Modeling the contribution of Nitrogen Dioxide, Vertical pressure velocity and PM2.5 to COVID-19 fatalities

Alaba Boluwade1 · Amna M.1 · Al Ruheili2

Abstract
The COVID-19 caused by the severe acute respiratory syndrome coronavirus was reported in China in December 2019. The severity and lethality of this disease have been linked to poor air quality indicators such as tropospheric nitrogen dioxide (NO$_2$) and dust surface mass concentration particulate matter (PM2.5) as possible contributors. The Arab League has 22 member countries and is home to almost 420 million people. The primary objective of this study is to assess the relationship between NO$_2$, PM2.5 and vertical pressure velocity (hereafter: OMEGA) (extracted from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) database), socio-economic factors (the population and geographic area of each member country) and COVID-19 deaths using Bayesian model averaging. The total plausible models ($2^5$) were estimated. The results show that the posterior inclusion probability (PIP), which indicates the probability that a particular indicator is included in the best model, was 0.69, 0.94, 0.68, 0.47, and 0.61 for OMEGA, PM2.5, NO$_2$, geographical area, and population, respectively, meaning that these variables are important contributors in predicting COVID-19 fatalities in the Arab League states. This study shows that atmospheric satellite measurements from MERRA-2 datasets are capable of being used to quantify trace gases in pandemic studies.

Keywords COVID-19 · PM2.5 · Geospatial mapping · Bayesian Model Averaging · Nitrogen Dioxide · Arab League States

1 Introduction
The Arab League of States (ALS) is a global sub-region with a population of more than 390 million people (World Bank, 2020). This region is comprised of 21 countries (Algeria, Bahrain, Djibouti, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Mauritania, Oman, Qatar, Kingdom of Saudi Arabia (KSA), Somalia, Sudan, Syria, United Arab Emirates (UAE), Tunisia, Palestine, and Yemen) that are contiguous (except Comoro Islands) and share similar cultural histories. Due to rising population growth and poor health facilities in some member states, the risk of the rapid spread of epidemic communicable diseases is high. The novel severe acute respiratory syndrome coronavirus 2, known as SARS-CoV-2 disease (COVID-19) has had an unprecedented impact on people’s lives through economic losses and fatalities. The first case of COVID-19 was reported in Egypt on February 14, 2020, followed by Oman which was reported on February 24, 2020. Since then, there have been more than 200,000 confirmed active cases with more than 16,645 fatalities (as of July 28th, 2020). COVID-19 is a respiratory and infectious disease whose symptoms include tiredness, shortness of breath, fever, etc. The focus on reducing the spread of the virus has been from human to human, but consideration needs to be given as to how environmental factors such as air, soil, and pre-existing health conditions such as mortality rate or access to health facilities play a role in transmission and deaths. A very conducive environment with a relatively medium air moisture level has also been reported to aid in severe acute respiratory syndrome (SARS) transmission (Chan et al. 2011). In some other studies, the impacts of lockdowns on the reduction of these
air contaminants have also been evaluated (Sekmoudi et al. 2020; Singh and Chauhan, 2020; Mishra et al., 2021; Chauhan and Singh, 2021). From the spatio-temporal changes standpoint, Liu et al. (2021) evaluated the impacts of lockdown policy on the NO$_2$ emissions in six countries (China, South Africa, Brazil, India, and the United Kingdom). The authors concluded that the lockdown policy reduced the NO$_2$ trend in all the countries using 10-year climatology as a reference. Air quality is not restricted to a local area. This is because air pollution can be transported from one country to another depending on wind distribution & patterns in the atmosphere (Copernicus, 2020). Poor air quality has a direct link to heart diseases (EPA, 2020). In the same EPA (2020) report, it was established that “healthy individuals exposed to air particle pollution over the long term had accelerated cases of atherosclerosis”. Pre-existing health conditions and factors such as coronary heart disease have been associated with COVID-19 death cases. Wichmann et al. (2020) conducted a study on autopsies of 10 in-patients and two outpatients. The authors reported that coronary heart disease was among the most common conditions. Globally, around 7 million people die from coronary heart disease annually (WHO, 2020). More than three-quarters of these cases are in developing countries (WHO, 2020). The COVID-19 virus thrives as tiny water droplets in the atmosphere and therefore natural/artificial surfaces (i.e. soil, concrete, glass) could be potential hosts. Remote sensing techniques, especially from the free and widely available Modern-Era Retrospective Analysis and Research and Application, version 2 (MERRA-2) sensor, cannot be used to detect COVID-19 occurrence or cases. They can, however, be used to study and understand subtle changes and interactions in environmental variables such as temperature, air quality, precipitation, relative humidity, and soil moisture content. Various studies have reported on the impacts of environmental factors such as air pollution (i.e. nitrogen dioxide (NO$_2$), particulate matter (PM$_{2.5}$ (< 2.5 μm) and PM$_{10}$ (< 10 μm)) and ozone (O$_3$) on the COVID-19 fatality rate (Ogen, 2020; Wu et al. 2020a; Wu et al. 2020b; Contini and Costabile, 2020; Boluwade, 2020). Conticini et al. (2020) provided evidence showing that those living in areas with high levels of pollutants are prone to chronic respiratory health conditions and eventual deaths. The authors concluded that a high level of air pollutants should be considered as an additional co-factor in the high level of COVID-19 lethality. Sarmadi et al. (2020) conducted a study in the United States that related COVID-19 distribution with factors such as total air quality index (AQI), carbon monoxide, ozone, NO$_2$, population density, and other geographic covariates. The authors concluded that these important variables were significantly related to COVID-19 deaths. Coccia (2020) also conducted a study in Italy that compares the COVID-19 distribution with air quality indicators such as PM$_{10}$. The author concluded that the number of infected people was higher in cities where the population is exposed to “more than 100 days per year exceeding the limit set for PM$_{10}$”. In another study, Pansini and Fornacca (2020) concluded that there were more viral infections in areas with high carbon monoxide, formaldehyde and NO$_2$, meaning that poor air quality seems to be a risk factor in the prevalence of COVID-19 cases and deaths. Liang et al. (2020) conducted a study using 3122 counties in the United States from January 22, 2020 – April 29, 2020. The authors also concluded that county level NO$_2$ concentrations were positively correlated with COVID-19 fatality and cases. Hendryx and Luo, (2020) also concluded that air pollution contributed to the COVID-19 prevalence and fatality rates using cumulative COVID-19 fatality data. Reviewing the impacts of air qualities on COVID-19 prevalence and fatality rates in various regions of the world is beyond the scope of the current study, however, comprehensive study on the impacts of air pollution on COVID-19 cases and fatalities can be found in a review paper such as Woody et al. (2020).

