



Article

Nine Months of COVID-19 Pandemic in Europe: A Comparative Time Series Analysis of Cases and Fatalities in 35 Countries

David Meintrup ^{1,*}, Martina Nowak-Machen ² and Stefan Borgmann ³

¹ Faculty of Engineering and Management, University of Applied Sciences Ingolstadt, 85049 Ingolstadt, Germany

² Department of Anaesthesia and Intensive Care Medicine, Ingolstadt Hospital, 85049 Ingolstadt, Germany; martina.nowak-machen@klinikum-ingolstadt.de

³ Department of Infectious Diseases and Infection Control, Ingolstadt Hospital, 85049 Ingolstadt, Germany; stefan.borgmann@klinikum-ingolstadt.de

* Correspondence: david.meintrup@thi.de

Abstract: (1) Background: to describe the dynamic of the pandemic across 35 European countries over a period of 9 months. (2) Methods: a three-phase time series model was fitted for 35 European countries, predicting deaths based on SARS-CoV-2 incidences. Hierarchical clustering resulted in three clusters of countries. A multiple regression model was developed predicting thresholds for COVID-19 incidences, coupled to death numbers. (3) Results: The model showed strongly connected deaths and incidences during the waves in spring and fall. The corrected case-fatality rates ranged from 2% to 20.7% in the first wave, and from 0.5% to 4.2% in the second wave. If the incidences stay below a threshold, predicted by the regression model ($R^2 = 85.0\%$), COVID-19 related deaths and incidences were not necessarily coupled. The clusters represented different regions in Europe, and the corrected case-fatality rates in each cluster flipped from high to low or vice versa. Severely and less severely affected countries flipped between the first and second wave. (4) Conclusions: COVID-19 incidences and related deaths were uncoupled during the summer but coupled during two waves. Once a country-specific threshold of infections is reached, death numbers will start to rise, allowing health care systems and countries to prepare.

Keywords: COVID-19; SARS-CoV-2; corrected case fatality rate; time series analysis; multiple regression; flip effect; death threshold



Citation: Meintrup, D.; Nowak-Machen, M.; Borgmann, S. Nine Months of COVID-19 Pandemic in Europe: A Comparative Time Series Analysis of Cases and Fatalities in 35 Countries. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6680. <https://doi.org/10.3390/ijerph18126680>

Academic Editor: Jahar Bhowmik

Received: 30 April 2021

Accepted: 14 June 2021

Published: 21 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

On 24 January 2020, the first three cases of COVID-19 in continental Europe were announced in France [1]. Two months later on 17 March 2020, Montenegro was the last country in Europe to report at least one case of COVID-19. In the early phase of the pandemic, transmission rates in Europe were comparably low [2]. Nevertheless, by the end of the year 2020, more than 22 million confirmed cases and more than 500,000 COVID-19 related deaths were counted by the European health authorities. However, infection rates with SARS-CoV-2 and the related mortality of COVID-19 have not been equal among European countries. Significant differences both in infections and COVID related death rates have been observed on a local, regional and national level [3–5].

An uncontrolled spread of SARS-CoV-2 infections would have disastrous consequences for individuals and health care systems worldwide. However, the severe global economic and social impact of lockdown measures needs to be taken into consideration. Gaining a better understanding of dynamic patterns exhibited by the pandemic in Europe is paramount as political inequalities between countries followed by at times opposing strategies to fight the pandemic has made Europe a challenging place to contain the virus.

Observing variables in fixed time intervals, like the daily numbers of COVID-19 incidences and related deaths in each country, naturally produces time series. In the ongoing research focussing on these time series, one can distinguish three main directions: First, the temporal dynamics of COVID-19 incidences and related deaths at the onset of the pandemic in a specific country [6–9]; second, the correlation between COVID-19 incidences and another factor of interest, ranging from air pollution [10] through social capital [11] to stroke admissions [12]; third, the effect of political measures—from social distancing to complete lockdown of public life—on COVID-19 incidence and related death numbers [13–15].

In the present study, we sought to model the temporal relation between COVID-19 incidences and related deaths for 35 European countries with at least 1 million inhabitants, extending our analysis over a period of nine months, including both spring and fall waves of 2020 and the relatively quiet summer months of 2020 observed in most European countries.

We hypothesized that incidence rates and death rates are coupled at certain times, whereas an uncoupling of incidence rates and death rates can be observed when incidences stay below a certain infection threshold. The goal of our study was to describe the dependency of death rates from COVID-19 incidences, taking into account the time delay between infection and death, the changes in mortality over time, and the differences from country to country. Additionally, we aimed to estimate the threshold value of coupled death and incidence rates for each country. The geographical structure of the clusters can provide insights about the dynamic of the spread of the SARS-CoV-2 virus across Europe.

2. Materials and Methods

2.1. Time Series Analysis

Our analysis is based on the COVID-19 incidences and deaths data from the European Centre for Disease Prevention and Control [1]. We restricted our attention to the 35 European countries with a population of more than 1 million. In order to obtain stable results, we used the 7-day moving average of the COVID-19 incidences and of the COVID-19 related deaths.

By 30 November, the pandemic progressed in three distinct phases in most European countries. During the first and third phase, commonly referred to as first and second “waves”, the case fatality rate (CFR) was strongly connected to the incidence rate, typically with a higher percentage of COVID-19 related deaths in the first than in the second wave. In the phase in between the two waves, the CFR was independent from incidence appearing as a random variable around a constant value. In most countries, the first and third phase occurred in the spring and fall of 2020, while the second “quiet phase” was observed during the summer months of 2020 [1].

In order to validate the three different phases of the pandemic for Europe, the following time series model was used:

$$y_t = \begin{cases} \alpha_1 \cdot x_{t-d_1}, & \text{for } 1 \leq t \leq t_1, \\ p_2, & \text{for } t_1 < t < t_2, \\ \alpha_2 \cdot x_{t-d_2}, & \text{for } t \geq t_2. \end{cases} \quad (1)$$

We evaluated a period of nine months, beginning on 1 March 2020 ($t = 1$), and ending on 30 November 2020 ($t = 275$). In this model, (x_t) is the 7-day moving average of COVID-19 incidences, and (y_t) is the (predicted) 7-day moving average of COVID-19 related deaths on day t . The model parameters α_1 and α_2 ($0 \leq \alpha_{1,2} \leq 1$, discretized with an increment of 0.001 for optimization), correspond to the percentage of incidences that led to a fatal outcome. We will refer to α_1 and α_2 as corrected case fatality rates (cCFR). In order to obtain proper case fatality rates, it was necessary to calculate country specific shift parameters (d_1 and d_2 in the model) to account for the correct censoring of the data [16]. The integer-valued shift parameters d_1 and d_2 represent the number of days between the

date of reporting the number of COVID-19 cases and the date of reporting the number of COVID-19 associated deaths. In addition, differences in the efficiency of treatment and in the case reporting systems of the countries can affect the time between a positive test result and death, and hence impact these shift parameters.

The model was fitted individually to each country using a minimum error variance approach. The optimizer that we wrote iterated through all possible integer values for d_1, d_2, t_1, t_2 (with the constraint $t_2 > t_1$) and all $0 \leq \alpha_1, \alpha_2 \leq 1$ (with step size of 0.001) in order to find the minimum of the loss function. It is crucial to note that the optimization did not only include the cCFRs (α_1, α_2), the average number of deaths in phase 2 (p_2), and the shift parameters (d_1, d_2), but also the points in time t_1 and t_2 , where the switch between the phases happened. Hence, in the fitting process, all subdivisions of the nine months in three phases were considered in order to find the best possible time intervals. This implies that the model-fit automatically detects the time interval in which a constant death rate, unconnected to the number of COVID-19 incidences, fits the data better than a coupled death rate proportional to the number of incidences. More precisely, a better model fit means a smaller value of the loss function, which is given by the mean squared error.

2.2. Clustering

The time series analysis described above provides a set of 7 parameters ($\alpha_1, \alpha_2, p_2, d_1, d_2, t_1, t_2$) for each country. These parameters were used to assign countries to specific groups (“clusters”) with similar epidemiology. The relative death rate per 100,000 inhabitants in the second phase was calculated for each country:

$$p_{2rel} = p_2 \cdot \frac{100,000}{\text{Population Size}}. \quad (2)$$

The three parameters ($\alpha_1, \alpha_2, p_{2rel}$) represent relative, and hence comparable measures of the case fatalities for the three phases in each country. As the absolute magnitude of the α -values changes significantly between the first and the second wave, we assigned each country to its corresponding α -ranks r_1 and r_2 . The ranks are ascending, meaning that smaller α -values correspond to lower ranks. The rank difference

$$\Delta r = r_1 - r_2 \quad (3)$$

is an expression of the change in rank of each country from the first to the second wave and reflects aspects of the dynamics of the pandemic. Hence, the three severity measures and the rank difference ($\alpha_1, \alpha_2, p_{2rel}, \Delta r$) were used for the cluster algorithm. The hierarchical clustering was performed with Ward’s method and standardized variables [17].

