Dynamics of COVID-19 Development: New Data and New Estimates from Wavelet Analysis

Vyacheslav Lyashenko

Correspondence: lyashenko.vyacheslav@gmail.com

Department of Informatics, Kharkiv National University of RadioElectronics, Ukraine

Abstract

Investigation of the dynamics of diseases from viruses is a key issue in the understanding of their distribution. This is especially important when viruses are dangerous. This requires the use of various analysis tools. We used wavelet coherence. We obtained results that explain some of the dynamics of the COVID-19 pandemic. We also conducted a comparative analysis of the development of the pandemic between individual European countries.

Keywords: Viruses, COVID-19, Wavelet Analysis, Wavelet Coherence, Statistics

Introduction

Viruses are an integral part of human life. Viruses infect all types of organisms. There are various groups and types of viruses. Viruses are the most common biological form. Viruses cause various infectious diseases (Krupovic, Dolja & Koonin, 2019; Dolgin, 2019). Among such diseases, an infection that is caused by the SARS-CoV-2 virus occupies a special place (Nickbakhsh et al., 2020; Petersen et al., 2020). This virus is of the type coronavirus. This is the most dangerous form of viruses.

Currently, the SARS-CoV-2 virus has provoked the emergence and development of the COVID-19 pandemic (Petersen et al., 2020; Spinelli & Pellino, 2020). The main danger of such a pandemic is its rapid spread and the lack of effective tools to combat such an infection. At the same time, this is complicated by possible mutations in SARS-CoV-2, which makes it difficult to quickly develop the necessary vaccine.

Thus, there is an objective need for various studies to counter coronaviruses. Among such studies, it is worth highlighting the analysis of statistical data that describe the dynamics of infection with such a virus and the dynamics of the processes of the corresponding disease. To address such issues, it is important to obtain reliable statistical data. Another important aspect in solving the question posed is the analysis of the data that are available. At the same time, the emergence of new data leads to new research in order to obtain new results to understand what is happening. Based on this, in this work, we continue the data analysis that was started in the study (Lyashenko, Saad, & Babker, 2020).

Brief Literary Review

To begin with, we will conduct a small literature review of the methods that are used to analyze statistical data on the COVID-19 pandemic. As our research shows, there are a lot of such methods (Lyashenko, Saad, & Babker, 2020; Mustafa et al., 2020; Babker, Mohammed, & Lyashenko, 2020).

Nevertheless, M. Wolkewitz and Puljak emphasize that special attention should be paid to methodological problems of data analysis on COVID-19 (Wolkewitz & Puljak, 2020). First

64

of all, this concerns the generalization of the obtained data and their generalization for different periods of time.

The study authors (Weinberger et al., 2020) also emphasize the importance of methodological aspects when analyzing data on COVID-19. In their work, they take these aspects into account when estimating the total number of deaths associated with the COVID-19 pandemic. For this, not only test results are considered. For such an analysis, the general clinical picture of the course of the disease is taken into account.

S. Deb and M. Majumdar consider the data on the COVID-19 pandemic in the form of time series (Deb & Majumdar, 2020). Then their analysis is based on time series analysis methods. This approach allows for the assessment of future trends in the incidence of COVID-19. The authors emphasize that such a statistical analysis makes it possible to understand the trends of the COVID-19 outbreak, gives an idea of what epidemiological stage the country (region) is at. They also point out that this could help develop policies to combat the COVID-19 pandemic in different countries (Deb & Majumdar, 2020).

Various methods of statistical analysis for studying data on COVID-19 are considered in their work by M. Yousaf, S. Zahir, M. Riaz, S. M. Hussain and K. Shah (Yousaf et al., 2020). For this, the authors use regression analysis methods to study the corresponding time series. Various forecasts are also built using the ARIMA model. Based on these approaches, the authors examine the future dynamics of the COVID-19 pandemic using data from Pakistan as an example.

