



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.

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## International Journal of Hygiene and Environmental Health

journal homepage: [www.elsevier.com/locate/ijheh](http://www.elsevier.com/locate/ijheh)

## Effects of climate variables on the COVID-19 outbreak in Spain

José María Loché Fernández-Ahúja, Juan Luis Fernández Martínez \*

Group of Inverse Problems, Optimization and Machine Learning, Department of Mathematics. University of Oviedo. C/ Federico García Lorca, 18. Oviedo 33007, Spain

## ARTICLE INFO

**Keywords:**  
 COVID-19  
 Climate variables  
 PCR  
 IgG  
 Correlation

## ABSTRACT

An outbreak of the novel COVID-19 virus occurred during February 2020 onwards in almost all the European countries, including Spain. This study covers the correlation found between weather variables (Maximum Temperature, Minimum Temperature, Mean Temperature, Atmospheric Pressure, Daily Rainfall, Daily Sun hours) and the coronavirus propagation in Spain. A strong relationship is found when correlating the virus spread to the mean temperature, minimum temperature, and atmospheric pressure in different Spanish provinces. In this analysis we have used the ratio of the PCR COVID-19 positives with respect to the population size. A linear regression model using the mean temperature is implemented. Moreover, an analysis of variance is used to confirm the influence of mean temperature on the spread of virus. As a second measurement of the COVID-19 outbreak we have used the results of the antibodies tests carried out in Spain that provide an estimation of the heard immunity achieved. Based on this analysis, an estimation of the asymptomatic population is performed. All these results exhibit significant correlation with weather variables. The most affected provinces were Soria, Segovia and Ciudad Real, which are the coldest. On the opposite side, places such as Southern Spain, the Balears, and Canary Islands showed a lower rate of spread. This might be related to the warmer climate and the insularity of these islands. Besides, the coastal influence and the daily sun hours might also influence the lower rates in the east and west regions in Spain. This analysis provides a deeper insight of the influence of weather variables onto the COVID-19 spread in Spain.

## 1. Introduction

This paper aims to understand the influence of the weather variables in the COVID-19 spread in Spain. Many studies have only focused on the importance of the density of population in major cities with high connectivity worldwide. Nevertheless, there are other factors that could influence the transmission. For instance, it has been seen in South Korea that only one COVID-19 infected person, known as “patient 31”, transmitted the virus to 4482 persons in the first stage of the pandemic (Shim et al., 2020). Therefore, the potential spread of the COVID might not be only related to the number of infected people and other factors might have a great influence in the strength of the transmission. Several weather variables have been mentioned among the various potential factors that might have played a crucial role in the spread of this virus. One of the first studies regarding the effect of the temperature showed that the COVID-19 survived better with mean temperatures in the range between  $-10$  and  $15$  °C (Bannister-Tyrrell, M et al., 2020). Moreover, a robust negative association with high temperature and high humidity as factors, that reduced the COVID-19 transmission, was found over for

Chinese and USA cities (Wang et al., 2020). Following these statements, neither extreme cold nor hot conditions would favour the virus transmission. Regarding humidity, for Iran cities seems it has a negative correlation with COVID-19 propagation (Ahmadi, M et al., 2020). However, the authors of that study say the correlations are not enough significant. Another study focused on the main Turkish cities concluded that the climate factors that best correlate with COVID-19 spread are the mean temperature on the day of positives measurement and the wind speed 14 days before the positive detection (Şahin, M., 2020). Nevertheless, this result should be interpreted with caution. Another study, focused on Brazil pointed that high mean temperatures (around  $27.5$  °C) favoured COVID-19 propagation (Auler, A.C. et al., 2020). The same authors pointed that other demographic and social-economic factors, such as health conditions, might be crucial. Beside Brazil is located in the tropical zone with temperatures that are approximately steady along the year. In addition, the distance among large metropolitan areas are considerable. In another study, concluded that mean and minimum temperatures correlated well with COVID-19 pandemic in New York (Farhan Bashir et al., 2020). Since the major affected cities in Spain are

\* Corresponding author.

E-mail addresses: [jmloche96@gmail.com](mailto:jmloche96@gmail.com) (J.M. Loché Fernández-Ahúja), [jlfm@uniovi.es](mailto:jlfm@uniovi.es) (J.L. Fernández Martínez).<https://doi.org/10.1016/j.ijheh.2021.113723>

Received 1 August 2020; Received in revised form 16 February 2021; Accepted 17 February 2021

Available online 27 February 2021

1438-4639/© 2021 Elsevier GmbH. All rights reserved.

not such air polluted the role of air quality has been not focused in this study. Regarding a high populated country as India, air temperature was mentioned as a non-significant factor for the virus spread, as well as the lockdowns were said to have not a correlation to the reduction of the virus spread (Gupta, A et al., 2020). However, the virus spread was reduced by a 45% in Italy during the lockdown (Gatto, M et al., 2020).

This paper analyses the influence of the main meteorological variables on the Covid spread, treating these variables as spatial-temporal processes, analyzing their correlations through time and space in the Iberian Peninsula. Although the methodology that we used is well-known in applied mathematics and statistics, this serves to clarify and shed light on the importance of meteorological variables in the spread of the pandemic in all the Spanish provinces.

## 2. Methodology and data

### 2.1. Data sources

Climate variables values were obtained from the Spanish Agency of Meteorology (AEMET) platform OpenData (AEMET OpenData). Appendix I shows the list of used meteorological stations. The range of dates covers from 31st January to 20th June. The percentiles for each climate variable were computed for each meteorological station representative of each Spanish province. COVID-19 positives data was collected from Carlos III, 2020 Institute website. Spain has notified 237 096 COVID-19 cases and has conducted 2 536 234 COVID-19 PCR Tests by the May 28, 2020. At the start of the pandemic, Spain notified 2968 COVID-19 cases by the March 12, 2020. The distribution of PCR positives is not uniform along the country. Finally, the population data is obtained from the Spanish National Institute of Statistics (INE, 2020).