2.3. Regression

After a relatively quiet summer, in most countries, the COVID-19 incidences started to rise again in the fall, and, at a certain individual time point, COVID-19 related deaths started to rise and reconnected with the incidences. We hypothesize that each country has an individual threshold value C_T . Once this limit is exceeded, a re-coupling of the death rate with the incidence rate can be observed. The beginning of phase 3, the day t_2 , is an optimized parameter in our time series analysis. It can be used to estimate the threshold value C_T that leads to reconnection of the COVID-19 incidences and the related death rate in a specific country:

$$C_T = x_a, \quad \text{with } a = t_2 - d_2. \quad (4)$$

The day t_2 is the point in time, where the connected model fits the data better (i.e., leading to a smaller value of the loss function) than the unconnected model. We find the corresponding threshold value of incidences by going back d_2 days, the country-specific delay between incidence and death.

It is plausible to assume that the threshold value C_T mainly depends on the population size of each country, but also on the severity of the first and second phase, expressed in α_1

and p_{2rel} . We therefore fitted a multiple linear regression model with $\ln(C_T)$ —the natural logarithm of the threshold value—as response variable, and $\ln(P)$ —the natural logarithm of the population size—and the parameters α_1 and p_{2rel} as factors:

$$\ln(C_T) = c_0 + c_1 \cdot \ln(P) + c_2 \cdot \alpha_1 + c_3 \cdot p_{2rel}. \quad (5)$$

When this linear regression model fits the data well, it can be used to predict the threshold value of COVID-19 incidences for each country that will lead to a reconnection of COVID-19 cases and related deaths.

3. Other Data

Information about gross domestic product (GDP) of the countries analyzed herein provided by Wikipedia [18]. Cumulative number of infected and deceased individuals in these countries was obtained from the homepage of the World Health Organization (WHO) [19] on 13 April during the third wave.

All statistical analyses were performed with the statistical software JMP® Pro [20].

4. Results

Table 1 contains the results of our statistical analysis for the 35 European countries that were included in our study. Following the country name, the next seven columns contain the parameter estimates of our model (see Equation (1)). Based on these parameter estimates, countries were assigned to groups with similar outcomes (clusters). A typical representative of the time series model for each cluster is displayed in Figure 1: Greece, representing cluster 1, Germany, representing cluster 2, and the Netherlands, representing cluster 3. The table is organized by cluster, and alphabetically by country name within each of the three clusters.

Cluster 1 consists of countries that were mildly affected by the first wave, but suffered from high cCFRs in the second wave (Figure 2a red cluster, Figure 2b rank parallel plot). In addition, these countries exhibit high relative death rates (p_{2rel}) during the “quiet” phase 2. Cluster 1 can be described as countries in which the CFRs increased slowly over the course of two waves, but, unfortunately, the situation declined constantly and ultimately led to the highest cCFRs in all of Europe. Countries in cluster 1 are all geographically located in the South-East of Europe. A typical representative of cluster 1 is Greece (see Figure 1a), switching from a relatively low $\alpha_1 = 4.7\%$ to a relatively high $\alpha_2 = 4.2\%$, which was the highest value of all considered countries in the second wave.

Cluster 2 includes Central European countries such as Germany, Poland, and the Czech Republic (Figure 2a). These countries started with the lowest cCFRs across Europe in the first wave. Cluster 2 also showed a slow progression of CFRs during the pandemic; however, compared to cluster 1, the CFRs in cluster 2 never reached the same increase in CFR during the second wave. In Germany, for example (Figure 1b), during the first wave, $\alpha_1 = 4.5\%$ was similar to $\alpha_1 = 4.7\%$ of Greece. However, Germany’s CFR during the second wave (phase 3) remained much lower than the CFR of Greece during the same time period ($\alpha_2 = 1.5\%$). Furthermore, in Germany, the incidences during the summer months (phase 2) declined substantially when compared to the first wave (phase 1) which marks a difference between the countries in cluster 1 where the incidences failed to drop during phase 2.

The third cluster, located in Western Europe (blue countries in Figure 2), showed an opposite course to cluster 1. As demonstrated in the rank parallel plot (Figure 2b), countries in cluster 3 started out with the highest cCFRs across Europe in the first wave, and switched to relatively low CFRs in the second wave. France, for example, had the highest cCFR of all 35 countries in the first wave ($\alpha_1 = 20.7\%$), while in the second wave, despite a high number of incidences, the cCFR was low ($\alpha_2 = 1.3\%$). In the Netherlands (cluster 3, Figure 1c), the first wave led to a very high $\alpha_1 = 13.1\%$, followed by an absolute and relative decrease to $\alpha_2 = 0.8\%$.

Table 1. Results of the time series analysis for 35 European countries. The first 7 columns contain the parameters of the time series model (Equation (1)): α_1 and α_2 = corrected case fatality rates of phase 1 and phase 3. d_1 and d_2 = shift parameters (days) of phase 1 and 3. Date of switch between phase 1 and phase 2: March 1 + t_1 (days). Date of switch between phase 2 and phase 3: March 1 + t_2 (days). p_2 = average deaths in phase 2. Δr = the difference of the α -ranks. C_T = threshold value (Equation (3)). The last column contains the result of the cluster algorithm with three clusters. H. = Herzegovina.

Country	α_1	d_1	t_1	t_2	α_2	d_2	p_2	Δr	C_T	Cluster
Albania	3.2%	4	141	187	1.9%	3	4.173	-18	136.1	1
Bosnia and H.	5.4%	21	136	228	3.5%	10	7.892	-11	168.0	1
Bulgaria	4.1%	15	140	221	3.7%	14	7.125	-22	129.6	1
Greece	4.7%	8	164	252	4.2%	17	5.100	-17	487.0	1
Macedonia	4.9%	6	131	223	2.7%	8	4.545	-12	132.6	1
Moldova	4.4%	11	110	166	2.1%	11	7.642	-14	326.4	1
Romania	6.0%	10	147	221	1.9%	2	39.945	-1	1627.4	1
Austria	3.4%	13	97	220	1.3%	13	1.107	-3	736.1	2
Belarus	0.6%	1	108	237	0.8%	23	4.896	-4	308.0	2
Croatia	4.1%	18	137	196	1.8%	10	1.486	-10	269.1	2
Czech Rep.	3.4%	11	130	198	1.7%	10	1.460	-10	450.3	2
Estonia	3.5%	11	81	250	1.8%	20	0.059	-13	31.4	2
Germany	4.5%	14	117	227	1.5%	18	6.457	-1	1778.6	2
Latvia	2.6%	24	133	228	1.7%	10	0.109	-18	52.1	2
Lithuania	3.2%	8	114	212	1.1%	9	0.175	-2	46.0	2
Norway	2.9%	16	81	257	0.5%	18	0.303	4	156.6	2
Poland	5.8%	6	78	211	2.0%	11	11.327	-5	515.4	2
Portugal	4.3%	5	100	195	1.3%	6	4.024	0	340.4	2
Serbia	2.3%	1	190	231	1.0%	8	1.245	-5	97.4	2
Slovakia	2.0%	10	122	204	1.0%	17	0.145	-7	72.3	2
Slovenia	7.2%	13	112	225	1.5%	14	0.302	9	126.0	2
Switzerland	4.8%	10	103	227	1.2%	14	1.054	8	323.4	2
Belgium	16.9%	5	102	213	1.4%	15	4.643	19	879.7	3
Denmark	5.2%	1	108	198	0.5%	25	0.377	20	124.3	3
Finland	5.9%	15	85	257	0.6%	2	0.348	21	212.4	3
France	20.7%	7	97	220	1.3%	13	26.506	22	6011.3	3
Hungary	13.7%	8	168	202	2.2%	1	1.943	0	779.3	3
Ireland	6.8%	10	125	212	0.6%	19	0.716	23	150.7	3
Italy	14.3%	4	120	217	2.0%	11	12.429	5	1535.0	3
Netherlands	13.1%	4	97	206	0.8%	14	2.834	24	702.3	3
Spain	11.6%	5	93	177	1.4%	9	6.521	13	4064.4	3
UK	15.4%	2	130	203	2.0%	21	15.033	4	1190.1	3
Kosovo					Model not applicable					
Sweden					Model not applicable					
Ukraine					Model not applicable					

The 5-point summary statistics for the five model parameters ($\alpha_1, d_1, \alpha_2, d_2, p_{2rel}$) are summarized in Table 2.

