The effect of weather data on the spread of COVID-19 in Jordan

Eman Abdelhafez1 · Loai Dabbour2 · Mohammad Hamdan3

Received: 16 September 2020 / Accepted: 30 December 2020 / Published online: 9 January 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract
This study aims to analyze the correlation between the daily confirmed COVID-19 cases in Jordan and metrological parameters including the average daily temperature (°C), maximum ambient temperature (°C), relative humidity (%), wind speed (m/s), pressure (kPa), and average daily solar radiation (W/m²). This covers the first and the second waves in Jordan. The data were obtained from both the Jordanian Ministry of health and the Jordan Metrological Department. In this work, the Spearman correlation test was used for data analysis, since the normality assumption was not fulfilled. It was found that the most effective weather parameters on the active cases of COVID-19 in the initial wave transmission was the average daily solar radiation \( (r = -0.503; p = 0.000) \), while all other tests for other parameters failed. In the second wave of COVID-19 transmission, it was found that the most effective weather parameter on the active cases of COVID-19 was the maximum temperature \( (r = 0.394; p = 0.028) \). This was followed by wind speed \( (r = 0.477; p = 0.007) \), pressure \( (r = -0.429; p = 0.016) \), and average daily solar radiation \( (r = -0.757; p = 0.000) \). Furthermore, the independent variable importance of multilayer perceptron showed that wind speed has a direct relationship with active cases. Conversely, areas characterized by low values of pressure and daily solar radiation exposure have a high rate of infection. Finally, a global sensitivity analysis using Sobol analysis showed that daily solar radiation has a high rate of active cases that support the virus’ survival in both wave transmissions.

Keywords COVID-19 · Coronavirus · Weather data · Correlation test · Sensitivity analysis multilayer perceptron

Introduction
COVID-19, which was originated in the area of Wuhan, China, late December of 2019, is believed to be associated with the coronavirus 2 (SARS-CoV-2) (Zhou et al. 2020; Lam et al. 2020). SARS-CoV-2, which is believed to be from a bat origin (Huang et al. 2020; Wue et al. 2020), is a single-stranded positive-sense RNA virus (Satija and Lal 2007) with a nanosize ranging from 60 to 140 nm in diameter (Satija and Lal 2007; Hsiao et al. 2020). Currently, there are over sixty-seven million confirmed COVID-19 cases throughout the globe, and this number is expected to escalate due to the highly contagious nature of the virus.

As a result, early and accurate diagnosis of COVID-19 is of great importance for controlling the spread of the disease and to reduce its mortality. Emerging laboratory and epidemiological data suggest that environmental conditions may affect the current COVID-19 pandemic (Brassey et al. 2020). Consequently, the effect of meteorological parameters on the spread of COVID-19 should be explored to help in the prediction of its development over the coming months, although numerous factors may affect the progression of the COVID-19 pandemic.

Extensive research evolved to investigate the relationship between different metrological parameters and the COVID-19 daily cases (Cucinotta and Vanelli 2020). Pani et al. (2020) inspected the role of hot tropical weather in the transmission of COVID-19 by investigating the correlation between meteorological parameters and the COVID-19 pandemic cases in Singapore using Spearman and Kendall rank correlation. A positive significant correlation with the COVID-19 pandemic was found with temperature, dew point, relative humidity, absolute humidity, and water vapor.
Gupta et al. (2020) collected daily data of new cases of COVID-19 in 50 US states from January 1 to April 9, 2020, and two weather data (temperature and absolute humidity). They validated the results for 10-day intervals against monthly observations and also with worldwide trends. The obtained results were used to predict Indian regions which would be vulnerable to weather-based spread in the upcoming months of 2020. It was noticed that COVID-19 spread in the USA is significant for states characterized by a humidity ratio values in the range from 4 to 6 g/m$^3$. The number of new cases exceeds 10,000, irrespective of the chosen time intervals for the study parameters. These developments are consistent with worldwide observations but do not correlate well with Indian regions so far possibly due to the total cases reported per interval < 10,000. The results explained the relationship between weather parameters and COVID-19 spread.