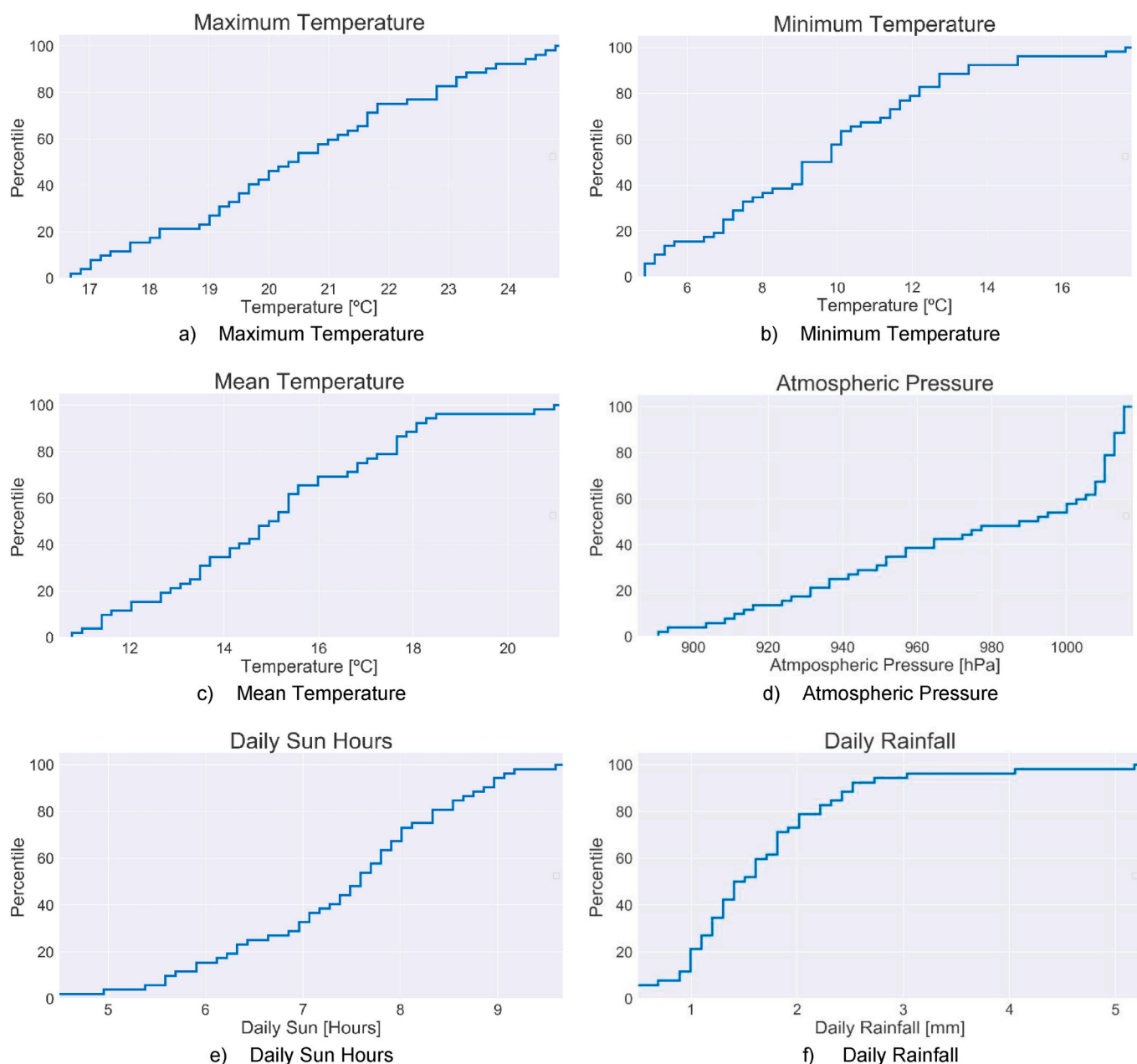


Fig. 1. Cdf (percentile curves) of the Climate Variables.

## 2.2. Climate variables

Six climate variables are used in this project: Maximum Temperature, Minimum Temperature, Mean Temperature, Atmospheric Pressure, Daily Sun Hours and Daily Rainfall. All these variables are obtained from AEMET for each Spanish province.

- Maximum, Minimum and Mean Temperatures are expressed in Celsius degrees [°C].
- Atmospheric Pressure is expressed in hectopascals [hPa] and is the mean of the minimum and maximum atmospheric pressures during the day.
- Daily sun hours are obtained directly from insolation.
- Daily rainfall is expressed in millimeters [mm] and it is the total rain gathered during the day.

Data about wind speed is available, although wind direction data is not available. From our point of view, wind direction data is quite as important as wind speed data. Hence, if both data are not available, we decided using only one is not a right approach. Another reason is that wearing a mask outside is compulsory. Therefore, most of the virus transmission happens in indoor scenarios, in which wind speed is not as important. We also think that in an urban scenario, the wind speed and direction can be modified by the buildings or the factories. Fig. 1 shows the Cumulative Distribution Function (CDF) of these climate variables. Table 1 shows the percentiles of these variables. Appendix II also shows the descriptive analysis of the climate variables for all the Spanish provinces.

## 2.3. COVID-19 variables

The cumulative number of COVID-19 tests carried by each Autonomous Community shows the total number of PCR tests done, positive or negative. It is estimated as a “density of tests” (*Testd*), defined as follows:

$$Testd = \frac{PCR}{Total\ Population} \tag{1}$$

The cumulative number of COVID-19 positives by province is taken from the data provided by the organization responsible for Carlos III, 2020 Spanish Institute, which is responsible to gather and cure this data. To estimate a “density of positives”, the population data from each province is needed.

The density of positives by province (*PCRd*) is defined as follows:

$$PCRd = \frac{PCR+}{Total\ Population} \tag{2}$$

which gives the ratio of the PCR positive over the total population.

To estimate a “daily density of positives”, the population data from each province is needed. The density of positives by province (*Daily PCRd*) is defined as follows:

$$Daily\ PCRd = \frac{PCR + (in\ a\ single\ day)}{Total\ Population} \tag{3}$$

**Table 1**  
Percentiles 10, 25, 50, 75 and 90 of the climate variables.

Variable	P10	P25	P50	P75	P90
Maximum Temperature [°C]	17.48	19.13	20.54	22.06	23.72
Minimum Temperature [°C]	5.44	7.26	9.62	11.86	13.45
Mean Temperature [°C]	11.74	13.48	15.15	17.00	18.22
Atmospheric Pressure [hPa]	915.34	941.26	991.03	1011.66	1015.40
Daily Sun Hours [h]	5.77	6.62	7.63	8.18	8.89
Daily Rainfall [mm]	0.99	1.12	1.53	2.05	2.55

which gives the ratio of the number of PCR positive in a single day over the total population.

To estimate a “daily increment of the density of positives”, the population data from each province is needed and the density of positives by province (*Daily PCRd*) for two consecutive days is required. The density of positives by province ( $\Delta Daily\ PCRd$ ) is defined as follows:

$$\Delta\ Daily\ PCRd = \frac{PCR + (in\ a\ single\ day) - PCR + (the\ day\ before)}{Total\ Population} \tag{4}$$

which gives the evolution of the daily density of PCR positives.