In general, cCFRs in Europe were lower in the fall (phase 3, second wave) than in the spring (phase 1, first wave) of 2020. While the median of the cCFRs during the first wave was 4.8%, it dropped to 1.5% in the second wave. The shift parameters increased from nine days in the first wave to 12 days in the second wave. Both results indicate a less vulnerable population and improved treatment of COVID-19 patients during the second wave.

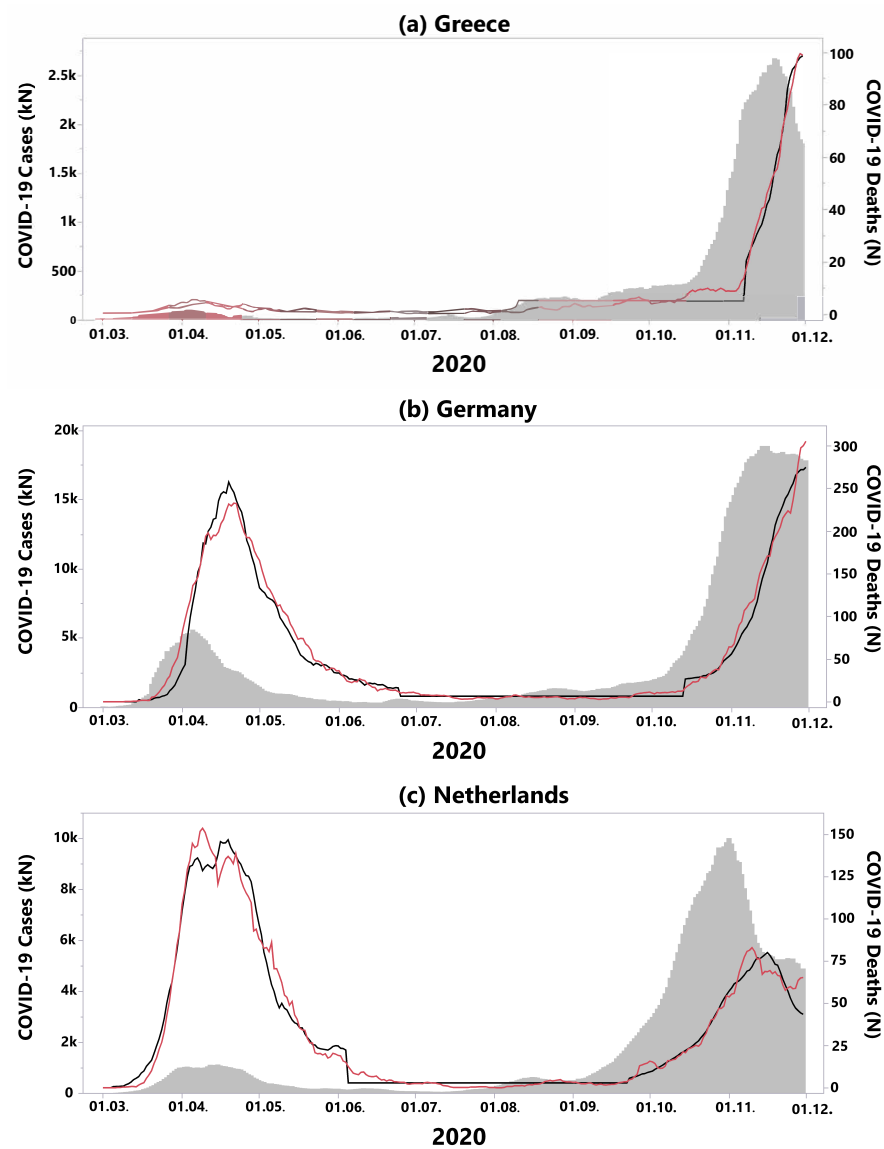


Figure 1. Examples of time series analyses performed on the temporal incidence of SARS-CoV-2 infections (grey area, left scale) and the incidence of COVID-19 associated deaths (red line, right scale). The three-phase-prediction model (black line, right scale) for COVID-19 associated deaths based on actual incidences has been estimated for each country. The second phase of the model with a constant number of deaths is identifiable by the horizontal black line. (a) Greece (cluster 1), exhibits a constant increase of incidences over time ($\alpha_1 = 4.7\%$ and $\alpha_2 = 4.2\%$). (b) For Germany ($\alpha_1 = 4.5\%$ and $\alpha_2 = 1.5\%$ —cluster 2), and (c) the Netherlands ($\alpha_1 = 13.1\%$ and $\alpha_2 = 0.8\%$ —cluster 3), the threshold effect is particularly visible: the case fatality rates did not increase during the summer months, despite a relevant increase of COVID-19 incidences.

Table 2. The 5-point summary statistics of the five model parameters estimated for 32 European countries.

	α_1	d_1	α_2	d_2	p_{2rel}
Minimum	0.006	1	0.005	1	0.003
First Quartile	0.034	5	0.010	9	0.009
Median	0.048	9	0.015	12	0.019
Third Quartile	0.071	13	0.020	17	0.046
Maximum	0.207	24	0.042	25	0.239

