Analysis of Mobility Data in Relation to the COVID-19 Pandemic

Lavinia Hriscu, Guillem Álvarez, and Oriol Bertomeu*
Enginyeria Física, Universitat Politècnica de Catalunya
(Dated: June 2020)

Abstract: In this paper we study the correlation between the measures taken by different nations in regard to the mobility of their citizens and the evolution of the ongoing COVID-19 pandemic over the process of control period. To do so, mobility indices provided by *Facebook Data for Good* are analysed in order to understand the different impacts of actions taken by governments.

I. INTRODUCTION

The COVID-19 pandemic has been caused by SARS-CoV-2. The first case known of COVID-19 was on November 2019, in China. Coronavirus is known to cause respiratory infections on humans. Specifically, the COVID-19 disease produces, in most cases, fever, dry cough, shortness of breath and tiredness. Less common symptoms are aches and pains, chest pressure, nasal congestion and headache, among others. Although anyone is liable to become seriously ill, the evolution of the disease is gradual and different for each person. While some only experience mild symptoms, others require hospital treatment (mostly older people and those with other medical problems).

COVID-19 disease is highly contagious and spreads easily through small droplets from nose or mouth coming from a person infected when speaking, coughing or sneezing [**1**]. Its explosive expansion throughout the world has changed our lifes in just a few months. Not only a huge number of cases have been diagnosed, but also daily habits and actions have been modified in order to prevent the spread of the disease.

This change in daily habits is due to the fact that pharmaceutical interventions are not available right now and herd immunity has not proved itself effective in fighting the pandemic yet. Therefore, nonpharmaceutical measures such as mobility reduction are really important in order to decrease the total amount of direct contacts between population and by this, the propagation of the virus.

Government authorities are the ones who establish the measures each country follows to face the pandemic. Using public data provided by the Security European Commission Joint Research Centre [**2**], we found the dates in which important decisions regarding mobility and confinement were taken (see TABLE I). We have focused on four countries of special interest that have been diversely impacted by the pandemic and that have taken contrasting measures, ranging from the severe confinement imposed on Spain and Italy to the relatively lax actions of the Danish and Swedish governments.

* We want to thank the BIOCOM-SC group at UPC, especially Enric Álvarez-Lacalle, Clara Prats and Miquel Marchena, and the Data for Good team at Facebook that provided the mobility data.

Country	Spain	Italy	Sweden	Denmark
School closures	14/03	04/03	16/03	18/03
National movement	13/04	28/03	25/03	18/03
International movement	14/03	28/03	14/03	13/03
Flight restrictions	14/03	28/03	18/03	N/A
Shop closures	14/04	11/03	N/A	18/03
Events stop	14/03	11/03	12/03	18/03

Colour legend
No restrictions
Partial restrictions
Full restrictions

TABLE I: Governamental restriction measures.

II. METHODS

Since the aim of the analysis is to find a relation between the citizens' mobility index and the number of cases, all data has been collected from different sources, then mathematical models have been applied.

A. Mobility Data

The assessment of mobility data has been done through an agreement between the Universitat Politècnica de Catalunya (UPC) research group of Computational Biology and Complex Systems (BIOCOM-SC) and Facebook (Data for Good). Information on people's movement and lockdown indices has been provided at provincial level.

1. *Mobility Indices*

Two mobility indices are used in this study:

- Short-Walk Confinement Index: Contains information about the number of people that stays confined, where staying confined means being enclosed in predefined tiles of an area of about $0.3km^2$.
- Mobility Index: Gives information of the total mobility by comparing the total number of visited tiles with respect to a baseline. This index gives a relative value in comparison with the corresponding value for a regular/non-confinement day.

The aim of our study is to characterize group behaviour: at no point we have had access to information about individual users. To reinforce that, knowledge on the position or path of a single user is not possible to obtain by any means, since the indices refer to provinces and give us information over every province's total population. Therefore, no individual tracking is possible, and privacy is not violated in any terms.

Territorial scales for mobility indices are given at provincial level for Spain, Denmark and Italy, whereas for Sweden the data is at county (*län*) level. Our interest in this study is to obtain knowledge at country and regions level. In order to do so, a weighted mean over data indices at province level is carried out. The weights are calculated from the percentage of population of each province/county over the total population of the country.

B. Number of Cases

For the number of confirmed cases at each country we have used public data from the European Centre for Disease Prevention and Control [**3**].

