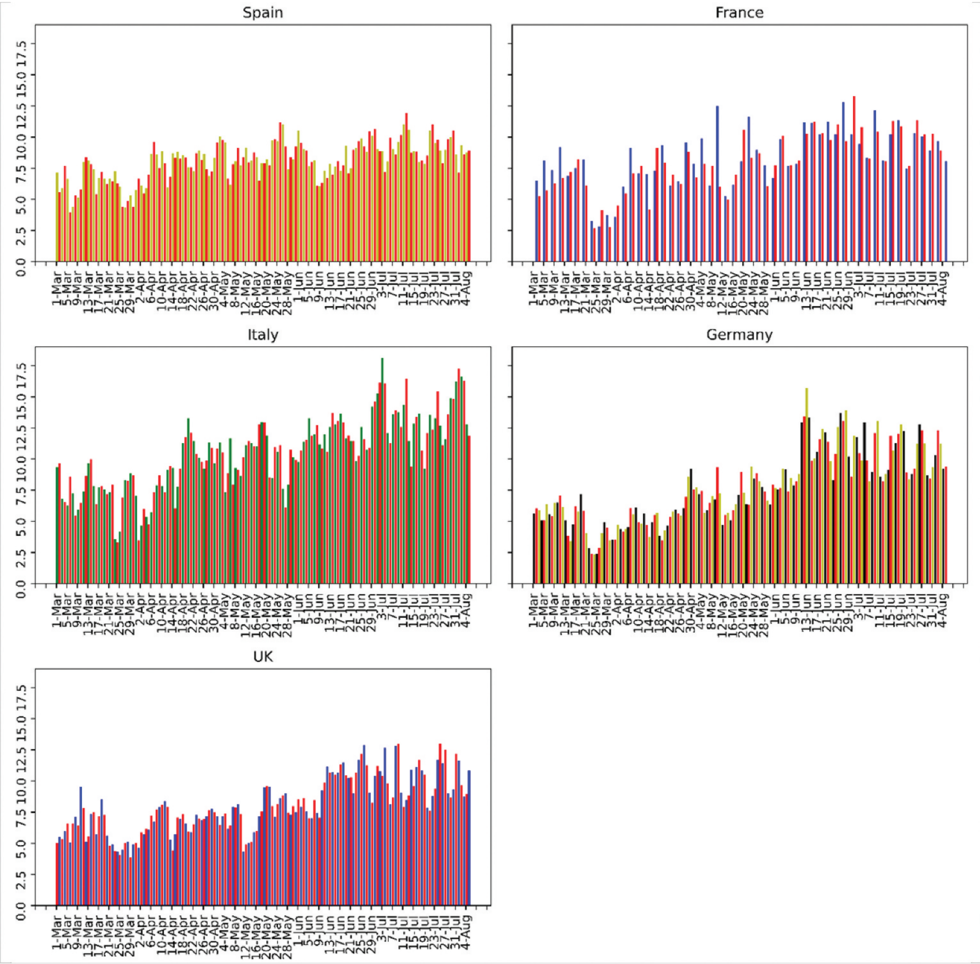

Figure 5. Bar graph comparison of absolute humidity by country.

Table 4. Poisson regression results for the number of daily new cases in Spain, France, Italy, Germany, and the UK.

Spain	Estimate	Standard Error	Z	P-value	95% Confidence Limits	
Intercept	9.2298	0.008	1091.696	<0.0001	9.213	9.246
Temperature	−0.0394	0.000	−109.484	<0.0001	−0.04	−0.039
Absolute Humidity	−0.1145	0.002	−75.054	<0.0001	−0.117	−0.111
France	**Estimate**	**Standard Error**	**Z**	**P-value**	**95% Confidence Limits**	
Intercept	8.6238	0.007	1202.506	<0.0001	8.61	8.638
Temperature	0.065	0.001	101.659	<0.0001	0.064	0.066
Absolute Humidity	−0.3123	0.002	−205.556	<0.0001	−0.315	−0.309
Italy	**Estimate**	**Standard Error**	**Z**	**P-value**	**95% Confidence Limits**	
Intercept	9.7661	0.006	1529.837	<0.0001	9.754	9.779
Temperature	−0.1705	0.001	−219.469	<0.0001	−0.172	−0.169
Absolute Humidity	0.0407	0.001	28.472	<0.0001	0.038	0.043
Germany	**Estimate**	**Standard Error**	**Z**	**P-value**	**95% Confidence Limits**	
Intercept	9.2119	0.006	1522.198	<0.0001	9.2	9.224
Temperature	0.0321	0.001	43.086	<0.0001	0.031	0.034
Absolute Humidity	−0.3617	0.002	−201.972	<0.0001	−0.365	−0.358
UK	**Estimate**	**Standard Error**	**Z**	**P-value**	**95% Confidence Limits**	
Intercept	8.9061	0.007	1291.287	<0.0001	8.893	8.92
Temperature	0.092	0.001	120.568	<0.0001	0.091	0.094
Absolute Humidity	−0.3398	0.002	−196.642	<0.0001	−0.343	−0.336