Iqbal et al. (2020) researched the connection between the fast spread of COVID-19 and regional climate parameters over a global scale. They collected data of COVID-19 cases from January 2020 to 5th of June 2020 in different affected countries and regional climatic parameters data. They found that the countries located in the lower temperature region had a rapid increase in the COVID-19 cases than the countries located in the warmer climatic regions. Also, they concluded that the average daylight hours and average high temperature were correlated to total COVID-19 cases, while average high temperature showed a correlation with total COVID-19 death cases.

Menebo (2020) obtained data of COVID-19 from the Norwegian public health institute and weather data from the Norwegian Meteorological Institute. He analyzed the correlation between weather and COVID-19 in the city of Oslo, Norway. He concluded that maximum temperature, normal temperature (°C), and precipitation level (mm) significantly affect the incidence rate of COVID-19.

Ahmadi et al. (2020) conducted a study to investigate the effect of the number of infected people with COVID-19 based on the population density, intra-provincial movement, and infection days to the end of the study period. The study investigated the average temperature, average precipitation, humidity, wind speed, and average solar radiation effect on COVID-19 spreading in Iran. They used the partial correlation coefficient (PCC) and Sobol-Jansen methods for analyzing the effect and correlation of variables with the COVID-19 spreading.
rate. Sensitivity analysis showed that the population density and intra-provincial movement have a direct relationship with the infection outbreak. However, areas with low values of wind speed, humidity, and solar radiation were found to be of high rate of infection that supports the virus’ survival.

Based on the above literature review, it is of importance to correlate weather variables (including average daily temperature, maximum temperature, relative humidity, wind speed, pressure, and average daily solar radiation) to COVID-19 active cases. Having achieved this, sensitivity analysis between variables was conducted using Sobol method, and independent variable importance of multilayer perceptron was found.

**Methods**

**Study area**

Jordan is located between longitudes 35° 40’ and 39° E and between latitudes 29° 30’ and 34° N in the northern hemisphere between the zone biome (moist cold winters and dry hot summers) and the arid one (Al-Quraín 2008). The total area of Jordan is 89,342 km² and according to the General Census of Population and Housing Institute the total population of Jordan exceeds ten million.

**Data collection**

In this work, the daily data on confirmed COVID-19 cases in Jordan over the period from March 15 to August 31, 2020, was obtained from the Jordanian Ministry of Health, while the weather data including the average daily temperature (°C), maximum ambient temperature (°C), relative humidity (%), wind speed (m/s), atmospheric pressure (kPa), and average daily solar radiation (W/m²) were obtained from Jordan Metrological Department.

Since the announcement made by WHO that COVID-19 is a pandemic, the number of confirmed cases is continuously escalating worldwide (Cucinotta and Vanelli 2020). Officially, the first confirmed COVID-19 case was declared by the Jordanian Health Ministry on March 2. The number of confirmed cases increased to 2739 on September 10, 2020, while the number of recovered cases was 1981, and 688 cases were still under treatment, and the total death numbers was 20 cases. Figure 1 shows the number of confirmed COVID-19 cases over the period from March 15 to August 31, 2020.

**Results**

**The initial wave of COVID-19 transmission**

Figure 2 shows the number of confirmed cases during the initial wave of COVID-19 transmission, which was over the period from March 15 to April 30, 2020. As shown in the figure, there exists a general trend that indicates a decrease in the positive cases over the test period.