The Spanish Ministry of Health carried out a prevalence of COVID-19 IgG Antibodies in all the provinces and the autonomous cities (Ceuta and Melilla). The result is an estimation of the percentage of inhabitants of each province who could have been infected. Three rounds were done to obtain the percentage of population for each province who had been infected (*IgGd* [%]). Theoretically, the *IgGd* is computed following equation (2). The *IgGd* [%] values are given in Appendix III.

$$IgGd = \frac{IgG+}{Total\ Population} \tag{5}$$

By using the data from the COVID-19 density of positives and the results from the IgG antibodies research study, an estimation of the percentage of asymptomatic population can be computed. Two assumptions are made. First, it is assumed all people who were in serious or critical conditions were tested with a PCR. Second, it is assumed all the infected people has generated IgG antibodies after recovering from the virus. Therefore, the percentage of asymptomatic population (*APr*) is the percentage of population with antibodies minus the percentage of positives:

$$APr[\%] = IgGd[\%] - PCRd[\%] \tag{6}$$

These metrics (density of tests, density of positives, increment of the density of positives, density of antibodies and percentage of asymptomatic population) provide different insights about the impact of COVID-19 in each Spanish province. The aim is to avoid any type of bias due to the number of inhabitants.

To reduce the impact of noise in data, we have computed for each day the moving averages for five days for each of these variables. Therefore, the correlation is performed between smoothed versions of these temporal series. This is particularly important for *PCRd*, since this kind of information is sometime delayed by the public health authorities. Then, the correlation between both temporal series is computed via the Pearson’s and Spearman’s correlation coefficients. That way is possible to produce regionalized variables using these coefficients for all the provinces.

## 3. Results

This section shows the results obtained in this study. First, the correlation between the COVID-19 density of positives *PCRd* and the climate variables is found, in subsection 3.1. Second, a linear regression model using the mean temperature as independent variable and the COVID-19 density of positives *PCRd* as dependant variable is presented, in subsection 3.2. Third, an Analysis of the Variance (ANOVA) between the mean temperature and the daily density of new COVID-19 positives. *Daily PCRd* is computed in subsection 3.3. Fourth, the results from the massive antibodies study *IgGd* carried out in Spain are correlated with the climate variables in subsection 3.4. An estimation of the percentage of Asymptomatic Population *APr*[%] is computed, and the correlation of this result with climate variables is also tested, in subsection 3.5. Last, a comparison between the *PCRd* and *Testd* is shown for each Autonomous Community in subsection 3.6.

These variables are considered spatial-temporal processes  $X(t, x)$ , where  $x$  stands for the spatial (or geographic) coordinates that depend on each province and  $t$  for the temporal dependency. Fixing a time  $t_0$ ,

$X(t_0, x)$  is a regionalized variable. In this case  $X(t_0, x)$  is a discrete sample of dimension 52, which is the total number of provinces and autonomous cities in Spain. On the other side fixing  $x$ , provides a temporal series  $X(t, x_0)$  that provides the temporal evolution of  $X(t)$  in the province  $x_0$ . Both aspects (spatial and temporal) are important in understanding the Covid-19 spread.

Summarizing, we have concluded that the mean and minimum temperatures are correlated to the density of PCR positives for Spanish provinces. Besides, the maximum temperature has a lower correlation than the minimum temperature. Therefore, the areas where the pandemic has been stronger are the cold regions of the Spanish “Meseta” (plateau) that are characterized by a continental climate. This fact has been also observed in New York (Farhan Bashir et al., 2020). Nevertheless, in the New York study there was not any evidence that the warm weather could suppress COVID-19 transmission. This paper shares this conclusion: climate plays a very important role in the COVID-19 transmission, mainly in the cold areas, reducing or amplifying its spread, but it cannot eliminate the outbreak, as it has been observed lately on the second wave, when mobility restrictions have been abandoned. Besides, we have also observed that atmospheric pressure has a negative correlation with the density of COVID-19 positives in all the Spanish provinces. Low atmospheric pressure is a quite significant factor and could be one of the most related with the virus transmission. In Spain the low pressures are typical of spring and autumn due to the Azorean anticyclone moving intermittently south. In winter high pressures predominate because of marine polar anticyclones, the Scandinavian anticyclone and its connection with that of the Azores. Finally, in summer the weather is controlled by the Azorean anticyclone again, resulting in warm, dry and stable weather with low surface pressures that reverse with altitude and generate summer storms. Therefore, the climate in Spain is very different from the rest of Europe, and is dominated by low pressure throughout the year, which would favour the spread of the COVID virus. We have also analyzed the correlation between PCR positives and the antigen tests, showing a very important correlation between the PCR tests that originally were only performed on people with hospital symptoms and the number of people with antibodies, which also takes into account those who are asymptomatic. Basically, the count is simple: the number of infected is approximately six times the number of PCR positive tests. This would explain partly the higher rates of mortality observed in Spain referred to the PCR tests with respect to other countries that initially made more massive use of rapid tests. This situation would also explain the higher number of positives observed in the second wave because of the greater mobility and increased testing of the population.

In conclusion, the climate variables can have a great influence in the propagation of the COVID virus and could be taken into account in the process of public health decision making to define the areas with a higher potential risk of COVID transmission.

### 3.1. COVID-19 density of positives analysis

A correlation analysis was performed among the different climate variables and the density of COVID-19 positives (PCRd). Two correlation coefficients were used: the Pearson’s and the Spearman’s correlation coefficients. Pearson’s coefficient measures the degree of the linear correlation between variable, whereas Spearman’s coefficient is a measure of the rank correlation. It serves to assess how well the relationship between both temporal series can be described using a monotonic function (linear or not).

For two series  $X_1, X_2$  of size  $n$ , both coefficients are defined as follows:

$$\rho_p = \frac{cov(X_1, X_2)}{\sigma_{X1} \sigma_{X2}} \tag{7}$$

$$\rho_s = \frac{cov(r(X_1), r(X_2))}{\sigma_{r1} \sigma_{r2}} \tag{8}$$

The Pearson coefficient  $\rho_p$  is related to the slope of the regression line between these variables ( $X_1, X_2$ ). Spearman’s correlation coefficient is the Pearson coefficient between the corresponding rank variables. Ranking is the data transformation in which the numerical values of  $X_1$  and  $X_2$  are replaced by their rank ( $r(X_1), r(X_2)$ ), when the data are sorted.

#### • Modelling as regionalized variables: spatial variability

Table 2 shows the relationship between the climate variables and the density of COVID-19 positives on 1st June when most of the pandemic had already taken place. The correlations with the different temperatures, the atmospheric pressure and the daily sun hours are negative. Besides, the correlations with mean, minimum temperatures, and atmospheric pressure are the most important. The correlation with the daily sun hours is very low. The analysis shows no correlation between the density of COVID-19 positives and daily rainfall. Daily sun hours and daily rainfall seem to be non-significant factors in the COVID-19 spread.

Fig. 2 shows the evolution of the correlation of the PCRd values with the climate variables, starting from 12th March. In this case for each date, the correlation between the corresponding spatial variables is performed. The minimum temperature (Fig. 2b) has a stronger correlation with the density of COVID-19 positives PCRd, than the maximum temperature (Fig. 2a). Mean temperature (Fig. 2c) is the temperature which best correlates with PCRd. The correlation is always negative and decreases with time from - 0.66 to - 0.78. Fig. 2d shows the correlation with the mean range atmospheric pressure, while Fig. 2e and f shows the evolution for the daily sun hours and rainfall. The conclusions are similar to those previously commented for the 1st June, that is, the analysis is consistent along the duration of the outbreak. Particularly, in the case of the atmospheric pressure, both correlation coefficients show similar and very close trajectories.