Table 5. Summary of the relationship between temperature & absolute humidity, and daily new cases.

Country	COVID-19	Temperature	Absolute Humidity
Spain	↑	↓	↓
France	↑	↑weak	↓
Italy	↑	↓	↑weak
Germany	↑	↑weak	↓
UK	↑	↑weak	↓

countries, and predicting the daily new cases with ANNs to find the best model for COVID-19.

6. Neural network analysis

In this study, after finding the statistical relationship between the daily number of new cases and weather

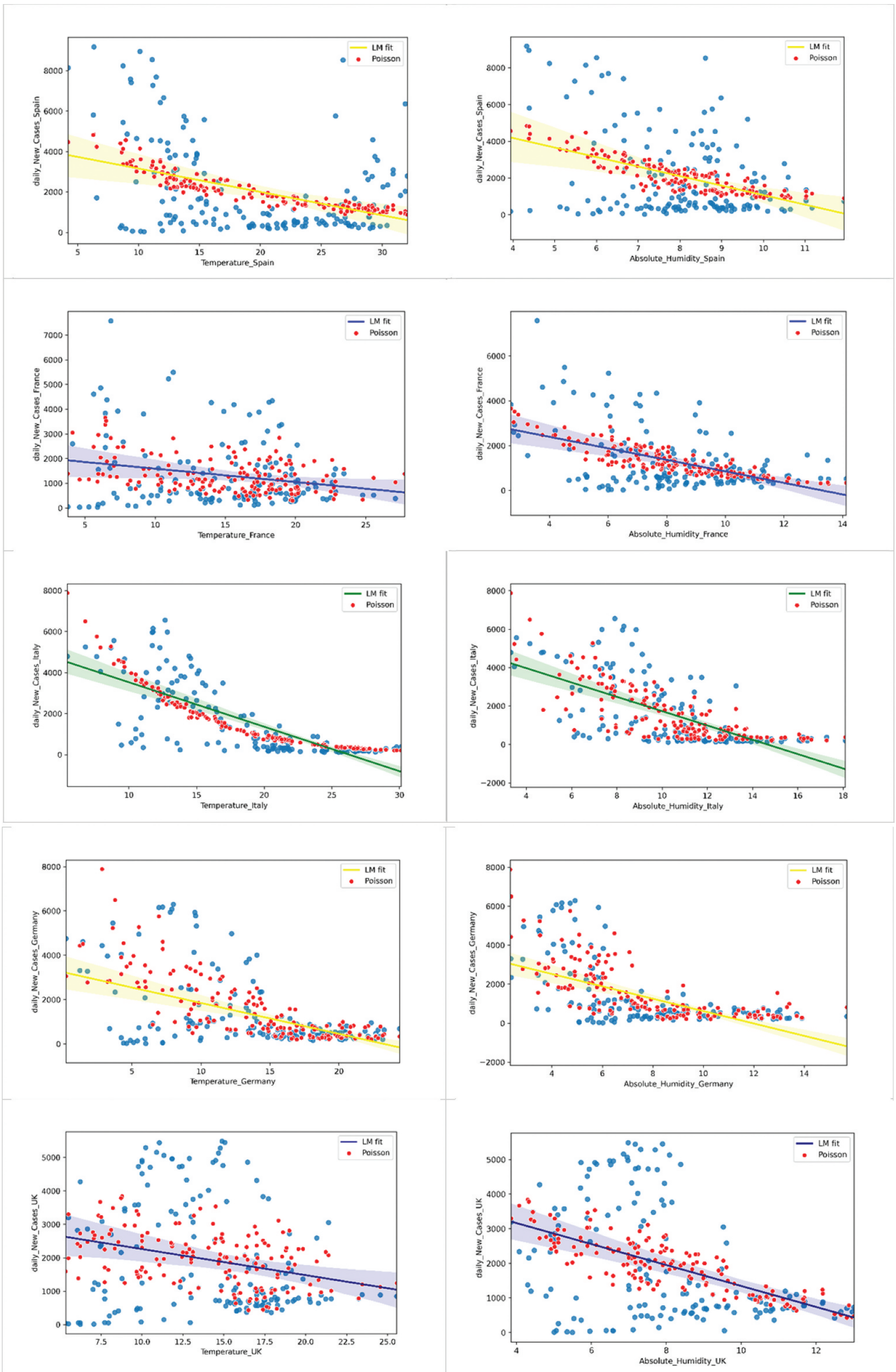

Figure 6. The goodness of fit plot for temperature and absolute humidity in Spain, France, Italy, Germany, and the UK.

factors, we also forecast the number of new cases in Europe's five most affected countries by neural network methods. The method that we implement in this research study is the GRNN, which uses a feedforward network with fast training ability (Bărbulescu, 2018; Majumder & Maity, 2018).

As we discussed above, some studies about SARS-CoV-1 (Bai & Jin, 2005; Hsieh et al., 2004; Lai, 2005; Wang & Ruan, 2004) found that this type of virus follows the Gaussian or Exponential distribution. The GRNN model in this study uses a Gaussian activation function in the hidden layer. Moreover, the structure of artificial networks is based on the statistics of data (Davenport, 2018), and the trend of the daily number of COVID-19 cases is like a discrete function with a specific number of cases, and in a particular period, which is daily. Therefore, we choose a feedforward neural network in our model and use the GRNN model to depict the trend of COVID-19.