The authors of the study (Mackey et al., 2020) use machine learning methods that can work with Big Data to analyze data on the COVID-19 pandemic. The authors use unsupervised machine learning methods. This allowed us to highlight the general characteristics of self-reported symptoms of COVID-19. As a result, clusters were obtained that combine cases of COVID-19 with identical symptoms. The authors note that this will allow for information surveillance to assess the severity of the disease COVID-19 (Mackey et al., 2020).

A. Srivastava and G. Chowell analyze the shape of the growth rate curves of data on COVID-19 (Srivastava & Chowell, 2020). This approach allows us to understand the spatial uncertainty of the COVID-19 pandemic. Various statistical methods are used to analyze the shape of the local COVID-19 growth rate curves. This makes it possible to identify different clusters that describe the same dynamics of the COVID-19 pandemic. The authors emphasize that this approach leads to a broad classification of spatial areas along different trajectories and complements the methodological toolkit for studying the COVID-19 pandemic (Srivastava & Chowell, 2020).

I. Ciufolini and A. Paolozzi used statistical and probabilistic methods to analyze data on COVID-19 (Ciufolini & Paolozzi, 2020). In particular, the authors use Gauss error function and Monte Carlo simulations. This allowed the authors to define a statistical prediction for the day when the number of daily positive cases will peak in Italy (Ciufolini & Paolozzi, 2020).

Thus, most authors conduct analyzes for univariate data that describe some parameter of the COVID-19 pandemic. At the same time, it is important to know the mutual dynamics of various parameters of the development of the COVID-19 pandemic.

Wavelet Coherence as a Tool for Analyzing Mutual Dynamics

Wavelet analysis tools can be used for mutual analysis of two time series. This tool is the wavelet coherence method. This method allows you to assess the impact of cross-referencing on the mutual dynamics of the time series being studied.

To implement wavelet coherence, the concept of cross wavelet spectra is used. Suppose we have two series of data $f(t)$ and $z(t)$ then we can determine the wavelet coherence (Torrence & Webster, 1999; Vasiurenko, Lyashenko, Baranova & Deineko, 2020):

$$R^2(a,b) = \frac{\left| S(a^{-1} V_{fz}(a,b)) \right|^2}{S(a^{-1} \left| V_f(a,b) \right|^2) S(a^{-1} \left| V_z(a,b) \right|^2)},\tag{1}$$

where: $V(a,b)$ – values of cross wavelet spectra; a – determines the size of the wavelet and is called the scale, b – specifies the temporal localization of the wavelet and is called a shift (the abscissa is the wavelet shifts (time axis – t), the ordinate is the scale; abscissa is the variable that displays the data number in the series under investigation and ordinate is the variable that displays the depth of cross-links for a time series); S – is a smoothing operator; $R^2(a,b)$ – the squared wavelet coherency coefficient. $0 \le R^2(a,b) \le 1$. If these values tend to zero, then we have a weak correlation. Otherwise, we have a strong correlation (Torrence & Webster, 1999).

Data for Analysis

For analysis, we will consider the general dynamics of diseased people, the dynamics of deaths and recovered from COVID-19 infection.