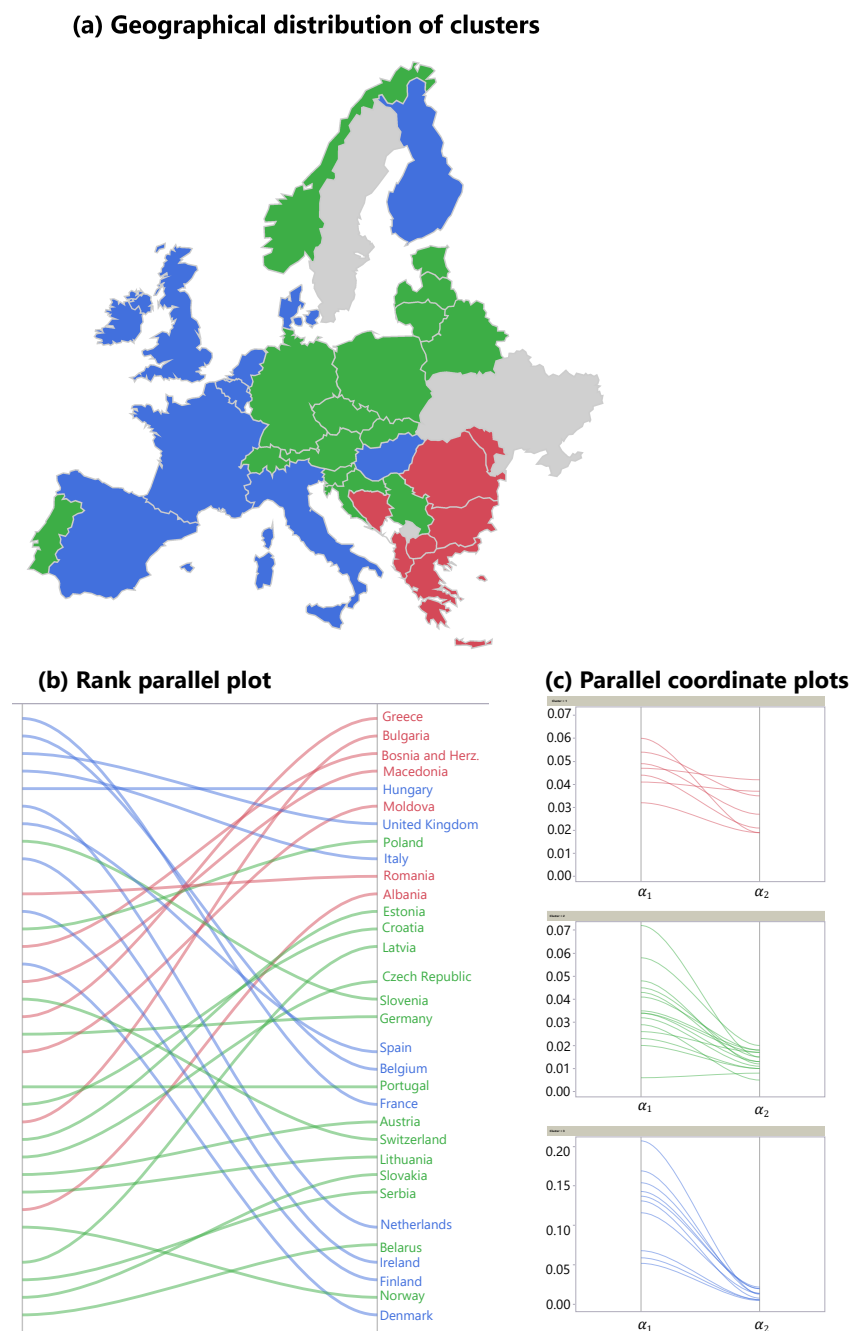


Figure 2. Graphical representations of the clustering. The three clusters were determined using the four parameters (α_1 , α_2 , p_{2rel} , Δr) deduced from the time series analysis. **(a)** Although these parameters did not contain any geographical information, the clusters represent different regions in Europe. The parallel plot of the α -ranks shows that the first cluster (red, southeast Europe) was characterized by low cCFRs in the first wave and high values in the second wave (compared to the rest of Europe). The third cluster (blue, western Europe) exhibited an opposite pattern. In the second cluster (green, central Europe), low cCFRs in the first wave were followed by slightly higher cCFRs in the second wave. In grey: countries not attributable to a particular cluster. **(b)** Overall, the rank plot shows evidence of the flip effect: in nearly all countries, the α -rank changed. **(c)** Absolute values of α_1 and α_2 of the three clusters. Note that these values are standardized for the clustering algorithm.

The fact that the constant model for the death rate was the best fit for the data in most countries during the summer months has another important consequence. In many European countries, the COVID-19 incidences were slightly increasing over the summer,

whereas death rates remained low. This result undermines our findings that incidences and death rates are not necessarily connected, as long as the incidences stay below a certain threshold, which can be described as an “uncoupling” of incidences and deaths. Once an individual threshold of infections in a country is reached, the death rates will begin to increase following the incidence rates, which can be described as the “re-coupling” of incidences and deaths. This “threshold” effect is shown in Figure 1 for representative countries of all three clusters.

Table 3 shows examples of uncoupling for Austria, France, Germany, and Denmark during the quiet summer months of the pandemic, where incidences were slowly increasing as the pandemic approached the second wave. In all four countries, the number of cases increased between 2- and 4-fold. In contrast to the first (phase 1) and second wave (phase 3), the absolute number of COVID-19 related deaths stayed constantly low.

Table 3. COVID-19 incidences and related deaths in July and August 2020 for four countries. Although the incidences increased significantly, the related deaths did not. Below a certain threshold, incidences and deaths are not necessarily connected (“uncoupling” of incidences and deaths).

Country	Incidences July	Deaths July	Incidences August	Deaths August
Austria	3343	15	6209	15
France	22,313	441	91,370	352
Germany	14,439	168	33,683	157
Denmark	974	10	2975	9

Column 9 in Table 1 contains difference in α -ranks (see Equation (3)), followed by the threshold value C_T (see Equation (4)). Based on the relative increase of the distance metric in the hierarchical clustering, three clusters were used with 7, 10, and 15 members, respectively. The mean values of the four parameters (α_1 , α_2 , p_{2rel} , Δr) that were used for the cluster algorithm are displayed in Table 4.

Table 4. Mean values per cluster. While cluster 1 shows a strong decrease of relative position, in cluster 3, the relative rank increases by almost the same amount. In cluster 2, the average rank decreases by 3.8 points.

Cluster	N	α_1	α_2	p_{2rel}	Δr
1	7	0.0467	0.029	0.164	−13.6
2	15	0.0364	0.013	0.017	−3.8
3	10	0.1236	0.013	0.020	15.1

The last column in Table 1 displays the cluster the corresponding country has been assigned to. A geographical representation of the clusters is shown in Figure 2a. The parallel coordinate plots in Figure 2c show the absolute values of the parameters (α_1 , α_2) for each country in the three clusters. The last parameter Δr is represented in Figure 2b. Each line in this graph connects the rank of α_1 on the left side with the corresponding rank of α_2 on the right side for the given country. The rank difference Δr between both sides represents the relative improvement or deterioration, and was used as a fourth parameter in the clustering process.

The linear regression model (see Equation (5)) is given by Equation

$$\ln(C_T) = -3.78 + 0.87 \cdot \ln(P) + 2.86 \cdot \alpha_1 + 1.50 \cdot p_{2rel}. \quad (6)$$

The variance inflation factors of all involved factors were below 2, so that there is no indication of multicollinearity in the data.

A graphical representation of the model is shown in Figure 3a. The predicted response value is displayed versus the actual response value. The better the points follow the

diagonal line, the better the model fits the data. The results of the *t*-tests for every parameter in the model are displayed in Table 5.

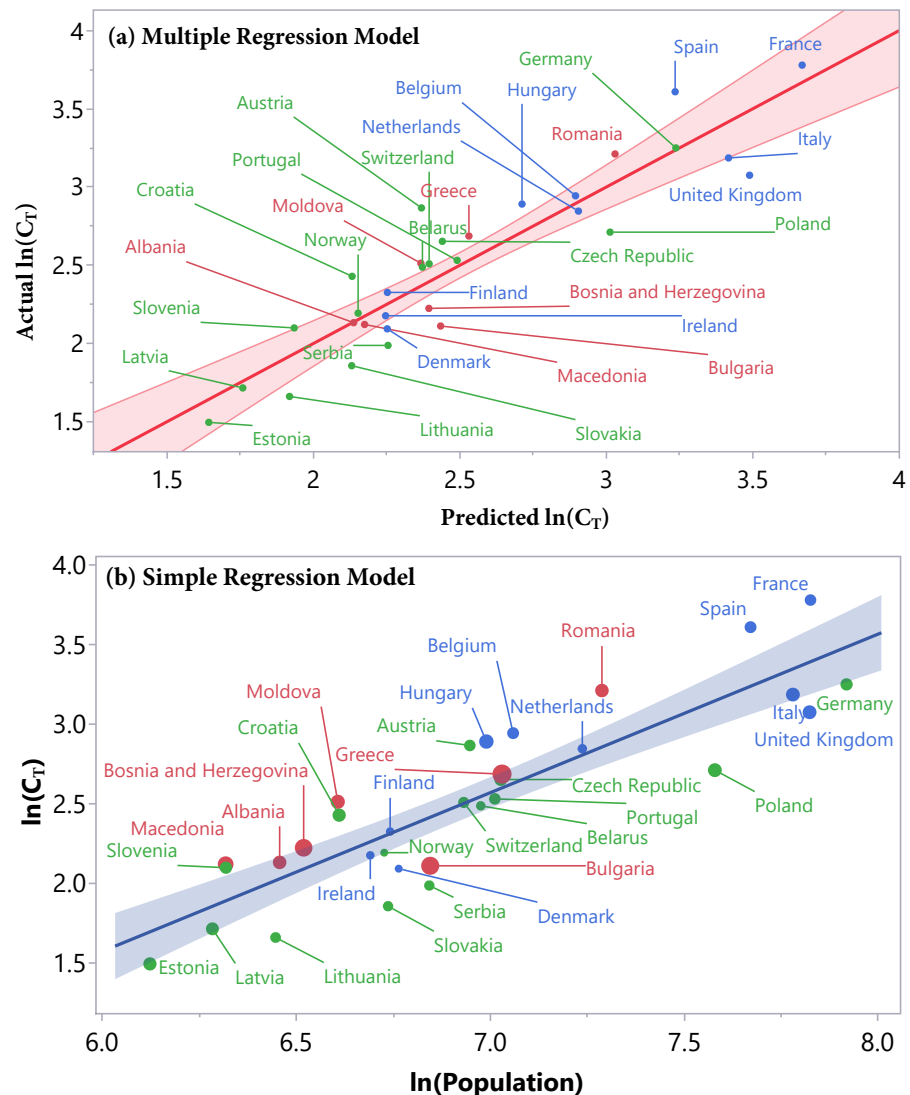


Figure 3. Graphical representations of the linear regression models for the logarithm of the threshold value $\ln(C_T)$. (a) The multiple regression model ($R^2 = 85.0\%$, root mean squared error (RMSE) = 0.23) contains three factors. The graph shows the predicted $\ln(C_T)$ versus the actual $\ln(C_T)$. The points roughly follow the diagonal line and are almost all inside the confidence region, showing a good model fit. For example, Germany is about four times bigger than Latvia, but both threshold values are very well predicted by the model. (b) The simple linear model ($R^2 = 77.2\%$, RMSE = 0.27) with $\ln(\text{Population})$ as the only factor shows the dominance of this factor. The size of points is proportional to the cCFR α_2 of the second wave. In both graphs, the colors represent the corresponding cluster. There is no apparent structure with respect to point size or cluster, the linear relation is independent of these factors.