Several of the modeling or prediction attempts on the impacts of air quality or socio-economic variables on COVID-19 cases and fatalities have been based on simple linear regression employing the classical statistical techniques (Rodriguez-Villamizar et al. 2021; Kodera, et al. 2020; Liang et al., 2020), but this approach suffers from a failure to account for the uncertainty in the model selection (Hoeting et al., 1999). In other words, this classical approach leads to risky decisions and inferences (Hoeting et al., 1999). A good inference should indicate the ambiguity of the model and the criteria for its selection. In addition, a statistical tool that is useful for analyzing the contribution of these variables to COVID-19 cases is Bayesian model averaging (BMA) (Ley and Steel, 2011; Eicher, et al., 2009). BMA is a statistical method that can combine all possible models and account for model uncertainty. BMA method would reduce the overconfidence that arises when uncertainty in model is not considered. In other words, BMA is robust in any cases of model misspecification, meaning that BMA provides ranges of rival models instead of just selecting a single model (Hinne, et al., 2020). The posterior inclusion probabilities (PIPs), which indicate the probability that each explanatory variable belongs to the best model, are also estimated. The primary objective of this paper is to explore the spatiotemporal patterns of air indicators and confirmed fatalities of COVID-19 and quantify the contributions and how environmental conditions (i.e. NO$_2$, PM$_{2.5}$, and OMEGA) impacted the COVID-19 fatality rate within the 52 months (January 2016–April 2020) preceding the pandemic fatalities in the Arab League region. Therefore, the aim of this study is to characterize the impacts of long-term
mixtures are frequently present near oil and gas production sites”.

The fluctuations in oil prices have created an economic recession as a consequence of the oil trade (Arvai et al., 2014). In response, the GCC and AL are encouraging diversification in their economies to secure more job opportunities, invest in non-oil industries such as tourism, and develop infrastructure, trade, logistics, transportation, and agriculture (Fayad & Rasmussen, 2012). This is encouraging news as this diversification could help in reducing the poor air quality from the crude oil production.