We will consider these data in the period 22.01.2020-01.11.2020 (summarized data from the site https://www.who.int/emergencies/diseases/novel-coronavirus-2019). Note that we are considering data that is presented on a cumulative basis. Figure 1 shows the dynamics of diseased people, the dynamics of deaths and recovered from COVID-19 infection.
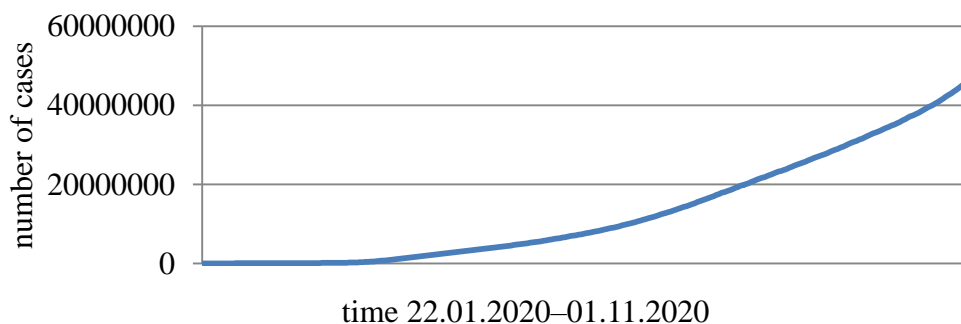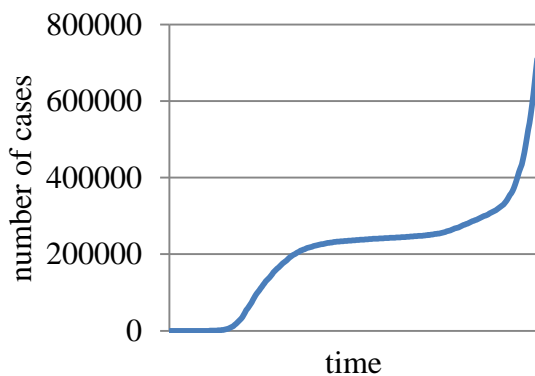

a) dynamics of diseased people
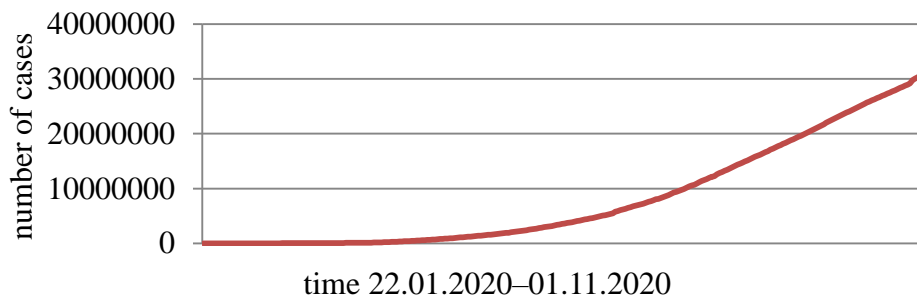
b) dynamics of recovered people
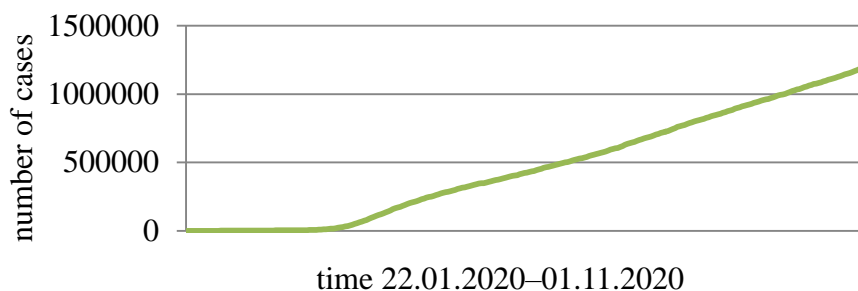
c) dynamics of deaths people

Figure 1. Dynamics of diseased people, the dynamics of recovered and deaths from COVID-19 infection (22.01.2020-01.11.2020)