Table 5. Parameters of the multiple regression model ($R^2 = 85.0\%$) for the prediction of $\ln(C_T)$. All parameters contribute significantly ($p < 0.05$) to the prediction, as shown by the p -values.

Term	Estimate	Standard Error	t Ratio	p-Value
c_0	−3.78	0.70	−5.43	<0.0001
c_1	0.87	0.10	8.31	<0.0001
c_2	2.86	1.01	2.82	0.0086
c_3	1.50	0.60	2.51	0.0182

For comparison, we also fitted a simple linear regression, with the logarithm of population size as only factor. This resulted in the following model:

$$\ln(C_T) = -4.39 + 0.99 \cdot \ln(P) \quad (7)$$

with R^2 dropping to $R^2 = 77.2\%$, and root mean squared error (RMSE) = 0.27. The corresponding graph with the factor $\ln(P)$ on the x -axis and the response $\ln(C_T)$ on the y -axis is shown in Figure 3b. The color of the points corresponds to the respective cluster, the size of the points is proportional to the cCFR α_2 during the second wave.

Three European countries displayed exceptional courses of the pandemic; hence, our three-phase-model was not suited to describe the course of the pandemic (Figure 4):

- In Kosovo, the second phase with generally low death rates and defined by an uncoupling of incidences and death rates is missing (Figure 4a). Instead, the country suffered from a second wave during summer, directly followed by a third wave.
- It is well known that Sweden chose a different way to respond to the pandemic compared to neighbouring states in terms of restrictive measures [21]. In particular, during the first wave, the recorded death rate was much higher than predicted by the incidences, probably due to a relatively low level of political protective measures (Figure 4b). In the meantime, Sweden changed its policies with strict protective measures for the population [22]. Consequently, in the second wave, our model fits the data well, with a cCFR of $\alpha_2 = 0.8\%$.
- In Ukraine, no waves could be observed to date, instead COVID-19 incidences and related deaths increased constantly over the course of the year 2020 (Figure 4c).

The phenomenon of clustering due to similar epidemiological patterns raises the question of whether it is favourable for a country to belong to a certain cluster. Therefore, current data from the WHO were used to calculate the cumulative incidence of COVID-19 associated deaths and the cumulative CFR for each of the countries examined herein. Figure 5 shows that countries belonging to cluster 1 and 3 had similar COVID-19 associated death incidences and CFRs, although there was a high variance especially of countries belonging to cluster 3. When calculating both values for the total population of all countries belonging to a cluster, cumulative incidence of COVID-19 associated deaths were 138.2, 108.1, and 166.3 for cluster 1, 2, and 3, respectively. The CFR of cluster 1, 2, and 3 was 2.9%, 1.8%, and 2.5%, respectively (stars in Figure 5). This finding suggests that, in summary, it was probably favourable for a country to have been affected by a moderate first wave (cluster 2) while it was unfavourable to miss out on the first wave (cluster 1), or to be struck by a disastrous first wave (cluster 3)—although Finland and Denmark, both belonging to cluster 3, exhibited a cumulative low COVID-19 burden. In Denmark, 42.08 of 100,000 individuals died from COVID-19 (CFR 1.01%) and in Finland 15.84 of 100,000 inhabitants (CFR 1.03%). The highest death incidence was noticed in Hungary (245.37/100,000 inhabitants) (CFR 3.29%) and in Bosnia and Herzegovina (226.81/100,000 inhabitants) (CFR 4.05%) belonging to cluster 3 and cluster 1, respectively.

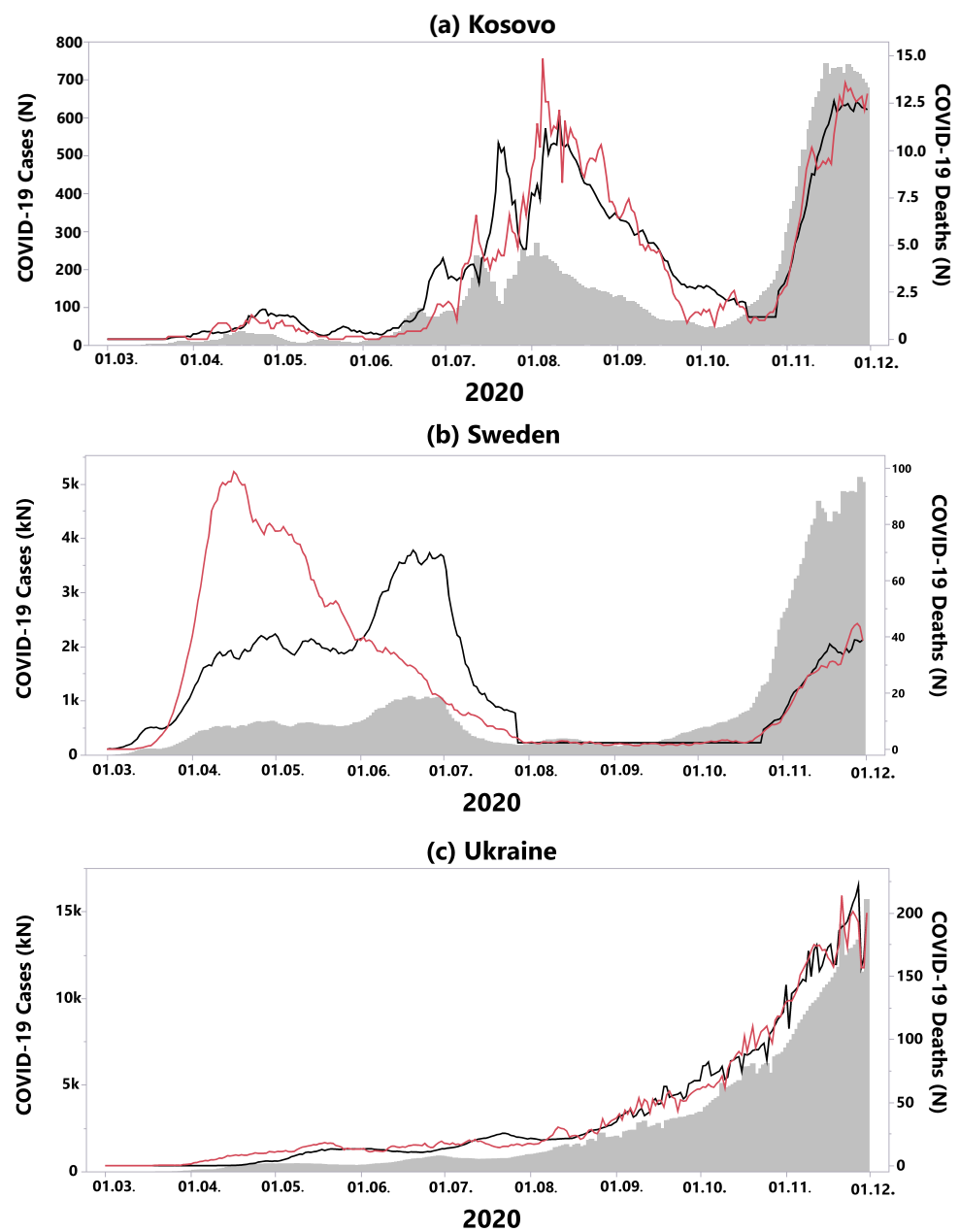


Figure 4. Countries where the three-phase-model was not applicable. (a) In Kosovo, the second phase with a relative low death rate was missing. Instead, the country suffered a second wave during summer, directly followed by a third wave. (b) In Sweden, at the beginning of the first wave, the recorded death rate was much higher than predicted by the incidences. In the second wave, the model fitted the data well, with a cCFR of $\alpha_2 = 0.8\%$. (c) In Ukraine, there were no waves; instead, COVID-19 incidences and related deaths increased constantly over the course of the year 2020.