From this data ρ index is obtained. This index is related with the number of new infections caused by a single case and gives us important information about the spread level of the pandemic, as if $\rho > 1$ the propagation is increasing, while when $\rho < 1$ means that the propagation of the virus is slowing down and the number of new cases decreases. Its value for a specific day is calculated as follows:

$$\rho(t-1) = \frac{N_{new}(t) + N_{new}(t-1) + N_{new}(t-2)}{N_{new}(t-5) + N_{new}(t-6) + N_{new}(t-7)}$$

Where $\rho(t-1)$ stands for the results of ρ for the $(t-1)$th day and $N_{new}(t)$ corresponds to the number of new confirmed cases at day t.

Working with the data the presence of many sources of noise that affect the results have been noticed. As a first approach to reduce it, a Gaussian-weighted moving average filter over a window of 4 days has been applied on $\rho(t-1)$ calculation in order to smooth the results.

Another important fact to be taken into account is that there is a delay between when an infected individual first experiences symptoms and the day in which the case is confirmed and reported. This delay has been characterized and depends on the country [**4**]. Moreover, in weekends there is a lack of reporting, which leads to an unrealistic decline of confirmed cases, producing artificial oscillations on the data. In contrast, mobility may increase with respect to business days. To deal with these issues, averages over time have been taken weekly. Therefore, working with a new index ρ_7, a more reliable evolution of the pandemic is obtained (see results in Section III C).

In the calculation of ρ_7 entire weeks were required to obtain a single result. Depending on the country the beginning of data was found in a different weekday. For cases in which the first day was not Monday, we included the data of those with at least 5 days in order to have trustworthy information. Otherwise, the data was neglected as the mean over less than five days was considered not reliable enough due to the artificial oscillations from data's noise.

III. RESULTS

A. Heterogeneity Between Countries at the Beginning of the Pandemic

In order to discern the degree of similarity between different regions the beginning of the pandemic is studied. To this end, an exponential fitting of the expansion period has been carried out for each country. In FIG.1 we find the graphical results representing the number of cases with respect to time (days).

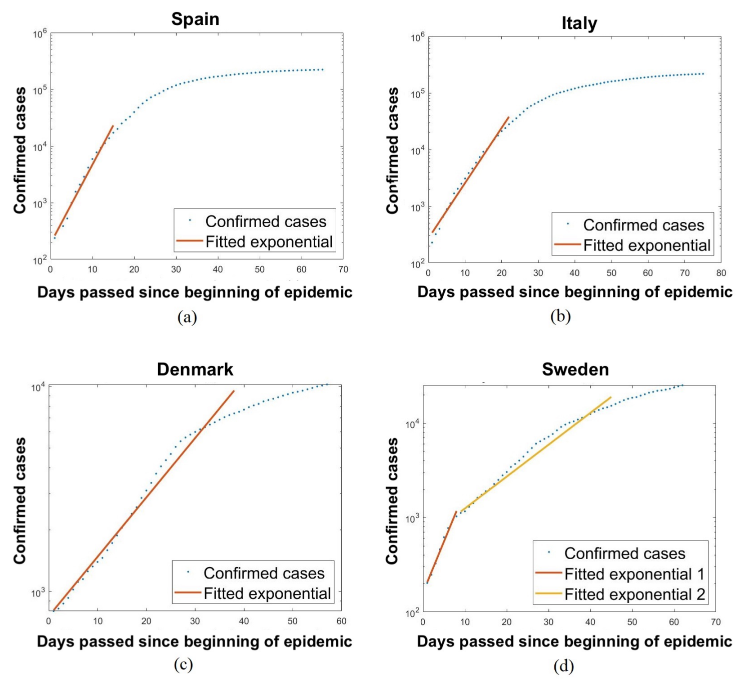

FIG. 1: Exponential fitting of the evolution of the number of cases.

The beginning of the epidemic in each country is established as the day in which the cumulative cases overcome an amount of 200. From this date, an exponential fitting is carried out for the maximum possible number of days accomplishing $R^2 > 0.98$, where R^2 is the accuracy coefficient of the exponential fitting. Numerical values for exponential's time scale (τ), R^2 and the number of days in which an exponential fitting is carried out (T) are found in TABLE II.