The MENA region (Fig. 1) is characterized by hot-arid climate in a similar pattern with the GCC. For instance, these regions are characterized by two main seasons; winter with a minimum temperature of 15 °C at night and a maximum temperature reaching up to 50+ °C during summer days with low relative humidity (Al-Maamary, Kazem & Chaichan, 2017). Moreover, various studies suggested a reduction in precipitation by 10–30% in the next century due to climate change (Sowers et. al., 2011). Furthermore, an increase in extreme heat events is anticipated by the end of the century due to climate change (Lelieveld et al., 2016).

Within the GCC alone, the number of the urban populations will be in continuous growth as 83% of the urban area had increased since 2008 (Al-Maamary, Kazem & Chaichan, 2017). Pollution factors such as sand and desert dust have been linked with harmful PM2.5 particles in AL countries such as Egypt, Saudi Arabia, Iraq, Kuwait and Bahrain (Malek, 2019).

2 Materials and methods

2.1 Study Area

The Arab League (AL), a regional organization in the Middle East and North Africa (MENA) region, was established in 1945. The AL was originally formed by six members: Egypt, Iraq, Jordan, Lebanon, Saudi Arabia, and Syria. The aim of the AL is to foster cooperation between its members and to safeguard the members’ independence and sovereignty (Vlahovic et al., 2018). Moreover, the AL members share similar norms, values, identities, understandings, and institutions. The AL involves 22 members from Asia and Africa, and it has been recognized by the United Nations. The AL principles are formed on the basis of promoting good relations between Arab countries, to harmonize their interests, to solve problems within the scope of the AL members, to promote the interests of Arab states, and to support members to become independent states (Dakhllallah, 2012; Hedstrom, 2020).

The MENA region is located in the Middle East and North Africa (Jayakumar, 2019). MENA countries play a major role in the international oil industry. For example, in the Middle East, members of the AL are known for their oil and gas reserves, which comprise about 39.6% of the world’s crude oil (Durrat & Orthy, 2017). Saudi Arabia, Kuwait, Bahrain, Qatar, Oman, and the UAE have the highest income per capita in the Middle East. Moreover, these six countries established the Gulf Cooperation Council (GCC) in 1981 to maintain the stability of their members (Maulidya, 2020). This has allowed the Middle East to become a key area for the oil market and trade, as the area nourishes energy industries that rely on oil and natural gas extraction and refinement. In addition, the majority of Arab states’ GDP comes from hydrocarbon and various governmental activities except for Algeria, Yemen, UAE, and Bahrain. Thus, oil is the main revenue source for most Arab states. For instance, in half of the states of the GCC, more than 80% of exports are based on oil (Jayakumar, 2019).

The huge dependence on oil as a source of income in the Middle East has had a massive impact on their development plans, as the oil price is declining and industrial countries are shifting their focus toward green energy and sustainability (Callen et al., 2014). According to EPA (2003), environmental hazards from refineries have been a major source of pollutants, especially in areas where they are located. Furthermore, Macey et al. (2014) concluded that “Air concentrations of potentially dangerous compounds and chemical exposures to poor air quality and the risk of COVID-19 death from a Bayesian statistical technique standpoint.

2.2 Datasets Description

The data used for this study are divided into two categories:

i. Response Variable.

Fig. 1 Geographical locations of Arab League Member States
COVID-19 Deaths: The number of cases and deaths was downloaded from the WHO dashboard (https://covid19.who.int/) as of July 31, 2020 [downloaded on January 9, 2022]. This period was selected to capture COVID-19 deaths at the onset of the pandemic. Long-term exposure to air pollutants can lead to health complications such as chronic respiratory and inflammatory stimulus conditions (Conticini et al., 2020). Therefore, at the onset of the pandemic, this affected population segment will be most heavily impacted due to their pre-existing health conditions. For BMA analysis, a log transformation was performed for all the variables. Although the COVID-19 death values may seem low in some AL countries when compared with other regions in the world, there is a high likelihood that these figures may be an underestimated because of a lack of proper recording and testing, and there may be deaths at home or remote villages that occur without any government certification.

ii. Explanatory Input Datasets: These are input variables (explanatory) for the BMA modeling.