Figure 7 displays the GRNN forecast trends for the five European countries. According to this Figure, the pattern of daily new cases in all five courtiers are almost like a bell curve, which first increasing and then decreasing, accordingly our proposed GRNN forecast model in Spain, France, Italy, Germany, and the U.K. are matched to the trend of actual numbers.

The overall prediction trends in all five countries follow the actual numbers, and it indicates that the GRNN is a promising prediction model for the COVID-19. In this study, we observe that the GRNN model shows more accurate capability with the actual data in Italy.

6.1. Validity of the model

In this section, we represent the actual & forecast numbers and the average forecast error of the proposed GRNN model for the last week of data in each of these five European countries: Spain, France, Italy, Germany, and the U.K.

Table 6 shows that the forecast number of daily new cases of these European countries is very close to the actual value of daily new cases. The proposed GRNN model shows an almost excellent fit model in the prediction of daily new cases. The only significant difference between actual and forecast numbers of COVID-19 for 35 days of forecast in the aforementioned European countries is on August 4 2020, for Spain, which has a sudden spike in the number of cases. In the next table, we show the average forecast errors of the model.

Table 7, shows the average forecast error of the last week of the COVID-19 cases, from July 30 2020, through August 5 2020. There is a rough rule of thumb that indicates %MAPE < 10% is an ideal forecast. The above table shows that the proposed GRNN model is a good fit model for almost all five European counties in the study. We would say that the model