To better understand the results, we have to understand what does each parameter mean. One one hand, τ is the exponential's time scale. Therefore, for lower τ

	$\tau(days)$	R^2	T(days)
SPAIN	3.1272	0.9816	15
ITALY	4.4578	0.9811	22
DENMARK	15.0148	0.9804	38
SWEDEN 1	4.0286	0.9803	8
SWEDEN 2	12.8539	0.9813	45

TABLE II: Results of the exponential fitting.

values a higher growth ratio of propagation of the virus is found, and the other way around. On the other hand, T is the maximum number of days in which the exponential fitting can be carried out while accomplishing the accuracy condition $R^2 > 0.98$. Therefore, T can be understood as the duration of the exponential growth period.

From TABLE II it can be stated that Spain had the most drastic growth of the four countries analysed, as it has the lowest τ value for the fitting done over a period of 15 days. For Italy and Spain exponential period resemblance was found: Both have a clear exponential expansion period for two-three weeks and later on the propagation diminishes and is no longer exponential.

Meanwhile, the cases of Sweden and Denmark differ. In order to obtain the exponential fitting for Denmark the first day had to be obtained by stablishing more than 700 cases reached (instead of 200 cases), as for the firsts days (from 200 to 700 cases) the behaviour of the data analysed could not be set as an exponential. Moreover, the results on the exponential fitting of Sweden gave us a peculiar result, as two different exponentials could be fitted (Sweden 1 and Sweden 2 on TABLE II), with a first exponential that is more severe than the second: the second value of τ is more than thrice that of the first value (see TABLE II).

Adding both exponential growth periods (T) we have a total of 53 days with exponential growth for Sweden. For Denmark it takes the value $T = 38$. Therefore it is clear to state that exists an hetereogenity on the propagation of the pandemic for different regions. For Spain and Italy the initial propagation rate is clearly exponential, but the period in which is found is smaller compared with the situation found in Denmark and Sweden.

Regarding the results obtained for Sweden, some factors need to be mentioned to understand the results. On one hand, more permissive measures could be logically related to the larger propagation period observed. On the other hand, the obtention of two different exponential regimes could be related with the different way of reporting data, which varies in different countries, and the criteria may have changed.

Focusing on this idea differences can depend on many factors, such as social, economic and political factors. Anyhow, in the following sections we prove that mobility is a transcendental factor to be taken into account.

B. Heterogeneity Between Countries in Mobility

From TABLE I we can deduce that mobility indices differ between countries because the restriction measures each government implemented are different. In particular, the Swedish government applied less restrictions than the Spanish and Italian ones. Then, it is interesting to compare the variation of the two mobility indices, one with respect to the other, for the four countries analysed.

In FIG.2 the correlation of mobility indices is compared. At first sight, the indices are correlated in all countries. In Spain and Italy strong measures of confinement were applied and consequently a high reduction in both Mobility and Short-Walk Confinement indices appears. In contrast, the less restrictive measures applied in Denmark and Sweden reflect a low decrease in mobility.

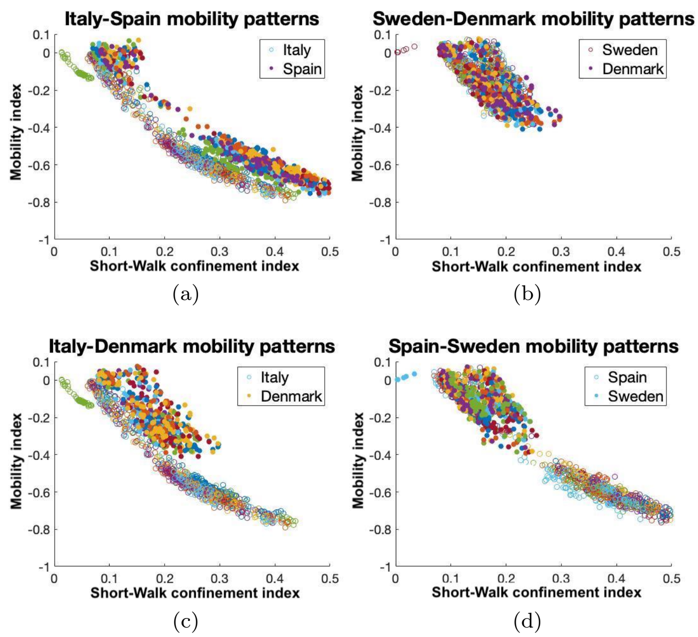

FIG. 2: Comparison between Short-Walk Confinement Index and Mobility Index. Each dot represents the value of both indices in a single day.