The virus propagation is also a temporal process  $X(t, x_0)$  when the spatial variable (province) is fixed. To show a detailed view of the Spain geography, Fig. 3 shows a map recorded using QGIS and Natural Earth Data. The map features latitude and longitude coordinates, a scale bar, and a north arrow. Then, the correlation between the climate variables and the density of daily positives can be computed for each province. For that purpose, first the right time span to analyze the outbreak must be selected. The analysis was performed between 11th March to 20th May where most of the infections have occurred. Fig. 4 shows the absolute value of the correlation between PCRd and the maximum, minimum and mean temperatures, using maps created with the online tool Mapchart (Mapchart). The daily increment of the density of positives has a negative correlation with the maximum, minimum and mean temperatures for all provinces. Pearson and Spearman’s correlation coefficients show the same tendency, with a few changes in the value of the correlation in some cases. Catalonia, La Rioja, two provinces of Comunidad Valenciana (Valencia and Castellón) and in some cases Madrid are strongly negative

**Table 2**  
Correlation coefficients (climate variables, density of positives).

Variable	Correlation Coefficient with the Density of Positives (June 1st)			
	Pearson’s		Spearman’s	
	Value	p-value	Value	p-value
<b>Maximum Temperature</b>	- 0.57	$1.14 \times 10^{-5}$	- 0.67	$6.09 \times 10^{-8}$
<b>Minimum Temperature</b>	- 0.62	$1.06 \times 10^{-6}$	- 0.75	$1.10 \times 10^{-10}$
<b>Mean Temperature</b>	- 0.65	$1.65 \times 10^{-7}$	- 0.78	$10^{-11}$
<b>Atmospheric Pressure</b>	- 0.70	$6.14 \times 10^{-9}$	- 0.73	$6.3 \times 10^{-10}$
<b>Daily Sun Hours</b>	- 0.30	0.03	- 0.42	$1.78 \times 10^{-3}$
<b>Daily Rainfall</b>	0.06	0.69	0.21	0.14

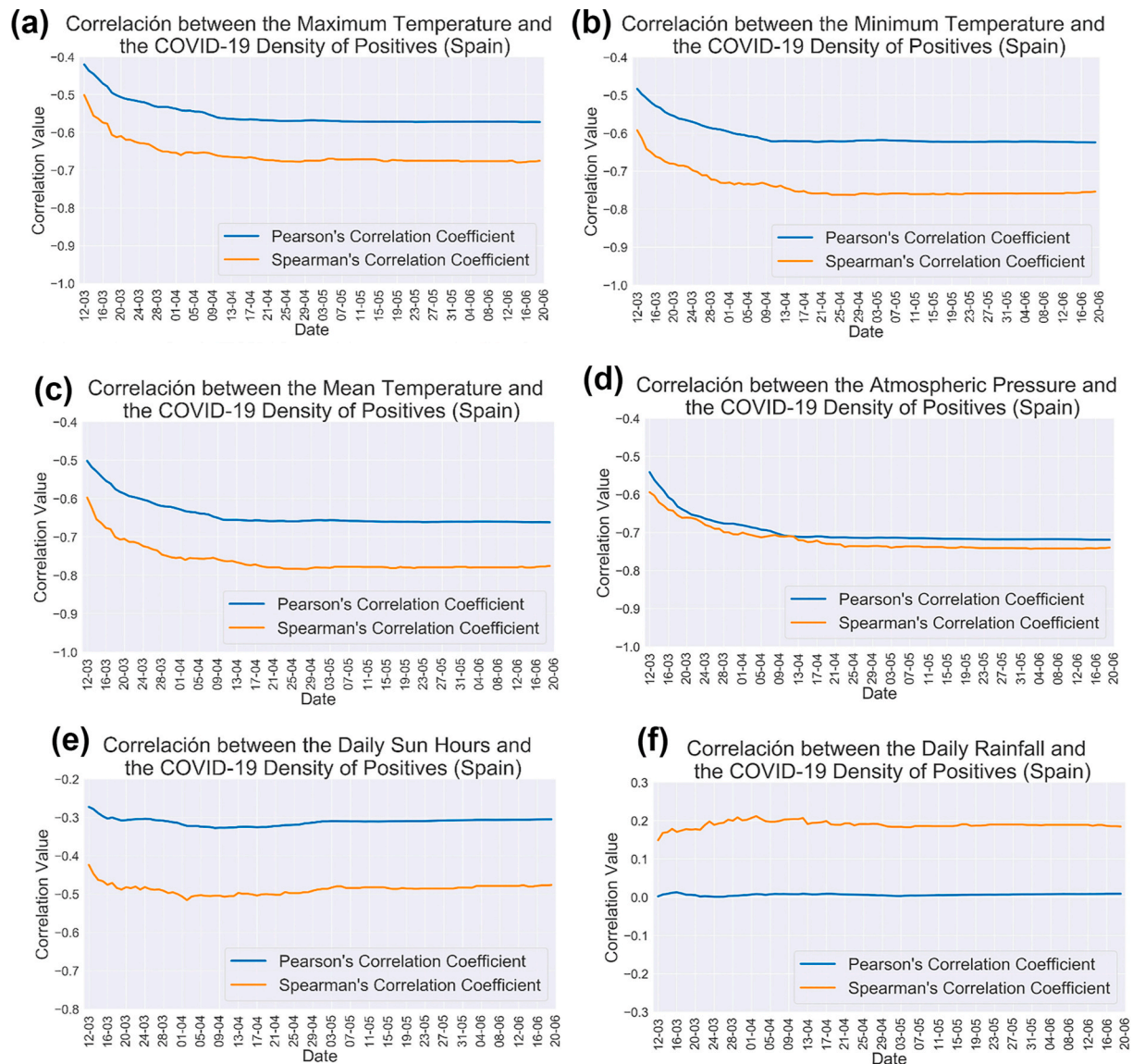


Fig. 2. Temporal evolution of the Correlation coefficients between the climate variables and the COVID-19 Density of Positives (PCRd).

• Modelling as temporal series

correlated with the daily increment of the density of positives, with absolute values higher than 0.7. However, most places of Northwest Spain are less strongly correlated, showing absolute correlation values lower than 0.5. Southern Spain, Ceuta, Melilla, and the Canary Islands show an absolute correlation coefficient value close to 0.6. Therefore, in the places that show strong correlated results it is easier to predict the daily number of cases based on their mean temperature.