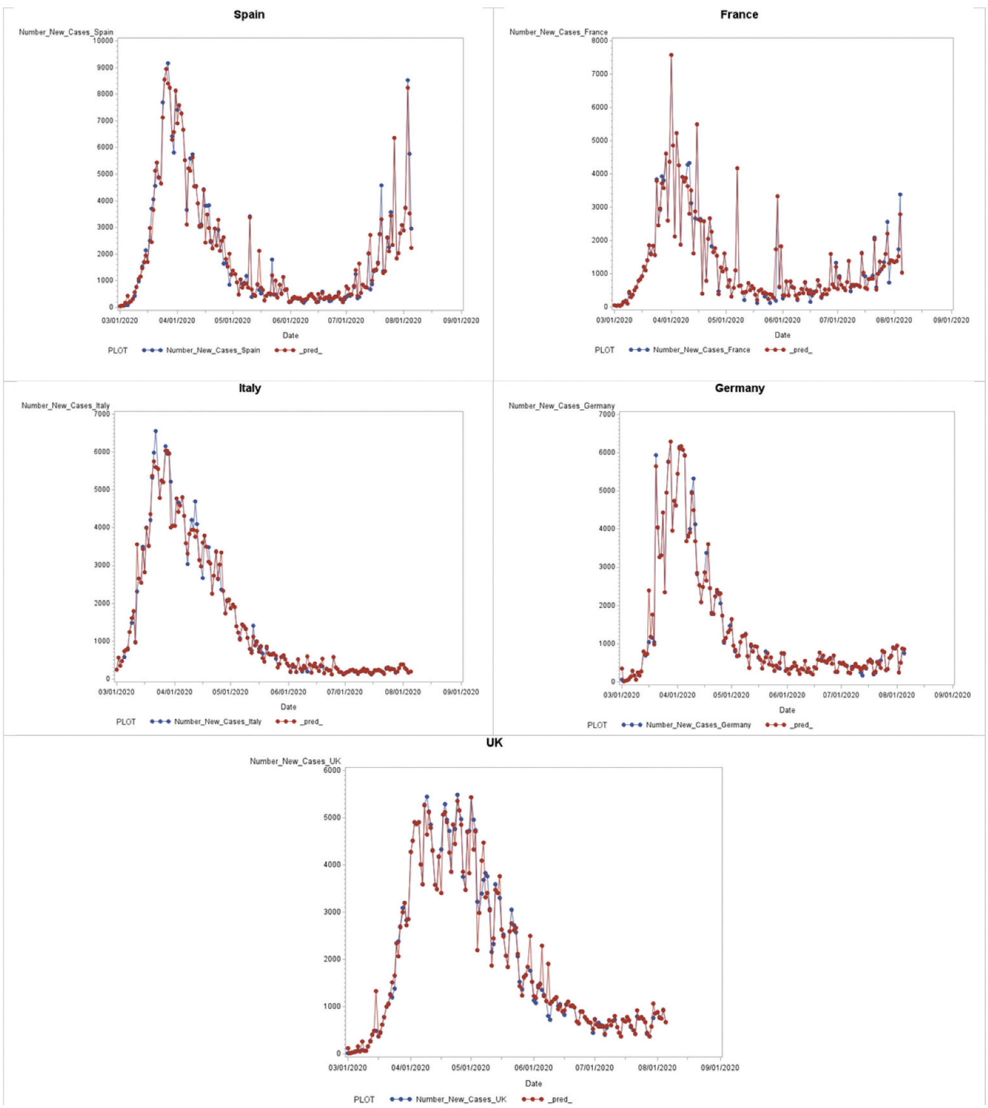

Figure 7. GRNN model prediction of the daily number of deaths for each country.

Table 6. Actual and forecast numbers of the proposed GRNN model.

	Spain		France		Italy		Germany		UK	
Country Date	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast
July-30-2020	2789	2789.00	1392	1392.00	288	288.00	902	879.54	763	1058.01
July-31-2020	3092	3092.00	1377	1377.00	382	382.00	870	869.96	846	846.00
August-1-2020	2896	2896.00	1346	1346.00	379	379.00	955	950.96	880	880.00
August-2-2020	3754	3726.03	1377	1377.00	295	295.00	240	240.76	771	764.26
August-3-2020	8532	8243.69	1742	1519.65	238	236.46	509	509.00	743	742.97
August-4-2020	5760	3528.97	3376	2791.87	159	160.01	879	871.28	928	924.51
August-5-2020	2953	2243.81	1039	1039.00	190	190.00	741	861.14	670	670.00

Table 7. Average forecast errors of the last week of COVID-19 cases for proposed GRNN model.

Country	Spain	France	Italy	Germany	UK
Mean Absolute Percentage Error (%MAPE)	10.93%	6.92%	0.03%	1.70%	5.08%

indicates very accurate forecasting results for Italy in this study.

7. Discussion & conclusion

In our study, we explored the relationship between the daily numbers of COVID-19 cases and weather factors. We conducted a Poisson regression analysis, and the findings support a statistically significant relationship between the daily number of new cases in Europe and the weather factors of temperature and absolute humidity.

Our findings reveal that for Spain, France, Germany, and the U.K., that an increase in absolute humidity is associated with a decrease in the number of daily new COVID-19 cases. In Italy, this trend is reversed, but it is fragile. Therefore, we can conclude that, in general, there is a negative relationship between the pattern of COVID-19 and the absolute humidity of weather. For instance, in the period between May 19 2020, and May 31 2020, the number of daily new cases in all five European countries is lower than in previous periods. The absolute humidity was a maximum during that May period.

Moreover, our results demonstrate that a decrease in Spain and Italy's temperature was associated with an increase in the number of daily COVID-19 cases. We see an inverse positive relationship for France, Germany, and the U.K. about temperature, but this positive relationship is very weak. An interesting observation that needs to be considered is that Spain and Italy were the first European countries with the most number of daily cases and deaths before the stay-at-home order. Therefore, we can conclude that the evidence generally supports a negative relationship between the temperature and the daily number of COVID-19 cases.