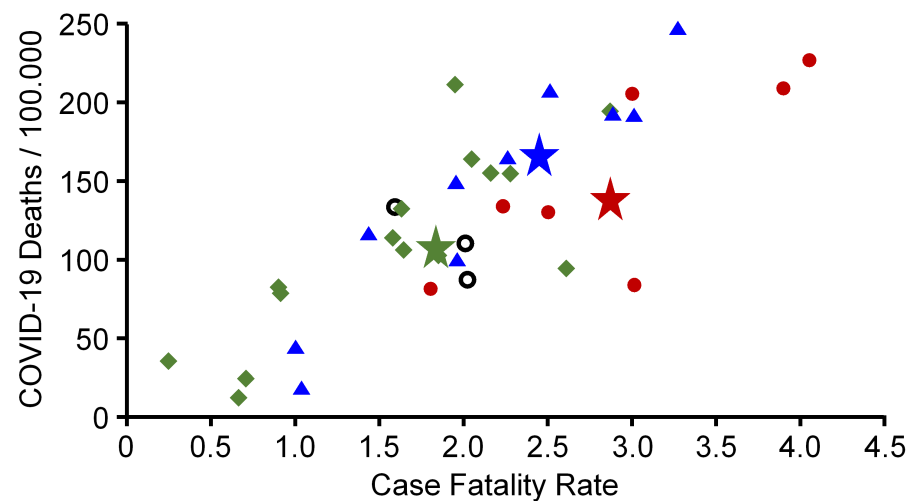


Figure 5. Cumulative COVID-19 associated death incidence per 100,000 inhabitants and cumulative case fatality rates of European countries with more than 1 million inhabitants. Stars (average) indicate the values for the total population of the corresponding cluster. Data were obtained on 13 April 2021 from the homepage of the World Health Organization [19]. Red circles, green squares, blue triangles = countries belonging to cluster 1, 2, and 3, respectively. Open black circles: countries belonging to no cluster.

5. Discussion

Our data show that COVID-19 has spread inhomogeneously across Europe during the first and second waves. We conducted a time series analysis, which showed that our three-phase-model (first wave–in-between waves–second wave) describes the course of the pandemic very well in 32 out of 35 European countries.

The coupling of COVID-19 incidences and death rates could be shown in two distinct phases, namely the first and second waves in the spring and fall of 2020, whereas an uncoupling could be found in the phase in between waves. Once a certain individual threshold of infection rates was reached, death rates and incidences started to re-couple and deaths started to increase. This finding is important as health care systems across Europe vary greatly and display fundamental differences in ICU capacities and medical resources. If the individual coupling threshold in certain countries could be predicted, preparations for distribution of medical resources and staff could be undertaken early and an overcrowding of hospitals as well as overflowing ICUs might be limited. We have previously shown that individual coupling thresholds can be successfully applied to smaller geographical regions within a country to further maximize preparation efforts as well as political measures for closures of schools and businesses [3]. These individual thresholds for countries as well as defined geographical regions could define the way countries and communities handle lockdown measures and guide school opening strategies as well as economy stabilizing business opening concepts.

In addition, we defined so-called “shift parameters” that represent the number of days between reported infections and reported COVID-19 associated deaths. The shift factors show a wide range in the two waves (phase 1 and phase 3) across Europe. The wide range of values emphasises that it is crucial to account for these shifts in order to obtain reliable cCFRs for individual countries. The shift parameters increased from nine days in the first wave to 12 days in the second wave across Europe. The increase in time between infection and death could possibly be explained by a certain immunity that had occurred by the second wave resulting in a less vulnerable population and a lower overall CFR in the second wave. The cCFRs in the second wave ranged from 0.5% in Denmark to 4.2% in Greece. Overall, these values are significantly lower than in the first wave. Possible explanations include increased test rates, protection of the most vulnerable individuals within a population and other political measures. In addition, medical knowledge of

the virus and its resulting disease had improved from wave one to wave two, adding more specific pharmacological treatments as well as progress in hospital organization and resource management.

We postulate that there are two additional effects that lead to lower cCFRs in the second wave. First, the most vulnerable individuals within a population had already been affected during the first wave, and outcomes were detrimental. Second, there is a large body of evidence showing a high number of unreported, often asymptomatic cases of SARS-CoV-2 infections [23,24] that might have contributed to a certain level of immunity across European countries with limited means for widespread testing. A recent meta-analysis of seroprevalence estimates a global factor between confirmed and actual infections of 11.9 on national levels [25], pointing towards a high level of undetected infections. Hence, the exposure to the virus had been much higher during the first wave than reported by the registered positive tests, which might have led to a partial immunity of the exposed population. Partial immunity of the population would also explain the observation that, after one year of the pandemic, countries belonging to cluster 2 in general had lower cumulative CFRs and lower cumulative COVID-19 associated death incidences than countries belonging to cluster 1 (Figure 5).

Adding both effects might have lowered the overall vulnerability of the population in the second wave and hence might have contributed to the lower cCFRs reported in our data. If these statements were true, we should also have observed the opposite: countries that were mildly affected in the first wave would have suffered from higher cCFRs in the second wave. These switches from “high to low” and from “low to high” respectively are called a “flip effect”.

Indeed, we could show that the geographical spread of SARS-CoV-2 in Europe followed a distinct pattern with similarities and discrepancies in death rates over time. Based on those similarities, we were able to define geographical clusters of countries showing a similar dynamic as the pandemic evolved during the first and second waves in 2020.

In summary, the western part of Europe (cluster 3) went from high cCFRs to low cCFRs, while the countries in the southeastern parts of Europe (cluster 1) exhibited the opposite starting off with low CFRs in the first wave and flipping to extremely high CFRs in the second wave, adding evidence to the “flip effect.” A particularly clear view of this flip effect is offered by the rank parallel plot in Figure 2b. While the red lines (cluster 1) are all trending upwards, the blue lines (cluster 3) are trending downwards. Without the presence of the flip effect, the lines in this diagram would be mostly horizontal. The situation in cluster 2 is similar to cluster 1 but less pronounced, with an average rank loss of -3.8 , compared to -13.6 in cluster 1. In cluster 1, the average duration of the first wave was 138.4 days, compared to 112.5 days in cluster 3, a difference of 26 days. The countries in cluster 3 were much more affected by the first wave, potentially leading to more drastic restriction measures and hence a faster decline of incidences than in less severely affected countries.

It is an interesting result by itself that, although the clustering algorithm did not use any geographical information, the clusters are clearly assignable to different regions of Europe (see Figure 2a). The first cluster is entirely located in the southeastern part of Europe, with Greece, Romania and Bulgaria being its three largest representatives. As mentioned, the southeastern part of Europe was less affected by the first wave, but heavily affected by the second wave. Cluster 3, showing the opposite behaviour, covers the whole western part of Europe, from Denmark to Italy, with France, UK, and Italy being its three largest representatives. It is worth noting two exceptions: on one hand, Portugal was warned by the terrible course of the pandemic in its only neighbouring state Spain, and therefore imposed lockdown measures at a very early stage [26], keeping the incidences and the death rate relatively low during the first wave ($\alpha_1 = 4.3\%$). On the other hand, Hungary had low incidence numbers, but an unusually high cCFR of $\alpha_1 = 13.7\%$. At least partially, this might be explained by a low test rate in Hungary during these months [27].

The questions arises as to why certain geographical regions in Europe follow certain similar patterns of viral spread and CFRs. Recent data point towards a link between

national gross domestic product (GDP), tourism and COVID-19 cases during the initial phase of the pandemic. Countries with a high GDP as well as high tourist activity had more COVID-19 cases during the initial first wave than countries with a lower GDP and lower tourist activity [28]. These findings support our data and are in keeping with our three geographical clusters. Clusters 2 and 3 representing the Central and Western parts of Europe, with higher GDPs on average as well as high levels of tourism as well as business travel facilitating the initial spread. Cluster 1 representing the southeastern parts of Europe with a relatively low average GDP suffered tremendously from the second wave after it had been spared from the first wave due to low tourist and business travel activity. Countries with higher GDP such as clusters 1 and 2 were less affected by the second wave as health care systems were able to adapt, and medical innovations could be widely applied and resources distributed.

Interestingly, the weather seems to have an impact on viral spread and COVID-19 [29,30]. Data from Spain, which is part of our “Western” cluster 3, show that UV-radiation during the winter months of 2019–2020 seemed to be inversely correlated with the number in infections in the spring of 2020. It remains unclear whether this effect is caused by direct effects of UV-radiation on virus replication or whether it is an effect of increased immunity by higher Vitamin-D levels [31]. The weather-hypothesis has been studied elsewhere. Incidences were much higher in the North of Italy for example than in southern Italian regions. Again, UV-radiation from North to South correlated inversely with COVID-19 [32]. Our cluster 3 is mainly composed of “southwestern” European countries defined by high UV-radiation and a dry southern climate which might partly explain the relatively mild second wave with low CFRs.