The same analysis was performed to study the correlation with the temporal evolution of the atmospheric pressure. The atmospheric pressure has a high dependence on the terrain elevation. Given the terrain elevation is a constant value over time, the results of the temporal analysis for all provinces are close to 0, some of them positive and others negative. In other words, atmospheric pressure does not influence the daily evolution of the pandemic in a province but can affect the differences in the number of positives observed among different provinces.

The main reason for using as a metric for measuring the virus outbreak the daily increment of positives and not the daily number of positives is that this measure shows more information. The number of positives only shows the virus impact during a day, but the daily increment also shows if the virus spread is increasing or decreasing.

3.2. Linear regression model (mean temperature and the COVID-19 density of positives PCRd)

Mean temperature is the most correlated variable with the density of positives, according to the Spearman's correlation coefficient (Table 2). Thus, a linear regression model, using the density of positives metric PCRd as the dependent variable can be constructed:

$$PCRd = a_0 + a_1 \times T_{mean} \tag{9}$$

Fig. 5 shows the correlation between COVID-19 density of positives metric PCRd and mean temperature  $T_{mean}$  for two weeks timestep. In this plot the atmospheric pressure variable is only used to colour the points, but not for the linear regression. The correlation is plotted for the March 15th, April 1st, April 15th, May 1st, and May 15th. The negative correlation between the density of positives and the mean temperature is found in all the plots. Table 3 shows the linear equation and the RMSE error for each regression. Hence, a coefficient which enables us to establish a relationship between temperature and the COVID-19 spread was found. The slope is always negative and is closer to zero at the beginning of the outbreak. From April 15th, the slope value stabilizes

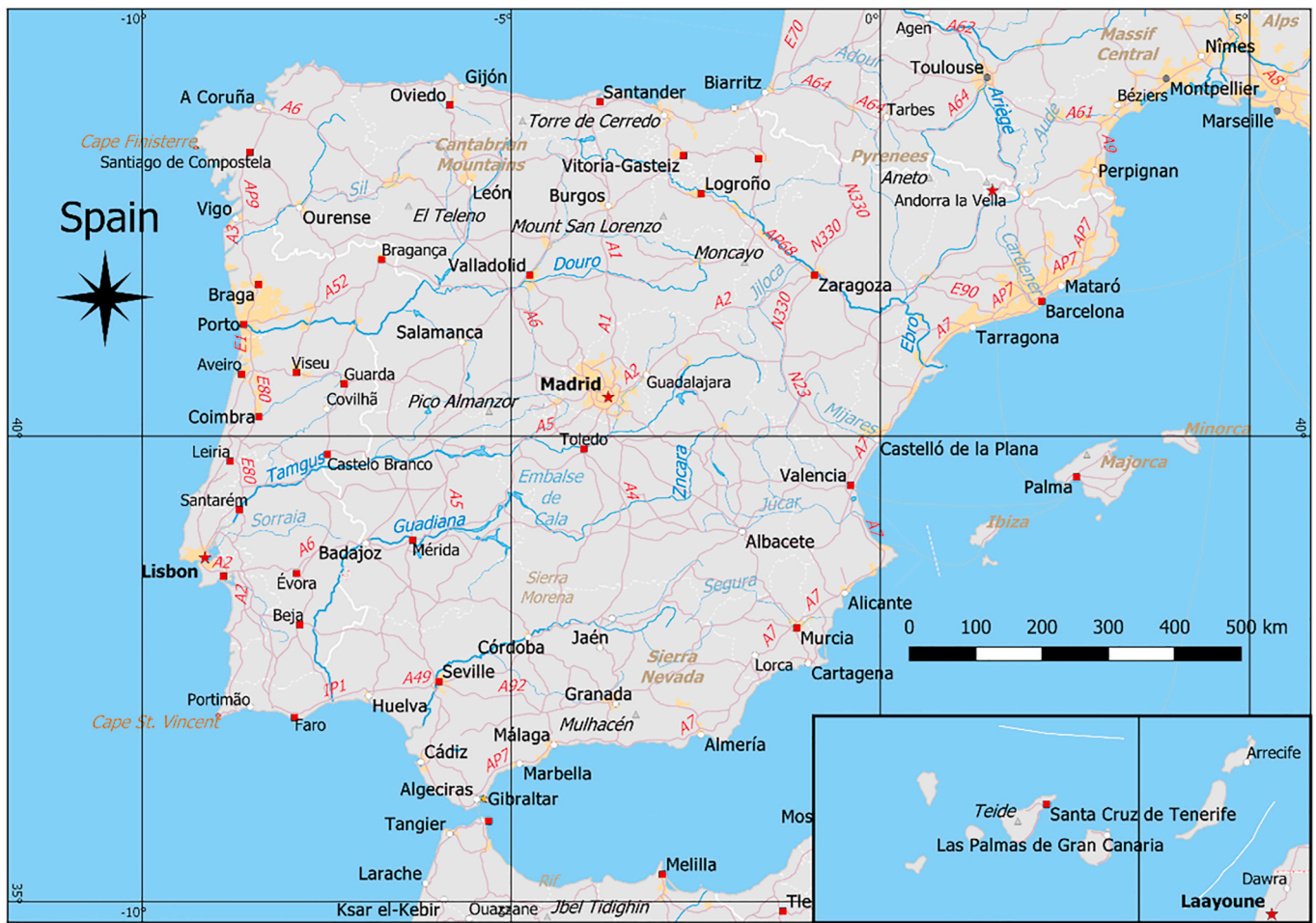


Fig. 3. Spain Map (it includes a North arrow on the left side and a scale bar on the right side).

between - 0.0012 and - 0.0014, that is, a decrease between 1.2 and 1.4 new PCRs per 1000 inhabitants and every degree of increase in the average temperature. The RMSE error is also very low in all the cases.

### 3.3. ANOVA (mean temperature and the daily COVID-19 density of positives DailyPCRd)

ANOVA serves to demonstrate that the mean temperature is correlated with the COVID-19 propagation. The virus expands in a similar way in two different cities, provided the daily mean temperature values are the same. As the reader can expect, there is always an error margin.

In this analysis, we have taken all the daily data of mean temperatures and density of COVID-19 positives between the 12th March and 20th June. To reduce the measurement error, the average mean of 10 days is computed for both the mean temperature and the daily new density of COVID-19 positives.

Fig. 6 shows a plot of the daily density of positives and the moving average of the mean temperature. To produce this plot, the mean temperature varies from 11.74 °C (P10) to 18.22 °C (P90). P10 and P90 were taken from Table 1. Within this range with step of 0.1 °C, the Daily PCRd was found for all the days between the 12th March and 20th June. For each temperature we scan the days were this temperature has occurred and we take the mean of the corresponding Daily PCRds. An average of these values along all provinces is finally taking place. Before doing these calculations, a moving average of 10 days of the daily mean temperatures and PCRd is performed to smooth the observed data and enforcing the dynamics of the virus propagation.