Some studies reveal that when temperature increases, the protein of COVID-19 may break down due to its heat intolerance, which is the result of the virus cover in a lipid bilayer. Similar to temperature, humidity is also a weather variable, which could affect the virus. This can be proven by the fact that once a pathogen is expelled from the respiratory tract by sneezing, it can stay in the air for a long period of time when humidity is low (Schoeman & Fielding, 2019).

The spread of SARS-CoV-2 is related to several factors. The results shown in some research studies demonstrate a relationship between COVID-19 and geographical and climate factors (Ficetola & Rubolini, 2020; Sajadi et al., 2020). Some studies show the role of absolute humidity and temperature (Gupta et al., 2020), relative humidity and temperature (Sajadi et al., 2020; Wang et al., 2020) in the number of daily new COVID-19 cases, but to our best knowledge, we can be the first group of scientists who have conducted a Poisson regression analysis to statistically substantiate this relationship.

These five countries in the study are among the top 10 countries in Europe with the largest total area, with France having the largest total area among these five countries in the study (United Nations website, 2020). According to our calculations, the number of COVID-19 cases per square kilometre in Spain, France, Italy, Germany, and the U.K. is around 0.5, 0.35, 0.8, 0.55, and 1.3, respectively, which indicates that France has a potential advantage over other European countries in this study in managing the COVID-19 pandemic among its population.

We should consider another item in our discussion: the median age of people in these European countries and its relationship with the number of daily new COVID-19 cases. Germany has the highest number of older people (World Health Organization, 2020), and also we found that it has the highest number of cold days among these five European countries in this study. Based on our calculation, France has the lowest number of COVID-19 cases per square kilometre, and it has the youngest median age after the U.K. among these European countries, for example, it has a difference of around seven years with Germany (World Health Organization, 2020). France is at an advantage with respect to its people's median age and the area of the entire land compared to other European countries in the study. Therefore, we can conclude that France may be able to control or reduce the mortality rate because it has a younger population, but controlling the spread of the disease may not be as easy. The U.K. has the youngest median age, but the number of COVID-19 cases per square kilometre in this country is the highest one among these European countries. The U.K. has the smallest total area among these countries (United Nations website, 2020).

In this study, we proposed a GRNN model to forecast the daily number of new COVID-19 cases in these five industrialised European countries, which are among the top European countries affected by COVID-19. Our proposed GRNN model shows a good convergence to the real data of COVID-19 in all countries in the study, and it shows a more accurate fit with the COVID-19 data in Italy. The trend of COVID-19 shows that the number of new cases is slowly increasing, especially in Spain; therefore, an accurate forecast model can help the governments of those aforementioned European countries to know better about the possible number of COVID-19 cases in the future. The impact of a precise forecast model's value can be more significant when our study shows a statistically negative relationship between the daily number of COVID-19 and the weather factors of absolute humidity and temperature. By moving from summer to winter gradually, the temperature and absolute humidity will decrease. Therefore, an accurate forecast model would be more necessary for the European governments to predict the possible number of COVID-19 cases in the future, helping the hospitals and healthcare in Europe provide enough equipment for their people.

In addition, the findings of this study may assist policymakers in Europe to understand the trend of the number of daily COVID-19 cases in their countries and to astutely monitor this trend to mitigate more disasters in the future. The proposed mathematical models in this study to forecast the number of new COVID-19 cases contribute to the literature on this new phenomenon and provide insightful associations that may support broad appropriate precautionary actions before finding a vaccine.

Acknowledgments

The authors are thankful to the respectful editors and three anonymous reviewers for their very constructive comments to improve the paper's quality. Moreover, Mahdi Fathi would like to thank Leon A. Kappelman, Professor of Information Systems & Chair, ITDS Department, Ryan College of Business, University of North Texas for his encouragement, mentoring, and support during this research.

Note

1. Machine learning and ANNs methods have been used to forecast typhoid fever (X. Zhang et al., 2013), haemorrhagic fever (W. Wu et al., 2015), hepatitis (Wei et al., 2016), dengue fever (Polwiang, 2020), and COVID-19 (Fong et al., 2020).