C. Weekly Evolution of the Pandemic

It is interesting to take a look at the evolution of the growth rate, the Short-Walk Confinement Index and the Mobility Index averaged over 7 days. This average is due to the fact that, as explained in Section II B, it allows us to smooth out artificial oscillations due to the noise of the data. In order to perform this analysis, a weekly average of each magnitude has been calculated. By doing this we obtain a better picture of the evolution of the pandemic for the countries studied and the differences among them.

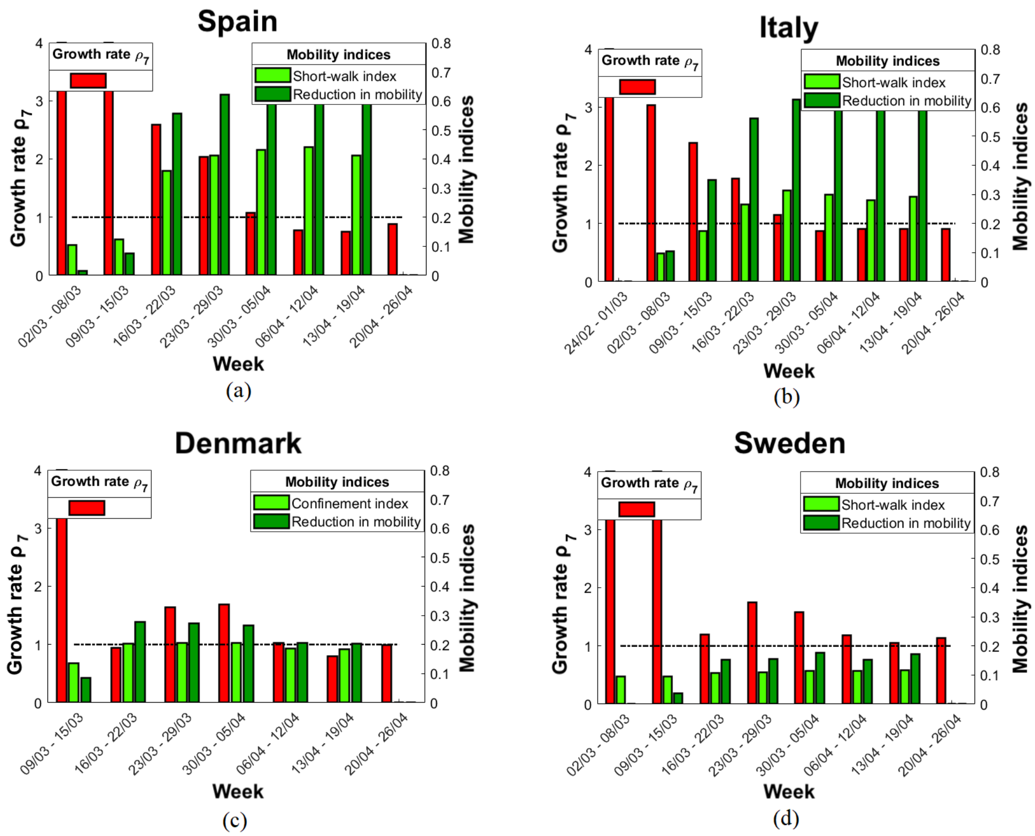

FIG. 3: Evolution of the weekly averages of ρ and the mobility indices.

We have restricted the value of ρ_7 to be found between 0 and 4 to remove the unrealistic high quantities obtained in the very first weeks of the pandemic. These abnormalities happened due to the poor testing capabilities at that stage, which led to large backlogs of cases that were subsequently reported in short spans of time.

From FIG.3 we can see how different levels of confinement affect the virus' propagation. On one hand, we find the strict lockdowns imposed by Spain and Italy in order to limit the spread of the virus by restricting mobility. On the other hand, we have the more relaxed measures imposed in Sweden. They are characterized by avoiding confining large sections of the population, locking down only the most vulnerable so as to avoid harming the economy. Meanwhile, Denmark has attempted to chart a middle course between these two extremes, as it can be seen on the measures in TABLE I.