This figure shows a negative correlation between mean temperature

and COVID-19 density of positives (independently of the provinces). The equation of the linear regression is:

$$Daily\ PCRd = 1.517 \times 10^{-4} - 8.329 \times 10^{-6} \times Mean\ Temperature\ [^{\circ}C] \tag{10}$$

that is, a decrease of round 8.3 cases by million every degree of increment in the mean temperature.

For example, 0.00005 ( $5 \times 10^{-5}$ ) Daily PCRd means 50 new daily positive cases that for 1 million habitants. Therefore, a decrease of 10° in the mean temperature during the winter implies around 83 new daily positive cases for 1 million habitants. This might be one the factors of the increase of cases in January 2021 during the impact of Filomena storm in the Spanish peninsula. Obviously, the mobility and the social networks are also very important in the dynamic of COVID spread.

For the ANOVA analysis, the null hypothesis  $H_0$  is that the mean temperature has not an effect in the COVID-19 propagation. The result of the ANOVA test is 11.60, with 51 (number of provinces-1) and 1661 (number of data – number of provinces) as degrees of freedom. The zero values in the data (temperatures in which there are not notified COVID-19 positives) are removed. The critical value is 1.32–1.39 for a significance level of 0.05. Therefore, the null hypothesis  $H_0$  is rejected and the conclusion is that mean temperature has a relationship with the COVID-19 propagation.

$$f(51, 1661) = 11.60 > 1.39 \tag{11}$$

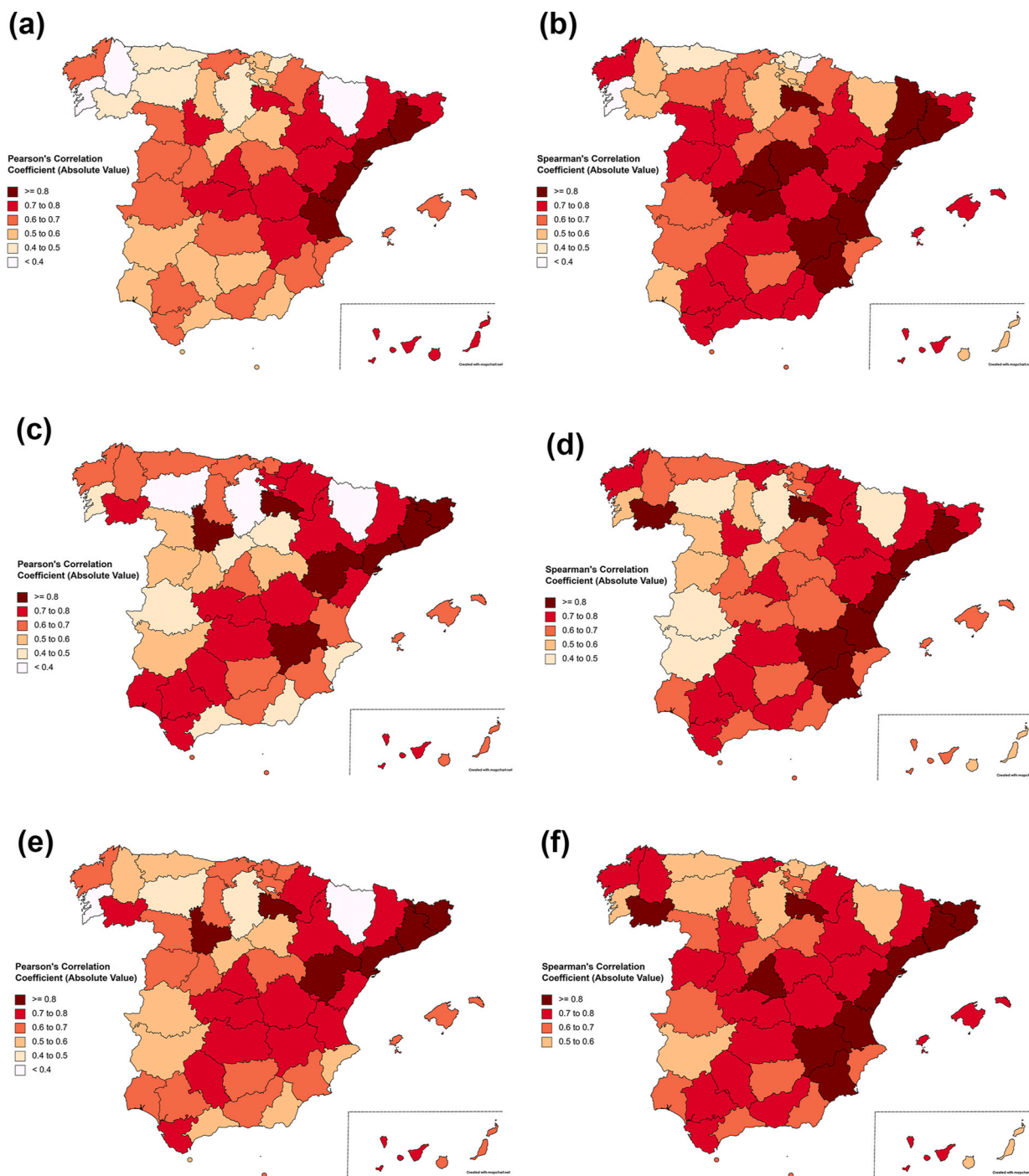


Fig. 4. Correlation between the temporal series for each province and the maximum, minimum and mean temperatures.

### 3.4. Antibodies analysis

The percentage of population with antibodies data *IgGd* [%] has also been correlated with the climate variables. First round analysis was carried from the 27th April to the 11th May. Second round analysis was carried from the 18th May to the 1st June. Third round analysis was carried from the 8th June to the 22nd June. Correlations with the climate variables are shown for each round: First Round (Table 4), Second Round (Table 5), Third Round (Table 6). As could be expected, correlation coefficients are quite similar when compared with the ones obtained from the COVID-19 density of positives metric *PCRd*.

Third Round analysis correlates better than the second and first round antibodies results *IgGd* [%] with the climate variables. Mean and

minimum temperature are more significant factors than the maximum temperature. Daily sun hours and rainfall seem not to be very significant factors, showing almost no correlation. Atmospheric pressure is the climate variable that correlates with *IgGd* [%] showing a bigger difference with respect to the temperature. These differences were smaller for the *PCRd* correlation analysis. The atmospheric pressure seems to impact more the COVID-19 transmission and the low temperatures with the new detected cases by PCR, which are more related to the severity of the infection, since it does not include the asymptomatic. Nevertheless, the impact of these two indicators must be implicit in the mean temperature (more daily sun hours, higher temperature) and in the atmospheric pressure (storms in Spain are usually associated to the entry of low pressures fronts from the North and West of the Spanish peninsula).