**Normality test**

The normality test results in the initial wave COVID-19 transmission are shown in Table 1. The basic assumption of Pearson’s correlation for the variable (number of daily new cases of COVID-19 in Jordan) was not achieved. The $p$ value from the Kolmogorov-Smirnov test of normality was found to be 0.022, which is less than 0.05. The $p$ value from the Shapiro-Wilk test of normality was found to be 0.00, which is less than 0.01. Based on these results, there was an assumption that the normal distribution of daily new cases of COVID-19 in the initial wave transmission in Jordan was found to be not acceptable; as a result, the Spearman correlation test (non-

| Table 1 | Normality test results |
|---------|------------------------|
| Kolmogorov-Smirnov* | Shapiro-Wilk |
| Statistic | df | Sig. | Statistic | df | Sig. |
| Cases | 0.140 | 47 | 0.022 | 0.878 | 47 | 0.000 |

* Lilliefors significance correction

| Table 2 | Spearman correlation test results |
|---------|----------------------------------|
| Correlations | Max. temp | Temp | Relative humidity | Wind speed | Pressure | Radiation |
| Spearman’s rho Cases | Correlation coefficient | −0.245 | −0.283 | 0.092 | 0.041 | −0.015 | −0.503** |
| Sig. (2-tailed) | 0.098 | 0.054 | 0.540 | 0.783 | 0.920 | 0.000 |
| $N$ | 47 | 47 | 47 | 47 | 47 | 47 |

**Correlation is significant at the 0.01 level (2-tailed)**

**Correlation is significant at the 0.05 level (2-tailed)**
parametric correlation estimation) was used to examine and determine the relationships between the COVID-19 pandemic in the initial wave transmission and weather variables. A bi-variate, two-tailed analysis at $\alpha = 95\%$ as a confidence interval was applied.

**Correlation test**

Since normality assumption was not fulfilled, Spearman correlation test was used for data analysis. Table 2 shows Spearman correlation test results. Daily temperature ($r = -0.283; p = 0.054$), maximum temperature ($r = -0.245; p = 0.098$), relative humidity ($r = 0.092; p = 0.540$), wind speed ($r = 0.041; p = 0.783$), and pressure ($r = -0.015; p = 0.920$) were not significantly correlated with COVID-19 pandemic positive cases because the $p$ value is more than 0.05. Average daily solar radiation ($r = -0.503; p = 0.000$) was negatively strong and significantly correlated with COVID-19 pandemic because the $p$ value is less than 0.05.

**Multilayer perceptron (MLP)**

Multilayer perceptron (MLP) is a popular feed-forward neural network architecture. MLP consists of three layers, namely, input, output, and hidden layers (Ahmadi et al. 2019). More details about MLP is found in Ahmadi et al. (2019).

Based on trial and error method, MLP technique, with 10 hidden neurons and using 70% of the data for training, 15% testing and 15% for validation, was conducted to obtain the best model with $R = 70.5\%$ for the initial wave of COVID-19 transmission as shown in Fig. 3.

Table 3 shows independent variable importance, which represents the measuring of how much the network’s model-predicted value changes for different values of the independent variable.

The results show that daily solar radiation and daily temperature are the most effective variables on the spread rate of COVID-19 virus with wind speed and pressure being the least effect.

**Sobol sensitivity analysis**

Sensitivity analysis between variables is conducted using Sobol sensitivity method to examine the importance of variables. More details about sensitivity analysis are outlined in Dorosti et al. (2020). It is based on the decomposition of model output variance into summands of variances of the input parameters in increasing dimensionality (Saltelli et al. 1999; Sobol 2001).

Sobol sensitivity analysis is applied in quantifying how much variability in the model output is depending on each of the input parameter or the interactions between parameters. It can be achieved by computing the first-order, second-order, higher-order, and the overall sensitivity indices.

In this research, the Sobol sensitivity analysis is conducted using a global sensitivity analysis toolbox in MATLAB. Figure 4 shows the significance and sensitivity analysis of the Sobol method. The vertical axis indicates the first-order effect of sensitivity. The results show that the most important

![Fig. 3 Regression plot of predicted values against actual ones for the MLP model in the initial wave of COVID-19 transmission](image)

![Table 3 Independent variable importance](image)
variable in COVID-19 virus infection rate is daily solar radiation at the initial wave transmission.