a) Dust surface column mass density (hereafter: PM2.5): This is the monthly average dust surface column mass density obtained from MERRA-2 global data at a spatial resolution of 0.5°x0.625° (GMAO, 2015) for January–April, 2020. This is considered as dust particles in the form of solid particles suspended in the atmosphere measured in kg/m² of each pixel of the image.

b) Tropospheric NO₂ column (hereafter: NO₂): This is the NO₂ tropospheric column (30% cloud screened) (Table 1) obtained from NASA. This is NO₂ measured by the Ozone Monitoring Instrument (OMI) on NASA’s Aura satellite. According to NASA, the NO₂ tropospheric column is the total number of NO₂ molecules in an atmospheric column. This space measurement is an indication of the fossil fuels being burned for transportation purposes and the generation of electricity (NASA, 2020). Although this is not the same as the NO₂ abundance at the surface, the satellite-based NO₂ measurement can serve as a proxy for the values measured at the earth’s surface. The version used in the study is the “OMNO2d: OMI/Aura NO₂ Cloud-Screened Total and Tropospheric Column L3 Global Gridded 0.25°x 0.25°/3” available at the NASA portal (Nickolay et al., 2019) for January–April, 2020. It should be noted that NO₂ observed by the satellite sensor is not the same as that on the earth’s surface due to cloud effects. However, it could act as a proxy for the NO₂ estimate at the earth’s surface. Compared with ground-based NO₂ measurements, which are based on specific locations and limited to that same location, the satellite-based NO₂ measurement is able to provide a seamless continuous estimate, especially in areas where no in-situ measurement is located.

c) The vertical pressure velocity (hereafter known as OMEGA [MERRA-2 Model M2IMNPASM v5.12.4]) also extracted from the NASA portal is an indication of vertical airflow during the same period of the event. In other words, OMEGA is an estimate of an indication of the atmospheric capability to disperse pollutants. The positive value is an indication of downward airflow, while the negative values show the upward movement of airflow. Therefore, in areas where the OMEGA is positive, the atmosphere will force the pollutant to be in close proximity to human exposure (Ogen, 2020).

2.3 Model Development

BMA is a statistical technique that can be used to account for model uncertainty by using plausible, possible, or ensemble numbers of models instead of just one model, as is the case in many conventional statistical inferences. For any linear

| Dataset used          | Period               | Source                                           |
|-----------------------|----------------------|--------------------------------------------------|
| COVID-19 Deaths       | Jan, 1-July, 2020    | WHO dashboard (https://covid19.who.int/)         |
|                        | 1-July, 2020         | Downloaded on 9th Jan, 2022                      |
| Dust surface column   | Jan, 2016, -April, 2020 | United States’ NASA’s Earth Observing System Data and Information System (EOSDIS) Channel (https://giovanni.gsfc.nasa.gov/giovanni/) |
| mass density          |                      |                                                  |
| spatial resolution     |                      |                                                  |
| 0.5°x0.625°           |                      |                                                  |
| [MERRA-2 Model         |                      |                                                  |
| M2TMNX-AER v5.12.4]    |                      |                                                  |
| Tropospheric NO₂ column at a 0.25° x 0.25° Spatial resolution [OMI OMNO2d v003] | United States’ NASA’s Earth Observing System Data and Information System (EOSDIS) Channel (https://giovanni.gsfc.nasa.gov/giovanni/) |
| Vertical pressure velocity monthly 0.5° x 0.625° @750 hPa [MERRA-2 Model M2IMNPASM v5.12.4] Pa s-1 | United States’ NASA’s Earth Observing System Data and Information System (EOSDIS) Channel (https://giovanni.gsfc.nasa.gov/giovanni/) |
structure model, having a response variable $Z$ (i.e. the percentage of COVID-19 deaths in AL), $\alpha$ as constant (intercept), $\beta$ as the coefficients of the explanatory variables $X$, and $\epsilon$ as the normal independent and identically distributed with variance represented as $\sigma^2$ can be written as (Zeugner, 2020):

$$Z = \alpha + X \cdot \beta + \epsilon \sim \mathcal{N}(0, \sigma^2 I)$$  \hspace{1cm} (1)

From Eq. (1), the classical conventional approach is to select a model based on certain criteria with the assumption that the model generated the response variable, $Z$. It is always a challenge when there are several variables in $X$, that have the possibility of being included in the final selected model according to the defined criteria. Inference through this approach will make it difficult to account for the uncertainty in the selected model. In other words, the parameters of the selected model are conditioned on such a model, and therefore the uncertainty involved in the selection is ignored. For instance, if there is $K$ number of variables in $X$, then there is potentially $2^K$ number of models and explanatory combinations.