We already elucidated how the fit of our three-phase-model implies the existence of a threshold value for each country individually at the beginning of the second wave. As long as the incidences stayed below this threshold value during the summer months, changes in COVID-19 incidences did not affect the number of related deaths in the same way as during the first and second wave.

The aim of the multiple regression model was to find out if it is possible to predict the threshold value, where the COVID-19 incidences and the death rates re-couple leading to increasing death rates and marking the beginning of a new wave. The multiple regression model has an R^2 value of $R^2 = 85.0\%$, which means that the model can explain 85% of the variation in the data (see Figure 3a). The comparison with the one-dimensional model (see Figure 3b) that exhibited an R^2 value of $R^2 = 77.2\%$ proves two statements: first, the logarithm of the population ($\ln(P)$) is the dominant factor for the prediction of the threshold. Second, adding the factors α_1 and p_{2rel} improves the model. This is also reflected by the significant effect tests (see Table 4). The fact that the parameter estimates for $\ln(P)$ are close to 1 (0.99 in the simple regression, 0.86 in the multiple regression) is highly plausible and implies the following interpretation: a 10-fold increase in population size leads to a ($10^{0.86}$) 7.24-fold increase in the threshold value. The two other factors in the multiple regression model, α_1 and p_{2rel} , can be interpreted as measurements of the severity of the pandemic before the second wave. A positive sign of these factors indicates the degree of preparedness: a more extended first wave, and a summer with slightly higher incidences, always staying under the predicted coupling threshold, might have produced higher rates of immunity within communities, potentially decreasing the severity of the following second wave and delaying the re-coupling of incidences and rising death rates.

We can draw another conclusion from Figure 3b by looking at the size and color of the points. The color represents the cluster, the size corresponds to the cCFR α_2 . There is no structure visible with respect to the color or size of the points. This means that the linear connection expressed by the regression model for the prediction of the threshold is independent of the other factors. It holds through all clusters and for any severity of the second wave, expressed by α_2 .

Our study has limitations. Our analyses are based on observational data for 35 European countries provided by the ECDC. The definition of a COVID-19 related death might

vary from country to country, as might the reliability of the reporting itself. In addition, it is important to keep in mind that we analyse case fatality rates and not the mortality of the novel coronavirus. Due to the high number of unreported cases, the cCFR are an overestimation of mortality. The analysis of excess mortality might provide additional information about the threat that COVID-19 poses for the population [5,33,34]. The testing strategies vary a lot from country to country, and changed within each country over the course of the pandemic. This introduces a source of variation to the incidence data that we did not control for.

In addition, our predicted thresholds are estimated ex-post. Further investigation is needed to assess the predictive power for future data. Another interesting aspect could be the analysis of possible interactions between, for instance, countries in the same cluster. Two other approaches could be used to study COVID-19 time series on a European level in further studies: entropy ratios [35] and a detrended cross-correlation analysis [36]. Finally, results of time series analysis might facilitate predicting the peak of the epidemic in terms of new cases or new deaths in each country [37].

Recently, mutations of the corona virus have been detected in many European countries. At this point, we cannot predict if our three-phase model could be adjusted to the seemingly more aggressive spread of newer virus mutations which seem to elucidate an entirely new dynamic in terms of infectiousness as well as aggressiveness of the resulting COVID-19.

6. Conclusions

Our results show that, in 32 out of 35 European countries, the COVID-19 pandemic can be accurately depicted by a three-phase-model. In the first and third phase, the corrected case fatality rates were coupled to the incidences. However, in the second phase, we see a random unconnected behaviour of incidences and deaths indicating an uncoupling of COVID-19 related deaths with incidences as long as the incidences stay below a specific threshold.

Three distinct clusters were obtained when clustering the parameters of the three-phase-model: The first cluster in the south east of Europe, with relatively low cCFRs in the first wave and high cCFRs in the second wave. A second cluster in the centre of Europe, with low cCFRs in the first wave, and slightly higher cCFRs in the second wave, and a third cluster in the western part of Europe, with very high cCFR during the first wave, and low cCFRs in the second wave.

Finally, a strong linear relationship was found for the prediction of the threshold value at the beginning of the second wave, where the incidences and the COVID-19 related death rates reconnected. Using the size of the population, the corrected case fatality rate from the first wave, and the relative death rate during the second phase, the model predicts a 7.24-fold increase of the threshold for a 10-fold increase of the population. Below this country-specific threshold, an increase in COVID-19 incidences did not lead to a significant rise of COVID-19 related deaths.