The second wave of COVID-19 transmission

Figure 5 shows the number of confirmed cases during the second wave of COVID-19 transmission, which was over the period from August 1 to August 31, 2020. As shown in the figure, there exists a general trend that indicates an increase in the positive cases over the test period.

Normality test

The normality test results in the second wave COVID-19 transmission are shown in Table 4. The basic assumption of Pearson’s correlation for the variable (number of daily new cases of COVID-19 in Jordan) was not achieved. The p value from the Kolmogorov-Smirnov test of normality was 0.007, which is less than 0.05. The p value from the Shapiro-Wilk test of normality was found to be 0.001, which is less than 0.01. Based on these results, there was an assumption that the normal distribution of daily new cases of COVID-19 in the second wave COVID-19 transmission in Jordan was found to be not acceptable. As a result, the Spearman correlation test (non-parametric correlation estimation) was used to examine and determine the relationships between the COVID-19 pandemic spread and weather variables during the second wave transmission. A bivariate, two-tailed analysis at $\alpha = 95\%$ as a confidence interval was applied.

Correlation test

Since normality assumption was not fulfilled, Spearman correlation test was used for data analysis. Table 5 shows Spearman correlation test results. Daily temperature ($r = 0.394; p = 0.028$), wind speed ($r = 0.477; p = 0.007$), pressure ($r = -0.429; p = 0.016$), and average daily solar radiation ($r = -0.757; p = 0.000$) were significantly correlated with COVID-19 pandemic spread because the p value is less than 0.05. Maximum temperature ($r = 0.339; p = 0.062$) and relative humidity ($r = -0.211; p = 0.255$) were not significantly correlated with COVID-19 pandemic spread because the p value is more than 0.05.

It may be noticed that maximum temperature and wind speed were positively moderate correlated. The pressure was negatively moderate correlated, and the average daily temperature was negatively strong correlated.

Multilayer perceptron (MLP)

Based on trial and error method, MLP technique with 10 hidden neurons and using 60% of the data for training, 20%
testing, and 20% for validation was used to obtain the best model with $R = 83.8\%$ for the second wave of COVID-19 transmission as shown in Fig. 6.

Table 6 shows independent variable importance. The results show that daily solar radiation and pressure are the most important variables on the COVID-19 virus infection rate, with the least important variables being wind speed and maximum temperature.

**Sobol sensitivity analysis**

The Sobol sensitivity analysis in this research was conducted using a global sensitivity analysis toolbox in MATLAB. Figure 7 shows the significance and sensitivity analysis of the Sobol method. The vertical axis indicates the first-order effect of sensitivity. The results show that the most important variable in COVID-19 virus infection rate is daily solar radiation at the second wave transmission.

**Discussion**

In this study, weather variables including average daily temperature, maximum temperature, relative humidity, wind speed, pressure, and average daily solar radiation were correlated with COVID-19 active cases in Jordan during the first and second waves. This was achieved by using non-parametric correlation test.

It was found that the most effective weather parameters on the active cases of COVID-19 in the initial wave transmission was the average daily solar radiation ($r = -0.503; p = 0.000$), while all other tests for other parameters failed as shown in Table 2.

In the second wave COVID-19 transmission, it was found that the most effective weather parameter on the active cases of COVID-19 was the maximum temperature ($r = 0.394; p = 0.028$). This was followed by wind speed ($r = 0.477; p = 0.007$), pressure ($r = -0.429; p = 0.016$), and average daily solar radiation ($r = -0.757; p = 0.000$) as shown in Table 5.

Independent variable importance of multilayer perceptron to examine the importance of variables was conducted. The best coefficient of determination ($R$) was 70.5% and 83.8% in the initial and second wave COVID-19 transmission, respectively. During the initial wave of COVID-19 transmission, the results showed that daily solar radiation and daily temperature were the most important variables on the COVID-19 virus infection rate, while the least important variables were wind speed and pressure.

In the second wave of COVID-19 transmission, the results showed that daily solar radiation and pressure were the most important variables in the COVID-19 virus infection rate, and the least important variables were wind speed and maximum temperature.