We will also look at the dynamics of diseased people for some European countries. This will help us understand the mutual dynamics of the COVID-19 pandemic in different countries. Figure 2 shows the dynamics of diseased people for some European countries.
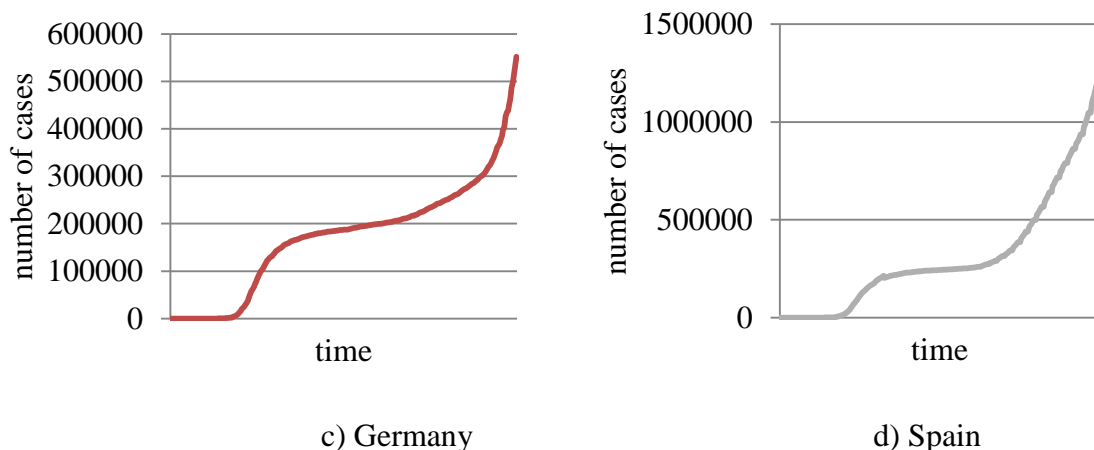

a) Italy

b) France

c) Germany   d) Spain
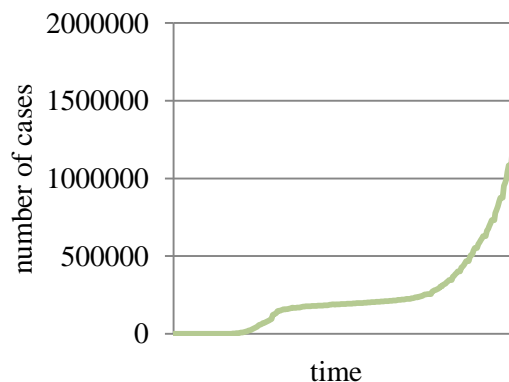
Figure 2. Dynamics of diseased people for some European countries (22.01.2020-01.11.2020)

We see that the data in Figure 1 and Figure 2 (respectively) are comparable to each other. But you can also see the differences in dynamics, which we investigate in Figures 1 or Figures 2. We will consider these differences using wavelet coherence.

Results and Discussion

Figure 3 shows the values of wavelet coherence between the total number of COVID-19 cases and the number of deaths from COVID-19 (a – 18.01.2020-01.04.2020; b – 22.01.2020-01.11.2020).
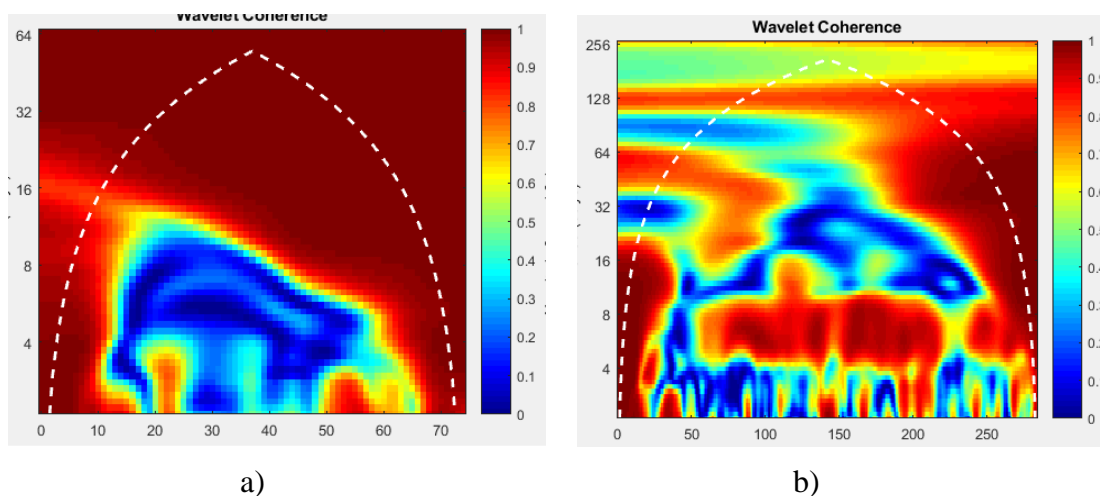

a)   b)