Highlights

- The trend of the daily number of new COVID-19 cases is almost decreasing in France, Italy, Germany, and the U. K. but this trend is increasing in Spain again.
- France has the advantage of controlling the COVID-19 mortality rate compared to other European countries in this study.
- There is a negative statistical relationship between the number of daily new COVID-19 cases and weather factors of absolute humidity and temperature in Western Europe.
- There is a statistically more reliable negative relationship between the number of daily new COVID-19 cases and absolute humidity than the temperature in Western Europe.
- GRNN shows a promising forecast model in the forecasting of the number of COVID-19 cases.

Disclosure statement

There is no conflict of interest by the authors.

ORCID

Mahdi Fathi http://orcid.org/0000-0003-3476-4722

References

Araujo, M. B., & Naimi, B. (2020). Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate. *medRxiv*. 2020.03.12.20034728. https://doi.org/10.1101/2020.03.12.20034728

Bai, Y., & Jin, Z. (2005). Prediction of SARS epidemic by BP neural networks with online prediction strategy. *Chaos, Solitons, and Fractals*, *26*(2), 559–569. https://doi.org/10.1016/j.chaos.2005.01.064

Bandyopadhyay, S. K., & Dutta, S. (2020). Machine learning approach for confirmation of covid-19 cases: Positive, negative, death and release. *medRxiv*.

Bărbulescu, A. (2018). Do the time series statistical properties influence the goodness of fit of GRNN models? Study on financial series. *Applied Stochastic Models in Business and Industry*, *34*(5), 586–596. https://doi.org/10.1002/asmb.2315

Bloom-Feshbach, K., Alonso, W. J., Charu, V., Tamerius, J., Simonsen, L., Miller, M. A., et al. (2013). Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): A global comparative review. *PLoS One*, *8*(2), 3–4. https://doi.org/10.1371/journal.pone.0054445

Bukhari, Q., & Jameel, Y. (2020). Will coronavirus pandemic diminish by summer? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3556998

Burrascano, P. (1991). Learning vector quantization for the probabilistic neural network. *IEEE Transactions Neural Networks*, *2*(4), 458–461. https://doi.org/10.1109/72.88165

Cacoullos, T. (1966). Estimation of a multivariate density. *Annals of the Institute of Statistical Mathematics*, *18*(2), 179–189. Tokyo. https://doi.org/10.1007/BF02869528

Caramelo, F., Ferreira, N., & Oliveiros, B. (2020). Estimation of risk factors for COVID-19 mortality Preliminary results. *Preprint at Med RXIV*. https://doi.org/10.1101/2020.02.24.20027268

Casanova, L. M., Jeon, S., Rutala, W. A., Weber, D. J., & Sobsey, M. D. (2010). Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Applied and Environmental Microbiology*, *76*(9), 2712–2717. https://doi.org/10.1128/AEM.02291-09

Chan, K. H., Peiris, J. S. M., Lam, S. Y., Poon, L. L. M., Yuen, K. Y., & Seto, W. H. (2011). The effects of temperature and relative humidity on the viability of the SARS Coronavirus. *Advances in Virology*, *2011*, 1–7. https://doi.org/10.1155/2011/734690

Chen, Z. L., Zhang, Q., Lu, Y., Guo, Z. M., Zhang, X., Zhang, W. J., et al. (2020). Distribution of the COVID-19 epidemic and correlation with population emigration from Wuhan, China. *Chinese Medical Journal*. PMID: 32118644.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, *368*(6489), 395–400.

Davenport, T. H. (2018). From analytics to artificial intelligence. *Journal of Business Analytics*, *1*(2), 73–80. https://doi.org/10.1080/2573234X.2018.1543535

Doremalen, N. V., Bushmaker, T., & Munster, V. J. (2013). Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. *Eurosurveillance*, *18*(38), pii=20590. https://doi.org/10.2807/1560-7917.ES2013.18.38.20590

Eslami, H., & Jalili, M. (2020). The role of environmental factors to transmission of SARS-CoV-2 (COVID-19). *AMB Express*, *10*(1). https://doi.org/10.1186/s13568-020-01028-0

Ficetola, G. F., & Rubolini, D. (2020). Climate affects global patterns of COVID-19 early outbreak dynamics. *MedRxiv*. https://doi.10.1101/2020.03.23.20040501

Fong, S. J., Li, G., Dey, N., Crespo, R. G., & Herrera-Viedma, E. (2020). Composite Monte Carlo decision making under high uncertainty of novel coronavirus epidemic using hybridized deep learning and fuzzy rule induction. *Applied Soft Computing*, *106282*. https://doi.org/10.1016/j.asoc.2020.106282

Gupta, S., Raghuwanshi, G. S., & Chanda, A. (2020). Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020. *Science of the Total Environment*, *728*, 138860. 1 August 2020. https://doi.org/10.1016/j.scitotenv.2020.138860

Gurney, K. (1997, August 5). *An introduction to neural networks*. UCL Press.