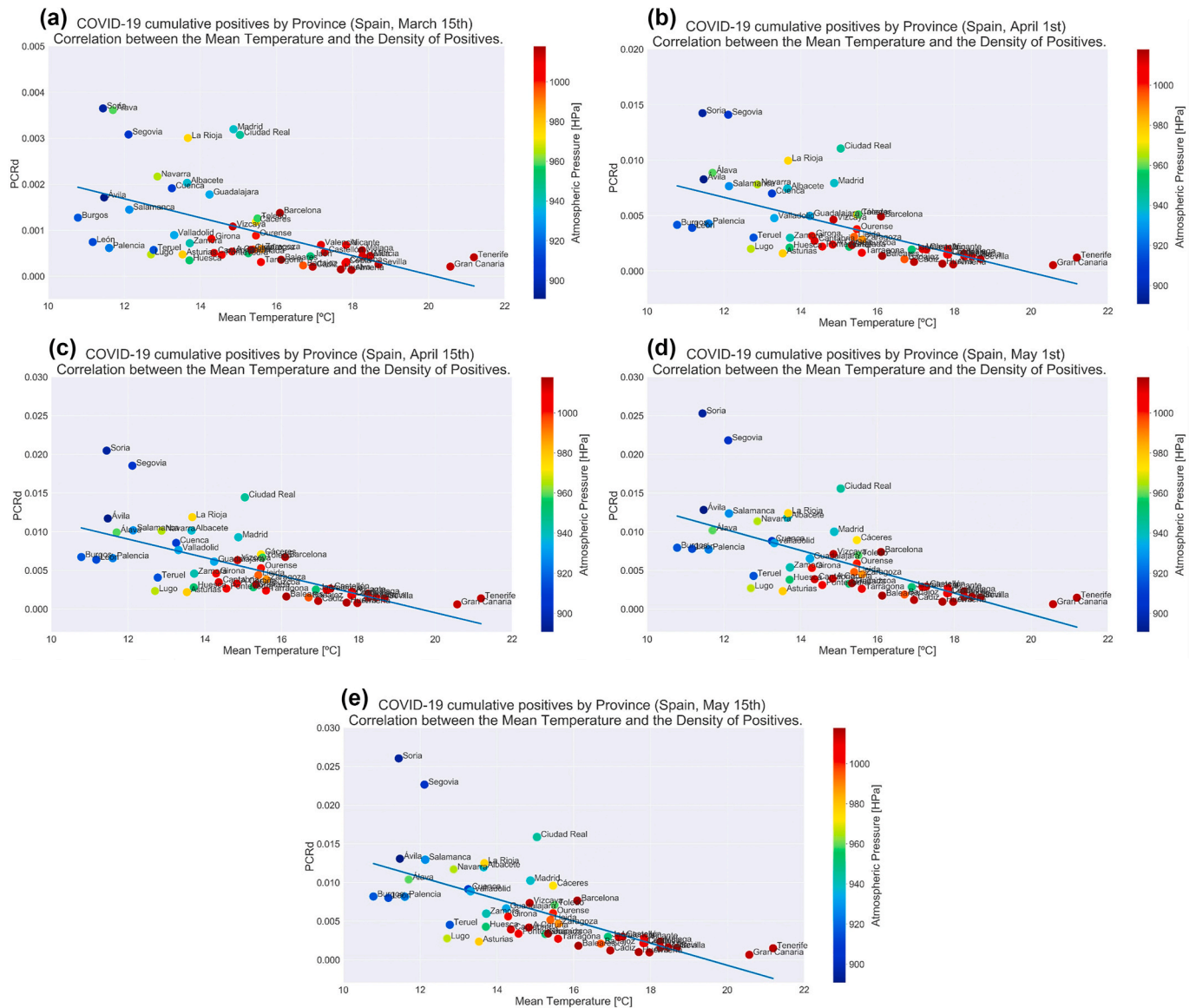


Fig. 5. Correlation between the COVID-19 density of positives *PCRd* and the mean temperature.

Table 3

Linear Regression (Mean Temperature, Density of Positives) coefficients.

Measurement Day	$a_0$ (Constant)	$a_1$ (Slope)	RMSE
March 15th	0.004161	- 0.000206	$10^{-6}$
April 1st	0.016826	- 0.000848	$7 \times 10^{-6}$
April 15th	0.023314	- 0.00119	$1.1 \times 10^{-5}$
May 1st	0.026891	- 0.00138	$1.5 \times 10^{-5}$
May 15th	0.027824	- 0.001427	$1.6 \times 10^{-5}$

Therefore, the virus could propagate along the whole year, more in winter than in summer, and the infections in summer would be more benevolent than in winter. These results suggest that during clear sky days, there would be lesser possibility of increment of positive cases, furthermore rainy winter could be very susceptible weather for Spain for the spread of this virus.

### 3.5. Estimation of the percentage of asymptomatic population

The linear regression between *PCRd* and *IgGd* is:

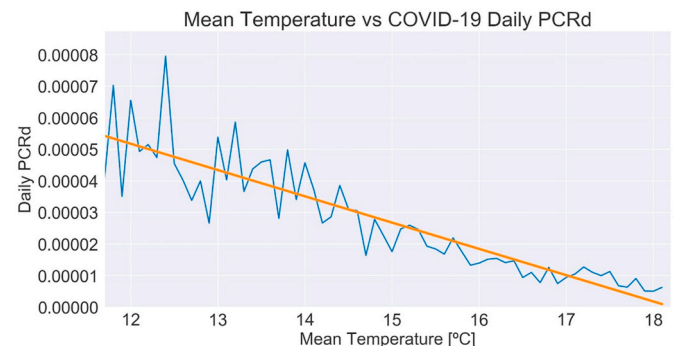


Fig. 6. Expected Daily *PCRd* depending on the mean temperature.

$$IgGd [\%] = 1.605 + 5.17 \times PCRd [\%] \tag{12}$$

This relation tells the percentage of herd immunity is between 5 and 6 times *PCRd*. Therefore, an estimation of the asymptomatic people can be computed based on the number of PCR positives known. To produce



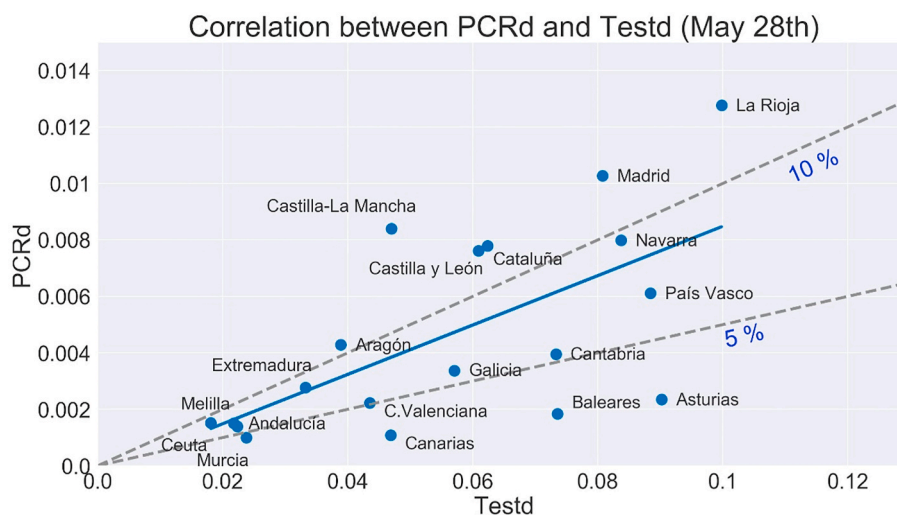


Fig. 8. Relationship between the positives *PCRd* and tests *Testd*.