Figure 3. Wavelet coherence between the total number of patients with COVID-19 and the number deaths from COVID-19

Figure 4 shows the wavelet coherence between the total number of COVID-19 cases and the number of recovered from COVID-19 (a – 18.01.2020-01.04.2020; b – 22.01.2020-01.11.2020).
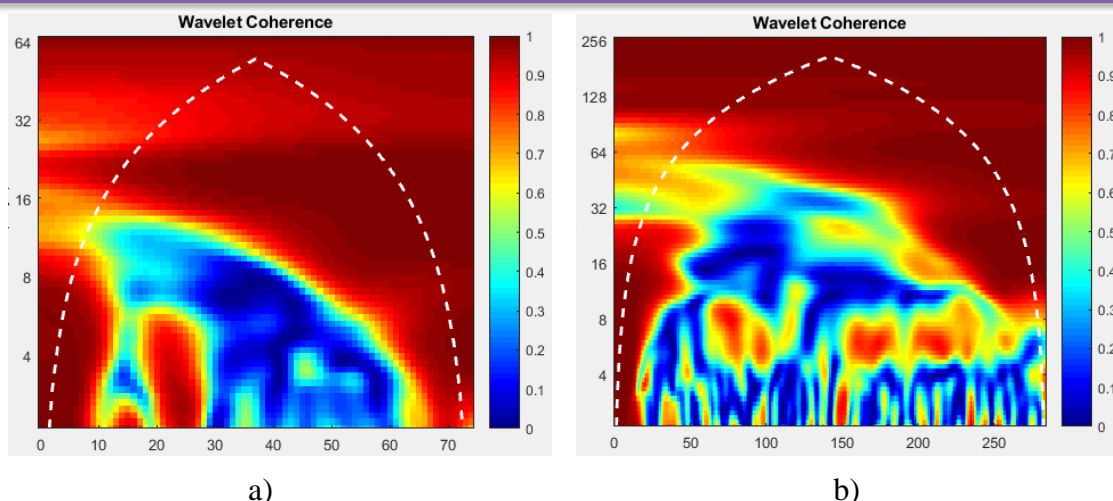
a)   b)

Figure 4. Wavelet coherence between the total number of patients with COVID-19 and the number of patients recovered from COVID-19

We see that the values of the wavelet coherence between the data that are being investigated have changed their significance during the development of the pandemic. In the period 22.01.2020-01.11.2020 (Figure 3b and Figure 4b), the values of the wavelet coherence between the data are most significant from 4 to 8 days. Consequently, patient mortality can occur after 8 days of disease detection (Figure 3b). This indicates a late visit for help or insufficiently effective treatment, a general increase in the mortality rate.

The number of days until recovery is also reduced. This suggests that the disease can be mild. But it should also be noted that this dependence is ambiguous for the entire time interval that we are studying (Figure 4b). Consequently, a number of cases of the disease are not recorded or such diseases are recorded with some time delay. Figure 5 shows the values of wavelet coherence between the total number of cases of COVID-19 in Italy and the same data in France (Figure 5a), Germany (Figure 5b) and Spain (Figure 5c), respectively.