Indeed, the graphs show much sharper reductions in Mobility Index and increases in the Short-Walk Confinement Index (indicating the level of confinement) for Spain and Italy. Meanwhile, mobility reduction has been low in Sweden, with mobility not having gone below 80% with respect to that of a normal day. Taking into account the more severe state in which Italy and Spain found themselves before the lockdowns, the measures taken have been more successful in cutting ρ_7 below one and by this stopping the propagation of COVID-19, while Sweden struggles to contain the pandemic as its ρ_7 value is still over 1 up to when this article is written (June 2020), as can be seen in TABLE III.

This result relays a very noteworthy fact, as it verifies that reduction of mobility has been proved a truly effective measure to fight the COVID-19 propagation.

Country	Week when $\rho_7 < 1$
Italy	30/03 - 05/04
Spain	06/04 - 12/04
Denmark	13/04 - 19/04
Sweden	N/A

TABLE III: Week in which $\rho < 1$ is achieved.

D. ρ vs Mobility with Diagnostic Delay

In this section the impact of mobility in the growth rate of the virus is further developed. This time, a representation of the growth rate ρ_7 and the two mobility parameters for Spain, Italy, Sweden and Denmark has been obtained in a scatter plot. A time shift is applied in order to introduce the effect of the diagnostic delay

(DD), which is the estimated delay period for the reported cases and depends on the country being studied. It is established from [**4**] and the values appear in TABLE IV. This shifts ρ values the number of days corresponding to DD. The results obtained are observed in FIG.4.

Country	DD (days)
Spain	14
Italy	14
Sweden	5
Denmark	14

TABLE IV: Diagnostic delay for the countries of interest.

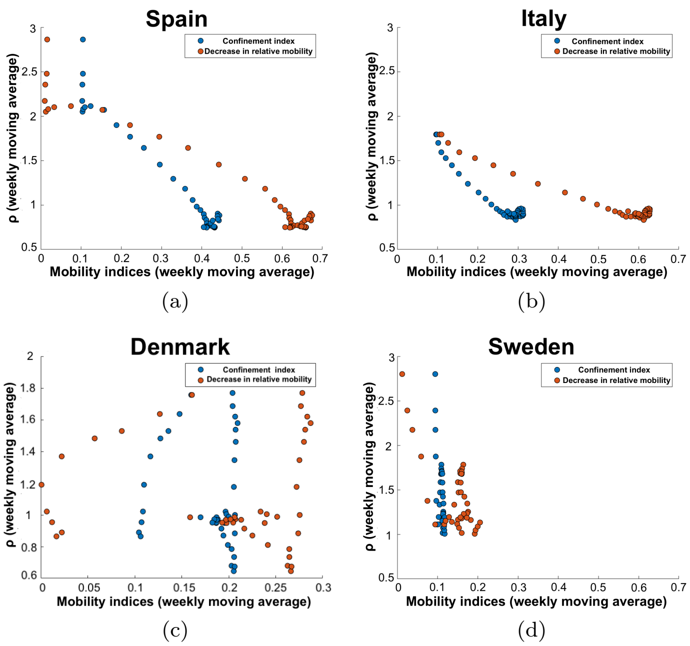

FIG. 4: Representation of ρ_7 with diagnostic delay against the mobility indices.

IV. DISCUSSION

From our study we obtained many interesting conclusions. First of all, we confirmed the point that we wanted to prove: mobility is a really important factor to take into account in order to regulate the pandemic's propagation.

We can clearly confirm this from the results of FIG.3 in Section III C. When mobility is strictly reduced, as in Spain and Italy, the goal of achieving $\rho < 1$ is reached, slowing down the propagation of the virus. Meanwhile, for more lax measures like the ones taken in Sweden and Denmark (see TABLE I), mobility is not drastically reduced, which implies $\rho > 1$. Therefore, the propagation of the virus is not controlled.

Refering to Section III D, the application of severe confinement as in Italy and Spain in order to reduce mobility and prevent the spread of the virus, produces a great deal of correlation between ρ and mobility, since the scatter plots are close to straight lines (see FIG.4 (a) and (b)).

That does not happen for the northern countries, where the results show a much less clear level of correlation between the measures taken and their effect on reducting the propagation of the pandemic. As can be seen in FIG. 4 (c) and (d), the level of correlation between the measures taken and ρ_7 in Sweden and Denmark is not comparable to the one reached in Spain and Italy. This shows again the relevance of mobility in the pandemic regulation and the effectiveness of more restricive measures.