Correlation Coefficient value is 0.64, whereas Spearman's Correlation Coefficient is 0.66. Two thresholds were included. Both represent the ratio between *Testd* and *PCRd*. A 5% threshold indicates that 1 out of 20 PCR tests provides a positive result. Similarly, a 10% threshold indicates that 1 out of 10 PCR tests is positive Fig. 8 can also be divided into four quadrants. The top left one would be Autonomous Communities with low *Testd* and high *PCRd* (Worst scenario). The bottom right quadrant would be Autonomous Communities with high *Testd* and low *PCRd* (Best scenario). The other two quadrants show low *Testd* and low *PCRd* (Second best scenario) and high *Testd* and high *PCRd* (Third best scenario).

#### 4. Conclusions

Climate variables provide an explanation for the COVID-19 propagation. Spain is a suitable country for this analysis since it exhibits four different types of climate: oceanic, Mediterranean, sub-tropical and continental. Temperature is an important factor to explain the COVID-19 propagation behaviour. The mean temperature is the one that best correlates with the density of positives *PCRd*. The correlation between the mean temperature and the density of COVID-19 positives *PCRd* increased from March to April, and then it is stabilized. Moreover, the minimum temperature correlates better than the maximum temperature with the density of positives *PCRd*. In colder places, the virus propagation can be extremely fast. Atmospheric pressure has a negative correlation with the density of COVID-19 positives *PCRd* in Spain provinces. Atmospheric pressure is a quite significant factor and could be one of the most related with the virus transmission. Daily sun hours and rainfall seem to be non-significant factors, although their contribution might be implicit in the mean temperature and also in the atmospheric pressure. Both values are considered not important to explain the COVID-19 spread behaviour. The ANOVA test has also confirmed the existing relationship between the mean temperature and the COVID-19 propagation. We also provide a plot for the Expected Daily PCR density depending on the mean temperature. According to our results there would be lesser possibility of increment of positive cases under good weather conditions, whilst cold weather is more susceptible for the COVID spread. Besides, the mild coastal weather conditions show an important influence in the Eastern and Western parts of Spain.

Finally, we have performed the estimation of the asymptomatic population based on the results of the antibodies analysis, showing that for each PCR positive case there are more than 4 people that have stayed in mild conditions or even asymptomatic. Besides we have shown a positive correlation between the density of the PCR positives and the

number of tests that have been performed in each state in Spain.

In conclusion climate variables explain some important aspects in the COVID transmission and its strength but cannot be used to justify the disappearance/vanishing of the outbreak without taken into consideration epidemiological measures of social distance.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijheh.2021.113723>.

#### References

- Aemet, OpenData, 2020a. Site for the distribution of Meteorological Spanish data. <https://opendata.aemet.es/centrodedescargas/inicio> accessed 2 July.
- Aemet, blog, 2020b. Climate data of Spain during the COVID-19 lockdown. <https://aemetblog.es/2020/05/05/una-cuarentena-en-penumbra/> accessed 2 July 2020.
- Ahmadi, M., Sharifi, A., Dorosti, S., Jafarzadeh Ghouschi, S., Ghanbari, N., 2020. Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Sci. Total Environ.* 729 <https://doi.org/10.1016/j.scitotenv.2020.138705>.
- Auler, A.C., Cássaro, F.A.M., da Silva, V.O., Pires, L.F., 2020. Evidence that high temperatures and intermediate relative humidity might favor the spread of COVID-19 in tropical climate: a case study for the most affected Brazilian cities. *Sci. Total Environ.* 729 <https://doi.org/10.1016/j.scitotenv.2020.139090>.
- Bannister-Tyrrell, M., Meyer, A., Faverjon, C., Cameron, A., 2020. Preliminary evidence that higher temperatures are associated with lower incidence of COVID-19, for cases reported globally up to 29th February 2020. *medRxiv*. <https://doi.org/10.1101/2020.03.18.20036731>.
- Carlos III. Institute COVID-19. Daily update data of the COVID-19 pandemic in Spain. <https://cnecovid.isciii.es/covid19/> accessed 2 July 2020.
- Farhan Bashir, M., Ma, B., Bilal, Komal, B., Adnan Bashir, M., Tan, D., Bashir, M., 2020. Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Sci. Total Environ.* 728 <https://doi.org/10.1016/j.scitotenv.2020.138835>.
- Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., Rinaldo, A., 2020. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. In: Proceedings of the National Academy of Sciences of the United States of America. <https://doi.org/10.1073/pnas.2004978117>.
- Gupta, A., Pradhan, B., 2020. Assessment of temporal trend of COVID-19 outbreak in India. *OSF Preprints*. <https://doi.org/10.31219/osf.io/qr6e6>.
- Gupta, A., Banerjee, S., Das, S., 2020a. Significance of geographical factors to the COVID-19 outbreak in India. *Modelling Earth Systems and Environment* 1–9. <https://doi.org/10.1007/s40808-020-00838-2>.
- Gupta, A., Pradhan, B., Nizam Abdul Maulud, K., 2020b. Estimating the impact of daily weather on the temporal pattern of COVID-19 outbreak in India. *Earth Systems and Environment* 1–12. <https://doi.org/10.1007/s41748-020-00179-1>.
- Instituto Nacional de Estadística (Ine), 1 January 2020. Data of the population of the provinces of Spain. <https://www.ine.es/jaxiT3/Tabla.htm?t=31304> accessed 2 July 2020.
- Mapchart, 2021. Website that allows the user to create a custom map. <https://mapchart.net/> accessed 15 February.

Şahin, M., 2020. Impact of weather on COVID-19 pandemic in Turkey. *Sci. Total Environ.* 728 <https://doi.org/10.1016/j.scitotenv.2020.138810>.

Shim, E., Tariq, A., Choi, W., Lee, Y., Chowell, G., 2020. Transmission potential and severity of COVID-19 in South Korea. *Int. J. Infect. Dis.* 93, 339–344. <https://doi.org/10.1016/j.ijid.2020.03.031>.

Wang, Jingyuan, Tang, Ke, Feng, Kai, Lin, Xin, Lv, Weifeng, Chen, Kun, Wang, Fei, 2020. High temperature and high humidity reduce the transmission of COVID-19 (March 9). Available at: SSRN: <https://ssrn.com/abstract=3551767>.