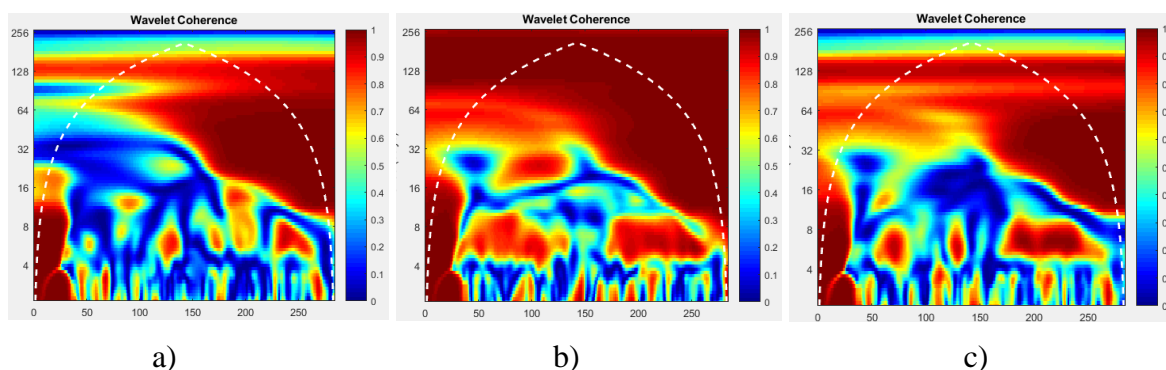a)   b)   c)

Figure 5. Values of wavelet coherence between the total number of cases of COVID-19 in Italy and the same data in France, Germany and Spain, respectively

Figure 6 shows the wavelet coherence values between the total number of COVID-19 cases in Spain and the same data in France (Figure 6a) and Germany (Figure 6b).

69

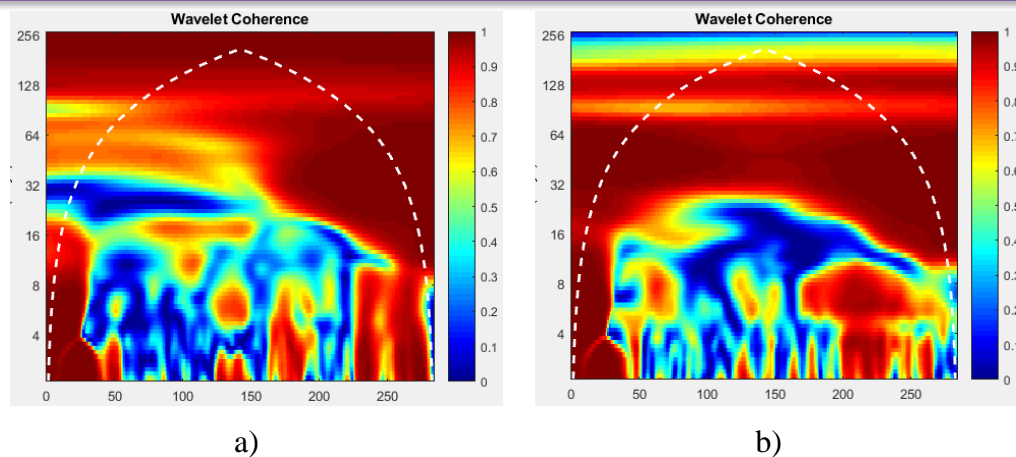a)   b)

Figure 6. Values of wavelet coherence between the total number of COVID-19 cases in Spain and the same data in France and Germany

We see that some of the figures have identical data. Especially these data are identical for the last days from the studied period of time. This suggests that the dynamics of the spread of the COVID-19 pandemic in individual European countries (which we are researching) is identical; over time, the identity of the dynamics for individual European countries (which we are studying) increases. At the same time, there is an increase in the incidence of COVID-19.

Thus, wavelet coherence allows you to determine the periods of identity in the spread of COVID-19 between countries. This allows us to build forecast models for the spread of the COVID-19 pandemic.

Conclusion

We conducted a comparative analysis of the dynamics of the development of the COVID-19 pandemic. To do this, we analyzed the development of the pandemic in the initial period and for the entire time of the pandemic. We used the coherence wavelet to investigate the dynamics of the COVID-19 pandemic. This analysis was carried out on the basis of general dynamics of diseased people, the dynamics of deaths and recovered from COVID-19 infection. We also made a comparative analysis of the dynamics of the pandemic development for individual European countries. Such an analysis allows us to speak about the identity in the development of the pandemic for individual countries and separate time periods. This can be the basis for building forecast models for the development of the COVID-19 pandemic.