The strength and advantage of BMA involve estimating all plausible combinations of the explanatory variables $X$ and also developing a weighted average for all these possible models. In other words, the model weights come from the posterior model probability, which can be expressed as a Bayes theoretical form:

$$p(M_i | Z, X) = \frac{p(M_i | X)p(M_i)}{p(Z|X)} = \frac{p(M_i | X)p(M_i)}{\sum_{i=1}^{2^K} p(Z|M_i, X)p(M_i)}$$  \hspace{1cm} (2)

Where $p(Z|X)$ represents the unified likelihood, $p(M_i | Z, X)$ is the posterior model probability (PMP), which is equivalent to the marginal likelihood of the model $p(Z|M_i, X)$ times the prior model probability $p(M_i)$ (Ley and Steel, 2009; Zeugner, 2020). Furthermore, the BMA approach is able to estimate all these combinations by estimating the posterior densities of the model parameters, $\theta$ through (Zeugner, 2020):

$$p(\theta | Z, X) = \sum_{\gamma=1}^{2^K} p(\theta | M_{\gamma}, Z, X)p(M_{\gamma} | X, Z)$$  \hspace{1cm} (3)

Where $p(\theta | M_{\gamma}, Z, X)$ is the posterior distribution for all possible sets of models, $M_\gamma$. $p(\theta | Z, X)$ is the weighted posterior distribution. A crucial parameter from the results that shows the importance of each explanatory variable is the PIP, which is the sum of the PMP for all the models where the explanatory models are included (Zeugner, 2020). More information about the BMA procedure can be found in Hoeting et al. (1999), Clyde, (1999), Amini and Parmeter (2011), Fernández et al., (2001), etc. BMA has been implemented in several packages in R statistical software. In this study, the Bayesian adaptive sampling (BAS) developed by Clyde (2020; 2021) is used.

## 3 Results and Discussions

### 3.1 Descriptive Statistics

The descriptive statistics of COVID-19 and other explanatory variables are shown in Table 2. Table 2 shows the coefficients of variation (CV) of 176%, 123%, 115%, 67%, 32% and 32% for COVID-19, area (geographical), population, OMEGA, PM2.5, and NO2 respectively. This shows the spatial heterogeneity and variability of COVID-19 and explanatory patterns as well as distributions in the AL areas. The minimum and maximum COVID-19 deaths were 11 (Jordan) and 4774 (Egypt) people, respectively. The table also shows that as of July 31st, 2020, there were a total of 17,207 fatalities due to COVID-19.

Figure 2 shows the correlation matrix of COVID-19 with the covariates. COVID-19 is strongly positively correlated with all the covariates except area, which is weakly positively correlated. This is an important finding, as it indicates that an increase in any of these four covariates could amplify

| Table 2: Descriptive statistics of COVID-19 and explanatory variables |
|---------------------------------------------------------------|
| **COVID-19 Deaths** | **Area (sq.km)** | **Population** | **OMEGA (Pa/s)** | **PM2.5 (10^3 kg/m^3)** | **NO2 (10^5 mol/cm^3)** |
| Mean | 819 | 626,164 | 19,967,311 | 0.472 | 0.118 | 0.119 |
| Standard Deviation | 1441 | 772,166 | 23,014,113 | 0.317 | 0.037 | 0.038 |
| Coefficient of Variation (%) | 176 | 123 | 115 | 67 | 32 | 32 |
| Minimum | 11 | 758 | 958,923 | 0.068 | 0.069 | 0.042 |
| Maximum | 4774 | 2,381,741 | 98,423,598 | 0.966 | 0.179 | 0.174 |
| Sum | 17,207 | 13,149,440 | 419,313,527 | 9.915 | 2.483 | 2.496 |
COVID-19 cases and fatalities. From a health management standpoint, this shows that worsening air quality (increased NO$_2$, PM2.5 & OMEGA) could exacerbate the pandemic.