The Sobol sensitivity analysis was also conducted using a global sensitivity analysis toolbox in MATLAB. The results showed that the most important variable in COVID-19 virus infection rate was daily solar radiation with a value equal to 0.3491 as shown in Fig. 4 at the initial wave transmission. Also, the results showed that the most important variable in COVID-19 virus infection rate was daily solar radiation with a value equal to 0.7452 as shown in Fig. 7 at the initial wave transmission.

Based on the above statistical analysis, correlation test, independent variable importance of Multilayer Perceptron, and Sobol sensitivity analysis, it may be noticed that the most effective weather parameter on the active cases of COVID-19 in the initial and the second waves transmission was the average daily solar radiation. It was found that the average daily solar radiation was negatively strong and significantly correlated, which means that the number of daily confirmed cases in Jordan decreases with the daily solar radiation.

**Table 4 Normality test results**

| Kolmogorov-Smirnov | Shapiro-Wilk |
|--------------------|-------------|
| Statistic | df | Sig. | Statistic | df | Sig. |
| Cases | 0.187 | 31 | 0.007 | 0.862 | 31 | 0.001 |

* Lilliefors significance correction

**Table 5 Spearman correlation test results**

| Correlations | Max. Temp | Temp Relative Humidity | Wind speed | Pressure | Radiation |
|--------------|-----------|------------------------|------------|----------|-----------|
| Spearman’s rho | 0.394* | 0.339 | -0.211 | 0.477** | -0.429* | -0.757** |
| Correlation coefficient | 0.028 | 0.062 | 0.255 | 0.007 | 0.016 | 0.000 |
| Sig. (2-tailed) | 31 | 31 | 31 | 31 | 31 | 31 |

* Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)
Time complexity of algorithm

In the presented algorithm used, three main approaches related to best model and time complexity were calculated. Xiao (2017) proved that the time complexity of the algorithm is $O(m_1, m_2)$ for Spearman’s rank correlation, where $m_1$ and $m_2$ are fixed integers. The running time of MLP approach with maximum number of indicators is $O(p.n.h.o.i)$, where $h$ and $i$ are the numbers of hidden layer and iteration, respectively (Serpen and Gao 2014). Using Monte-Carlo sampling technique, time complexity of computing an estimation of $E[\hat{E}_{\xi_{\alpha}^{u}}]^{2}$ expressions for one Sobol’ index term becomes $O(2n)$ (Kourany et al. 2017).

Conclusion

In this study, weather variables including average daily temperature, maximum temperature, relative humidity, wind speed, pressure, and average daily solar radiation were correlated with COVID-19 active cases in Jordan during the first and second waves. This was achieved by using non-parametric correlation test.

It was found that the most effective weather parameters on the active cases of COVID-19 in the initial wave transmission is the average daily solar radiation ($r = -0.503; p = 0.000$), while all other tests for other parameters failed.

In the second wave COVID-19 transmission, it was found that the most effective weather parameter on the active cases of COVID-19 is the maximum temperature ($r = 0.394; p = 0.028$). This was followed by wind speed ($r = 0.477; p = 0.007$), pressure ($r = -0.429; p = 0.016$), and average daily solar radiation ($r = -0.757; p = 0.000$).
In order to examine the importance of variables, independent variable importance of multilayer perceptron was conducted, and global sensitivity analysis by using Sobol analysis approved that the average daily solar radiation plays an important role in the COVID-19 outbreak.