References

Babker, A. M. A., Mohammed, O. I. A., & Lyashenko, V. (2020). COVID-19 data based on wavelet coherence estimates for selected countries in the Eastern Mediterranean. *World Journal of Advanced Research and Reviews*, 6(3), 110-120.

Ciufolini, I., & Paolozzi, A. (2020). Mathematical prediction of the time evolution of the COVID-19 pandemic in Italy by a Gauss error function and Monte Carlo simulations. *The European Physical Journal Plus*, *135*(4), 355.

Deb, S., & Majumdar, M. (2020). A time series method to analyze incidence pattern and estimate reproduction number of COVID-19. *arXiv preprint arXiv:2003.10655*.

Dolgin, E. (2019). The secret social lives of viruses. *Nature*, 570(7761), 290-292.

70

Krupovic, M., Dolja, V. V., & Koonin, E. V. (2019). Origin of viruses: primordial replicators recruiting capsids from hosts. *Nature Reviews Microbiology*, 17(7), 449-458.

Lyashenko, V., Saad, E. S. M., & Babker, A. M. A. (2020). Some data of wavelet analysis on the overall dynamics of the dead and recovered during infection COVID-19. *World Journal of Advanced Research and Reviews*, 6(2), 017-023.

Mackey, T., Purushothaman, V., Li, J., Shah, N., Nali, M., Bardier, C., ... & Cuomo, R. (2020). Machine Learning to Detect Self-Reporting of Symptoms, Testing Access, and Recovery Associated With COVID-19 on Twitter: Retrospective Big Data Infoveillance Study. *JMIR Public Health and Surveillance*, *6*(2), e19509.

Mustafa, S. K., Ahmad, M. A., Sotnik, S., Zeleniy, O., Lyashenko, V., & Alzahrani, O. (2020). Brief review of the mathematical models for analyzing and forecasting transmission of COVID-19. *Journal of critical reviews,* 7(19), 4206-4210.

Nickbakhsh, S., Ho, A., Marques, D. F., McMenamin, J., Gunson, R. N., & Murcia, P. R. (2020). Epidemiology of seasonal coronaviruses: Establishing the context for COVID-19 emergence. *The Journal of Infectious Diseases*, 222(1), 17–25.

Petersen, E., Koopmans, M., Go, U., Hamer, D. H., Petrosillo, N., Castelli, F., ... & Simonsen, L. (2020). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet infectious diseases*, 20 (9), e238-e244.

Spinelli, A., & Pellino, G. (2020). COVID-19 pandemic: perspectives on an unfolding crisis. *The British journal of surgery*, 107(7), 785–787.

Srivastava, A., & Chowell, G. (2020). Understanding Spatial Heterogeneity of COVID-19 Pandemic Using Shape Analysis of Growth Rate Curves. *medRxiv.*

Torrence, C., & Webster, P. J. (1999). Interdecadal changes in the ENSO–monsoon system. *Journal of climate*, *12*(8), 2679-2690.

Vasiurenko, O., Lyashenko, V., Baranova, V., & Deineko, Z. (2020). Spatial-Temporal Analysis the Dynamics of Changes on the Foreign Exchange Market: an Empirical Estimates from Ukraine. *Journal of Asian Multicultural Research for Economy and Management Study*, 1(2), 1-6.

Weinberger, D. M., Chen, J., Cohen, T., Crawford, F. W., Mostashari, F., Olson, D., ... & Watkins, A. (2020). Estimation of excess deaths associated with the COVID-19 pandemic in the United States, March to May 2020. *JAMA Internal Medicine*, *180*(10), 1336-1344.

Wolkewitz, M., & Puljak, L. (2020). Methodological challenges of analysing COVID-19 data during the pandemic. *BMC Medical Research Methodology*, 20, 81-84.

Yousaf, M., Zahir, S., Riaz, M., Hussain, S. M., & Shah, K. (2020). Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan. *Chaos, Solitons & Fractals*, 109926.