Author Contributions: Conceptualization: D.M., M.N.-M., S.B.; methodology: D.M., S.B.; software: D.M., S.B.; validation: D.M., M.N.-M., S.B.; formal analysis: D.M., M.N.-M., S.B.; investigation: D.M., M.N.-M., S.B.; data curation: D.M., M.N.-M., S.B.; writing—original draft preparation: D.M.; writing—review and editing: D.M., M.N.-M., S.B.; visualization: D.M., S.B.; supervision: M.N.-M., S.B.; project administration: S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available from the website of the European Centre for Disease Prevention and Control [1].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. European Centre for Disease Prevention and Control. *Geographical Distribution of 2019-nCov Cases Globally 2020*; ECDC: Stockholm, Sweden, 2020. Available online: <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases> (accessed on 24 December 2020).
2. Romero-Severson, E.O.; Hengartner, N.; Meadors, G.; Ke, R. Change in global transmission rates of COVID-19 through May 6 2020. *PLoS ONE* **2020**, *15*, e0236776. [[CrossRef](#)]
3. Borgmann, S.; Meintrup, D.; Reimer, K.; Schels, H.; Nowak-Machen, M. Incidence and Death Rates from COVID-19 Are Not Always Coupled: An Analysis of Temporal Data on Local, Federal, and National Levels. *Healthcare* **2021**, *9*, 338. [[CrossRef](#)]
4. Hauser, A.; Counotte, M.J.; Margossian, C.C.; Konstantinoudis, G.; Low, N.; Althaus, C.L.; Riou, J. Estimation of SARS-CoV-2 mortality during the early stages of an epidemic: A modeling study in Hubei, China, and six regions in Europe. *PLoS Med.* **2020**, *17*, e1003189. [[CrossRef](#)] [[PubMed](#)]
5. Fouillet, A.; Pontais, I.; Caserio-Schönemann, C. Excess all-cause mortality during the first wave of the COVID-19 epidemic in France, March to May 2020. *Euro Surveill.* **2020**, *25*, 2001485. [[CrossRef](#)]
6. Rotejanaprasert, C.; Lawpoolsri, S.; Pan-Ngum, W.; Maude, R.J. Preliminary estimation of temporal and spatiotemporal dynamic measures of COVID-19 transmission in Thailand. *PLoS ONE* **2020**, *15*, e0239645. [[CrossRef](#)] [[PubMed](#)]
7. Scortichini, M.; Schneider Dos Santos, R.; De' Donato, F.; De Sario, M.; Michelozzi, P.; Davoli, M.; Masselot, P.; Sera, F.; Gasparrini, A. Excess mortality during the COVID-19 outbreak in Italy: A two-stage interrupted time-series analysis. *Int. J. Epidemiol.* **2021**, *49*, 1909–1917. [[CrossRef](#)]
8. Tang, Y.; Serdan, T.D.A.; Masi, L.N.; Tang, S.; Gorjao, R.; Hirabara, S.M. Epidemiology of COVID-19 in Brazil: Using a mathematical model to estimate the outbreak peak and temporal evolution. *Emerg. Microbes Infect.* **2020**, *9*, 1453–1456. [[CrossRef](#)] [[PubMed](#)]
9. Wang, K.; Ding, L.; Yan, Y.; Dai, C.; Qu, M.; Jiayi, D.; Hao, X. Modelling the initial epidemic trends of COVID-19 in Italy, Spain, Germany, and France. *PLoS ONE* **2020**, *15*, e0241743. [[CrossRef](#)]
10. Li, H.; Xu, X.L.; Dai, D.W.; Huang, Z.Y.; Ma, Z.; Guan, Y.J. Air pollution and temperature are associated with increased COVID-19 incidence: A time series study. *Int. J. Infect. Dis.* **2020**, *97*, 278–282. [[CrossRef](#)] [[PubMed](#)]
11. Elgar, F.J.; Stefaniak, A.; Wohl, M.J.A. The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries. *Soc. Sci. Med.* **2020**, *263*, 113365. [[CrossRef](#)] [[PubMed](#)]
12. Gittins, M.; Ashton, C.; Holden, N.; Cross, S.; Meadipudi, S.; Kawafi, K.; Burger, I.; Rickard, S.; Vail, A.; Molloy, J.; et al. Environmental Factors and Hyperacute Stroke Care Activity During the COVID-19 Pandemic: An Interrupted Time-Series Analysis. *J. Stroke Cerebrovasc. Dis.* **2020**, *29*, 105229. [[CrossRef](#)]
13. Flaxman, S.; Mishra, S.; Gandy, A.; Unwin, H.J.T.; Mellan, T.A.; Coupland, H.; Whittaker, C.; Zhu, H.; Berah, T.; Eaton, J.W.; et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **2020**, *584*, 257–261. [[CrossRef](#)]
14. Pchetti, M.; Marini, B.; Giudici, F.; Benedetti, F.; Angeletti, S.; Ciccozzi, M.; Masciovecchio, C.; Ippodromo, R.; Zella, D. Impact of lockdown on Covid-19 case fatality rate and viral mutations spread in 7 countries in Europe and North America. *J. Transl. Med.* **2020**, *18*, 338. [[CrossRef](#)]
15. Soltesz, K.; Gustafsson, F.; Timpka, T.; Jaldén, J.; Jidling, C.; Heimerson, A.; Schön, T.B.; Spreco, A.; Ekberg, J.; Dahlström, Ö.; et al. The effect of interventions on COVID-19. *Nature* **2020**, *588*, E26–E28. [[CrossRef](#)]
16. Kobayashi, T.; Jung, S.M.; Linton, N.M.; Kinoshita, R.; Hayashi, K.; Miyama, T.; Anzai, A.; Yang, Y.; Yuan, B.; Akhmetzhanov, A.R.; et al. Communicating the Risk of Death from Novel Coronavirus Disease (COVID-19). *J. Clin. Med.* **2020**, *9*, 580. [[CrossRef](#)]
17. Everitt, B.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; Wiley: Chichester, UK, 2011.
18. Wikipedia, T.F.E. List of Sovereign States in Europe by GDP (Nominal). 2021. Available online: [https://en.wikipedia.org/wiki/List_of_sovereign_states_in_Europe_by_GDP_\(nominal\)](https://en.wikipedia.org/wiki/List_of_sovereign_states_in_Europe_by_GDP_(nominal)) (accessed on 22 March 2021).
19. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. 2021. Available online: <https://covid19.who.int/> (accessed on 13 April 2021).
20. JMP. *Version 15.2.0*; SAS Institute Inc.: Cary, NC, USA, 2020. Available online: <https://www.jmp.com> (accessed on 13 April 2021).
21. Helsingen, L.M.; Refsum, E.; Gjøstein, D.K.; Løberg, M.; Bretthauer, M.; Kalager, M.; Emilsson, L. The COVID-19 pandemic in Norway and Sweden - threats, trust, and impact on daily life: A comparative survey. *BMC Public Health* **2020**, *20*, 1597. [[CrossRef](#)]
22. Kavaliunas, A.; Ocaya, P.; Mumper, J.; Lindfeldt, I.; Kyhlstedt, M. Swedish policy analysis for Covid-19. *Health Policy Technol.* **2020**, *9*, 598–612. [[CrossRef](#)]
23. Percivalle, E.; Cambiè, G.; Cassaniti, I.; Nepita, E.V.; Maserati, R.; Ferrari, A.; Di Martino, R.; Isernia, P.; Mojoli, F.; Bruno, R.; et al. Prevalence of SARS-CoV-2 specific neutralising antibodies in blood donors from the Lodi Red Zone in Lombardy, Italy, as at 6 April 2020. *Euro Surveill* **2020**, *25*, 2001031. [[CrossRef](#)] [[PubMed](#)]
24. Havers, F.P.; Reed, C.; Lim, T.; Montgomery, J.M.; Klena, J.D.; Hall, A.J.; Fry, A.M.; Cannon, D.L.; Chiang, C.F.; Gibbons, A.; et al. Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23–May 12, 2020. *JAMA Intern. Med.* **2020**, *180*, 1576–1586. [[CrossRef](#)]
25. Bobrovitz, N.; Arora, R.K.; Cao, C.; Boucher, E.; Liu, M.; Rahim, H.; Donnici, C.; Ilincic, N.; Duarte, N.; Van Wyk, J.; et al. Global seroprevalence of SARS-CoV-2 antibodies: A systematic review and meta-analysis. *medRxiv* **2020**. [[CrossRef](#)]

26. Ricoca Peixoto, V.; Vieira, A.; Aguiar, P.; Carvalho, C.; Rhys Thomas, D.; Abrantes, A. Initial Assessment of the Impact of the Emergency State Lockdown Measures on the 1st Wave of the COVID-19 Epidemic in Portugal. *Acta Med. Port.* **2020**, *33*, 733–741. [[CrossRef](#)]
27. Röst, G.; Bartha, F.A.; Bogyta, N.; Boldog, P.; Dénes, A.; Ferenci, T.; Horváth, K.J.; Juhász, A.; Nagy, C.; Tekeli, T.; et al. Early Phase of the COVID-19 Outbreak in Hungary and Post-Lockdown Scenarios. *Viruses* **2020**, *12*, 708. [[CrossRef](#)] [[PubMed](#)]
28. Skórka, P.; Grzywacz, B.; Moroń, D.; Lenda, M. The macroecology of the COVID-19 pandemic in the Anthropocene. *PLoS ONE* **2020**, *15*, e0236856. [[CrossRef](#)]
29. Lin, S.; Fu, Y.; Jia, X.; Ding, S.; Wu, Y.; Huang, Z. Discovering Correlations between the COVID-19 Epidemic Spread and Climate. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7958. [[CrossRef](#)]
30. Agapito, G.; Zucco, C.; Cannataro, M. COVID-WAREHOUSE: A Data Warehouse of Italian COVID-19, Pollution, and Climate Data. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5596. [[CrossRef](#)]
31. Cacho, P.M.; Hernández, J.L.; López-Hoyos, M.; Martínez-Taboada, V.M. Can climatic factors explain the differences in COVID-19 incidence and severity across the Spanish regions?: An ecological study. *Environ. Health* **2020**, *19*, 106. [[CrossRef](#)]
32. Isaia, G.; Diémoz, H.; Maluta, F.; Fountoulakis, I.; Cecon, D.; di Sarra, A.; Facta, S.; Fedele, F.; Lorenzetto, G.; Siani, A.M.; et al. Does solar ultraviolet radiation play a role in COVID-19 infection and deaths? An environmental ecological study in Italy. *Sci. Total Environ.* **2021**, *757*, 143757. [[CrossRef](#)] [[PubMed](#)]
33. Vestergaard, L.S.; Nielsen, J.; Richter, L.; Schmid, D.; Bustos, N.; Braeye, T.; Denissov, G.; Veideman, T.; Luomala, O.; Möttönen, T.; et al. Excess all-cause mortality during the COVID-19 pandemic in Europe—preliminary pooled estimates from the EuroMOMO network, March to April 2020. *Euro Surveill.* **2020**, *25*, 2001214. [[CrossRef](#)] [[PubMed](#)]
34. Michelozzi, P.; de'Donato, F.; Scortichini, M.; Pezzotti, P.; Stafoggia, M.; De Sario, M.; Costa, G.; Noccioli, F.; Riccardo, F.; Bella, A.; et al. Publisher Correction to: Temporal dynamics in total excess mortality and COVID-19 deaths in Italian cities. *BMC Public Health* **2020**, *20*, 1325. [[CrossRef](#)] [[PubMed](#)]
35. Bandt, C. Entropy Ratio and Entropy Concentration Coefficient, with Application to the COVID-19 Pandemic. *Entropy* **2020**, *22*, 1315. [[CrossRef](#)]
36. Contreras-Reyes, J.E.; Idrovo-Aguirre, B.J. Backcasting and forecasting time series using detrended cross-correlation analysis. *Phys. A Stat. Mech. Appl.* **2020**, *560*, 125109. [[CrossRef](#)]
37. Li, S.; Linton, O. When will the Covid-19 pandemic peak? *J. Econom.* **2021**, *220*, 130–157. [[CrossRef](#)]