Herrmann, H., & Bucksch, H. (2014). Clausius-Clapeyron equation. *Dictionary Geotechnical Engineering/Wörterbuch GeoTechnik*. https://doi.10.1007/978-3-642-41714-6_ 32107.

Hopkins, J. (2020). *Track reported cases of COVID-19 Coronavirus resource center (WWW Document, online)*.

Hsieh, Y. H., Lee, J. Y., & Chang, H. L. (2004). SARS epidemiology modeling. *Emerging Infectious Diseases*, *10* (6), 1165–1167. https://doi.org/10.3201/eid1006.031023

Huang, C.-J., Chen, Y.-H., Ma, Y., & Kuo, P.-H. (2020). Multiple-input deep convolutional neural network model for covid-19 forecasting in china. *medRxiv*.

Ji, Y., Ma, Z., Peppelenbosch, M. P., & Pan, Q. (2020). Potential association between covid-19 mortality and health-care resource availability. *The Lancet Global Health*, *8*(4), e480. PMID: 32109372. https://doi.org/10.1016/S2214-109X(20)30068-1

Lai, D. (2005). Monitoring the SARS epidemic in china: A time series analysis. *Journal of Data Science*, *3*(3), 279–293.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*, *382*(13), 1199–1207. https://doi.org/10.1056/NEJMoa2001316

Lin, K., Fong, D. Y. T., Zhu, B., & Karlberg, J. (2006). Environmental factors on the SARS epidemic: Air temperature, passage of time and multiplicative effect of hospital infection. *Epidemiology and Infection*, *134*(2), 223–230. https://doi.org/10.1017/S0950268805005054

Lofgren, E., Fefferman, N. H., Naumov, Y. N., Gorski, J., & Naumova, E. N. (2007). Influenza seasonality: Underlying causes and modeling theories. *Journal of Virology*, *81*(11), 5429–5436. https://doi.org/10.1128/JVI.01680-06

Lowen, A. C., Mubareka, S., Tumpey, T. M., Garcia-Sastre, A., & Palese, P. (2006). The guinea pig as a transmission model for human influenza viruses. *Proceedings of the National academy of sciences of the United States of America*, 103: 9988–9992.

Lowen, A. C., Steel, J., Mubareka, S., & Palese, P. (2008). High temperature (30 degrees C) blocks aerosol but not contact transmission of influenza virus. *Journal of Virology*, *82*(11), 5650–5652. https://doi.org/10.1128/JVI.00325-08

Majumder, H., & Maity, K. (2018). Application of GRNN and multivariate hybrid approach to predict and optimize WEDM responses for Ni-Ti shape memory alloy. *Applied Soft Computing*, *70*, 665–679. https://doi.org/10.1016/j.asoc.2018.06.026

Munster, V. J., Koopmans, M., Doremalen, N. V., Riel, D. V., & Wit, E. D. (2020). A novel coronavirus emerging in China — Key Questions for Impact Assessment. *New England Journal of Medicine*, *382*(8), 692. https://doi.org/10.1056/NEJMp2000929

Oliveiros, B., Caramelo, L., Ferreira, N. C., & Caramelo, F. (2020). Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases. *medRxiv*. https://doi.org/10.1101/2020.03.05.20031872

Pal, R., Sekh, A. A., Kar, S., & Prasad, D. K. (2020). Neural network based country wise risk prediction of covid-19. *arXiv Preprint arXiv: 2004.00959*.

Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, *33*(3), 1065–1076. https://doi.org/10.1214/aoms/1177704472

Pica, N., & Bouvier, N. M. (2012). (Environmental factors affecting the transmission of respiratory viruses. *Current Opinion in Virology*, *2*(1), 90–95. https://doi.org/10.1016/j.coviro.2011.12.003

Polwiang, S. (2020). The time series seasonal patterns of dengue fever and associated weather variables in Bangkok (2003-2017). *BMC Infectious Diseases*, *20*(1), 208. https://doi.org/10.1186/s12879-020-4902-6

Riou, J., & Althaus, C. L. (2020). Pattern of early human-to-human transmission of Wuhan 2019-nCoV. https://doi.org/10.1101/2020.01.23.917351

Sajadi, M. M., Habibzadeh, P., Vintzileos, A., Shokouhi, S., Miralles-Wilhelm, F., & Amoroso, A. (2020). Temperature, humidity and latitude analysis to predict potential spread and seasonality for COVID-19 (March 5, 2020). *SSRN*. https://doi.org/http://dx.doi.10.2139/ssrn.3550308

Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: Current knowledge. *Virology Journal*, *16*(1). https://doi.org/10.1186/s12985-019-1182-0

Shaman, J., & Kohn, M. (2009). Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National academy of sciences of the United States of America*, 106: 3243–3248.