In addition, heterogeneity on the effects of COVID-19 over different regions has been demostrated. This is seen in Section III A (for the pandemic's beginning) and in FIG.3 and FIG.4 for the evolution of the pandemic and the effect of mobility restrictions on its development.

Another point to reinforce the heterogeneity between countries is found in FIG. 2 from Section III B, where the scattered points for mobility parameters have a resemblance for Spain and Italy, but are in general quite diverse. From this it is also seen that the relation between Short-Walk Confinement Index and Mobility Index changes among countries, indicating that the social mobility situation varies depending on the country, even for countries with similar measures taken.

Keeping this idea in mind, it can be stated that the different evolution of the pandemic is related to different social, economic and political situations of each region. In this sense, we clearly obtained remarkably different results from the different strategies implemented by governments to deal with the pandemic.

Concerning Sweden, it is still unclear whether its strategy will succeed, as it consists on confining only the most vulnerable sectors of the population to slow down the propagation of the virus, and try to gradually achieve group immunity while avoiding significant harm to the economy. As of now, in Sweden death rates are still increasing, overcoming levels achieved by their neighboring countries [**5**].

As a conclusion on our study, the results that we intended to prove have been demonstrated: we can conclude that there is an important contrast in the evolution of the pandemic between countries depending on the level of restrictions with respect to mobility and confinement measures that have been applied.

Another extra result that we had intended to obtain was to analyse the consequences of mobility on death rate. It would consist of using the number of deaths instead of the confirmed cases. This would have the advantage that data for death cases is expected to be a more robust indicator of the advance of the pandemic, because confirmed cases are influenced by the prevalence of testing and delays or troubles in reporting the information to authorities. In contrast, the number of deaths would show less noise, since most deaths by COVID-19 happen after the patient has been hospitalized already and every death can be reported immediately.

Our objective was to apply this idea on the data that we have about death cases in Spanish regions. However, we have been unable to perform this latter analysis due to the lack of reliable data: in Spain, the counting method has been changed several times, resulting in discontinuous time series, which are inadequate for investigating the spread of the pandemic.

Recently, the government has published a new and more precise list of confirmed cases. The reporting of death cases is being analysed in order to elaborate a new and more precise list, as done with the reported cases, but by the moment publication of new data for death cases in Spain has ceased.

We outline a very important fact that had to be taken into account while developing our study: the delay of the analysed data, which depends on many factors and changes over countries. It has also changed over time too, as the reporting criteria has changed along different states of the pandemic by different countries. Therefore, it is really important to study and characterize these delays in order to get more precise data [**4**].

In conclusion, it is hard to understate the importance of mobility and its study. It has been a key factor in the propagation of the pandemic. In this sense, we were able to see that something was wrong with Sweden at the beginning of the pandemic, since when we analysed the first data on mobility and compared it with the results on propagation of the virus we saw that the situation could become remarkably severe, given that mobility was not reduced enough: the country was not ready to avoid a considerable propagation of the pandemic and its harmful effects on the population.

Therefore, it can be stated that mobility is crucial in the propagation of COVID-19, and its study will be also really significant during the return to normality.

REFERENCES

1. World Health Organization. "*Q&A on coronaviruses (COVID 19)*". In: (2020). `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses`
2. Security European Commission Joint Research Centre -ISPRA - Space and Migration Directorate. "*Governmental measures database, ECML Covid*". In: (2020). `https://covid-statistics.jrc.ec.europa.eu/`
3. European Centre for Disease Prevention and Control. "*Geographic distribution of COVID-19 cases worldwide*". In: (2020) `https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide`
4. M. Català, D. Pino, M. Marchena, P. Palacios, T. Urdiales, P.J. Cardona, S. Alonso, D. López-Codina, C. Prats, E. Alvarez-Lacalle. "*Robust estimation of diagnostic rate and real incidence of COVID-19 for European policymakers*". In: (2020) `https://www.medrxiv.org/content/10.1101/2020.05.01.20087023v1`
5. M. Català, M. Marchena, P.J. Cardona, S. Alonso, D. López-Codina, C. Prats, E. Alvarez-Lacalle, D. Conesa. "*Analysis and Prediction of COVID-19 for different regions and countries*". In: (2020) `https://biocomsc.upc.edu/en/shared/20200601_report_web_77.pdf`,