Author contributions Conceptualization: Eman AbdelhafezMethodology: Eman AbdelhafezFormal analysis and investigation: Loai DabbourWriting - original draft preparation: Eman Abdelhafez, Loai Dabbour, and Mohammad HamdanWriting - review and editing: Mohammad Hamdan and Loai Dabbour

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Competing interests The authors declare that they have no competing interests.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

References

Ahmadi M, Jafzarzadeh-Ghoushchi S, Taghizadeh R, Sharifi A (2019) Presentation of a new hybrid approach for forecasting economic growth using artificial intelligence approaches. Neural Comput & Applic 31:8661–8680
Ahmadi M, Sharifi A, Dorosti S, Jafzarzadeh-Ghoushchi S, Ghanbari N (2020) Investigation of effective climatology parameters on COVID-19 outbreak in Iran. Sci Total Environ 729:138705
Al-Quraín S (2008) Taxonomical and pharmacological survey of therapeutic plants in Jordan. J Nat Prod 1:10–26
Brassej J, Heneghan C, Mahtani KR, Aronson JK (2020) Do weather conditions influence the transmission of the coronavirus (SARS-CoV-2)?. Retrieved from https://www.cebm.net/COVID-19/dow-weather-conditions-influence-the-transmission-of-the-coronavirus-SARS-CoV-2/
Cucinotta D, Vanelli M (2020) WHO declares COVID-19 a pandemic. Acta Bio-med 91:157
Dorosti S, Jafzarzadeh Ghoushchi S, Sobhrakshankhah E, Ahmadi M, Sharifi A (2020) Application of gene expression programming and sensitivity analyses in analyzing effective parameters in gastric cancer tumor size and location. Soft Comput 24:9943–9964
Gupta S, Raghushanshi GS, Chanda A (2020) Effect of weather on COVID-19 spread in the US: a prediction model for India in 2020. Sci Total Environ 728:138860. https://doi.org/10.1016/j.scitotenv.2020.138860
Hsiao TC, Chuang HC, Griffith SM, Chen SJ, Young LH (2020) COVID-19: an aerosol’s point of view from expiration to transmission to viral-mechanism. Aerosol Air Qual Res 20:905–910. https://doi.org/10.4209/aaqr.2020.04.0158
Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395:497–506. https://doi.org/10.1016/S0140-6736(20)30183-5
Iqbal MM, Abid I, Hussain S, Shahzad N, Waqas MS, Iqbal MJ (2020) The effects of regional climatic condition on the spread of COVID-19 at global scale. Sci Total Environ 739:140101
Kourany T, Ghoneima M, Hegazi E, Ismail Y (2017) Passiot: a pareto-optimal multi-objective optimization approach for synthesis of analog circuits using sobol’ indices-based engine. Integration VLSI J 58:9–21
Lam TTY, Shum MHH, Zhu HC, Tong YG, Ni XB, Liao YS, Wei W, Cheung WYM, Li WJ, Li LF, Leung GM, Holmes EC, Hu YL, Guan Y (2020) Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. Nature. 583:282–285. https://doi.org/10.1038/s41586-020-2169-0
Menebo MM (2020) Temperature and precipitation associate with COVID-19 new daily cases: a correlation study between weather and COVID-19 pandemic in Oslo, Norway. Sci Total Environ 737:139659
Pani SK, Lin N, RavindraBabu S (2020) Association of COVID-19 pandemic with meteorological parameters over Singapore. Sci Total Environ 740:140112
Saltelli A, Tarantola S, Chan K-S (1999) A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 41(1):39–56
Satija N, Lal SK (2007) The molecular biology of SARS coronavirus. Ann N Y Acad Sci 1120:26–38. https://doi.org/10.1196/annals.1408.002
Serpen G, Gao Z (2014) Complexity analysis of multilayer perceptron neural network embedded into a wireless sensor network. Procedia Comput Sci 36:192–197. https://doi.org/10.1016/j.procs.2014.09.078
Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simul 55(1–3):271–280
Wu Y, Jing W, Liu J, Ma Q, Yuan J, Wang Y, Du M, Liu M (2020) Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Sci Total Environ 729:139051. https://doi.org/10.1016/j.scitotenv.2020.139051
Xiao W (2017) An online algorithm for nonparametric correlations. arXiv preprint arXiv:1712.01521
Zhou T, Liu Q, Yang Z, Liao J, Yang K, Bai W, Lu X, Zhang W (2020) Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV. J Evid Based Med 13:3–7. https://doi.org/10.1111/jebm.12376

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.