Specht, D. F. (1991). A general regression neural network. *IEEE Transactions on Neural Networks*, *2*(6), 568–576. https://doi.org/10.1109/72.97934

Tamerius, J., Nelson, M. I., Zhou, S. Z., Viboud, C., Miller, M. A., & Alonso, W. J. (2011). Global influenza seasonality: Reconciling patterns across temperate and tropical regions. *Environmental Health Perspectives*, *119*(4), 439–445. https://doi.org/10.1289/ehp.1002383

Tan, J., Mu, L., Huang, J., Yu, S., Chen, B., & Yin, J. (2005). An initial investigation of the association between the SARS outbreak and weather: With the view of the environmental temperature and its variation. *Journal of Epidemiology and Community Health (1979-)*, *59*(3), 186–192. https://doi.org/10.1136/jech.2004.020180

Tou, J. T., & Gonzalez, R. C. (1974). *Pattern recognition principles*. Addison-Wesley.

United Nations website. (2020). https://www.un.org/en/

Van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., et al. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England Journal of Medicine*, *382*(16), 1564–1567. https://doi.org/10.1056/NEJMc2004973

Wang, J., Tang, K., Feng, K., Lin, X., Lv, W., Chen, K., et al. (2020). High temperature and high humidity reduce the transmission of COVID-19. *SSRN*. https://doi.org/http://dx.doi.10.2139/ssrn.3551767

Wang, W., & Ruan, S. (2004). Simulating the SARS outbreak in Beijing with limited data. *Journal of Theoretical Biology*, *227*(3), 369–379. https://doi.org/10.1016/j.jtbi.2003.11.014

Wei, W., Jiang, J., Liang, H., Gao, L., Liang, B., Huang, J., et al. (2016). Application of a combined model with autoregressive integrated moving average (ARIMA) and generalized regression neural network (GRNN) in forecasting hepatitis incidence in Heng County, China. *PLoS One*, *11*(6), e0156768. https://doi.org/10.1371/journal.pone.0156768

World Health Organization. (2003). *Consensus document on the epidemiology of severe acute respiratory syndrome (SARS)*. https://who.int/csr/sars/en/WHOconsensus.pdf

World Health Organization. (2020). *Novel coronavirus (2019-nCoV) situation reports*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet*, *395*(10225), 689–697. https://doi.org/10.1016/S0140-6736(20)30260-9

Wu, W., Guo, J., An, S., Guan, P., Ren, Y., Xia, L., et al. (2015). Comparison of two hybrid models for forecasting the incidence of hemorrhagic fever with renal syndrome in Jiangsu Province, China. *PLoS One*, *10*, 1–13. https://doi.10.1371/journal. pone.0135492.

Zhang, R., Liu, H., Li, F., Zhang, B., Liu, Q., Li, X., et al. (2020). Transmission and epidemiological characteristics of Novel Coronavirus (2019-nCoV)-Infected Pneumonia(NCIP): Preliminaryevidence obtained in comparison with 2003-SARS. *MedRxiv*. https://doi.org/10.1101/2020.01.30.20019836

Zhang, X., Liu, Y., Yang, M., Zhang, T., Young, A. A., & Li, X. (2013). Comparative study of four time series methods in forecasting typhoid fever incidence in China. *PLoS One*, *8*. https://doi.org/10.1371/journal.pone.0063116

Zhao, J., Zhao, J., Legge, K., & Perlman, S. (2011). Age-related increases in PGD(2) expression impair respiratory DC migration, resulting in diminished T cell responses upon respiratory virus infection in mice. *Journal of Clinical Investigation*, *121*(12), 4921–4930. https://doi.org/10.1172/JCI59777

Zietz, M., & Tatonetti, N. (2020). Testing the association between blood type and COVID-19 infection, intubation, and death. *medRxiv Preprint*. https://doi.org/10.1101/2020.04.08.20058073