Figure 3 shows the spatial distributions of the input variables (regressors). The spatial distribution shown in Fig. 3(a) has high NO$_2$ levels ($>4.5 \times 10^{15}$ mol/cm$^2$) in pockets of Egypt, the KSA, the UAE, and Iraq. It is clear that the hotspot of PM2.5 (Fig. 3(b): $>0.10 \times 10^3$ kg/m$^2$) is spread through the KSA, Algeria, Iraq, Oman, Libya, the UAE, Sudan, Mauritania, and part of Morocco. The vertical pressure velocity represented as OMEGA is the ability of the atmospheric air to disperse the pollutant and is relatively high in Egypt, Iraq, Mauritania, Sudan, parts of Morocco, and Libya. When comparing these findings with the COVID-19 fatalities in Fig. 4, it seems that several of the above indicators are spatially correlated with COVID-19. There are high values of fatalities in the KSA ($>725$), Iraq ($>725$), Algeria ($>725$), Egypt ($>725$), and Sudan ($>345$). Although Libya and Somalia seem to have low numbers of COVID-19 fatalities, the actual values may not be reported because of the civil unrest in these member states.

### 3.2 Bayesian Model Averaging Result

The BMA result (Table 3) shows the list of the top five models (in terms of posterior probability) with the zero-one indicators for variable inclusion (Clyde, 2020). In other words, 1 and 0 mean when the variable is either included or excluded respectively (from any of the models as shown in Table 3). It seems that Model #1 is the model with the highest $R^2$ (0.68). It is also interesting that this is the model that includes all the covariates. Figure 5 shows the image plots of the explanatory variables in the y-axis. The upper x-axis shows the best first 18 models out of the 32 possible models sorted according to posterior probabilities from best (left) to worst (right). Each column represents each model and associated variables that are included. Figure 5 also shows that Model #1 includes all the variables (plus the intercept), while Model #2 includes the intercept and all the variables (except area and population) and so on.

Figure 6 shows the PIPs for all the covariates used in the model prediction. The PIP for area, population, NO$_2$, OMEGA, and PM2.5 is 0.47, 0.61, 0.68, 0.69 and 0.94 respectively. The PIPs greater than 0.5 are all shown in blue color, meaning that these are variables with high inclusion.
3.3 Comparison with Other studies and future perspectives

This study assessed the impacts of environmental variables (PM2.5, NO$_2$, and OMEGA), all of which are extracted from MERRA-2 satellite data on the spatial distribution of COVID-19 fatalities in 21 AL states. To quantify the impacts of these variables on COVID-19 deaths, a BMA modeling approach was used. Using the PIP, the results show that PM2.5 has the largest contribution and importance (PIP ~ 94%) in COVID-19 fatality prediction in the Arab world. The next variables are OMEGA and NO$_2$, which have a PIP of 69% and 68% respectively. Also, the population has a PIP above 50%, meaning that the variable has inclusion probabilities that are important for COVID-19 death prediction. This result is consistent with other studies. Konstantinoudis et al. (2021) used the Bayesian hierarchical models to quantify the effect of air pollution (NO$_2$ and PM2.5 concentration during 2014–2018) on COVID-19 mortality risk. According to the authors, there is strong evidence of NO$_2$ impacts on COVID-19 mortality risk while the PM2.5 effect remains uncertain. Also, Zhang et al. (2021) concluded that there is a statistically significant relationship between air pollution variables (NO$_2$ and PM2.5) and COVID-19 cases. This result is important since these variables can cause other diseases such as poor heart conditions due to cardiovascular disease, which has been reported in other literature (Paramasivam et al., 2020; Bansal, 2020; Mehra, et al., 2020; Liu et al., 2020). Cardiovascular disease is caused by various factors such as tobacco use, poor air quality, hypertension, hypercholesterolemia, and physical inactivity. In the study conducted by Keates et al. (2017) on cardiovascular disease in Africa, the lowest levels of physical activity were found in Mauritania (a member of the AL).

Using the BMA statistical approach, the results show that all these factors contributed to fatality cases using the inclusion probability of each factor but at varying degrees. From the results, it seems that exposure to pollutants such as NO$_2$ and PM2.5 and airflow to disperse the pollutants all contribute to the relatively high COVID-19 fatality rates seen in countries such as the KSA, Egypt, and Iraq. The implication of these findings is that immediate health care should be provided for patients with pre-existing severe heart disease due to poor air quality, as they are the most vulnerable to the adverse impacts of COVID-19. Also, these results show that the AL region must improve health care delivery for the population. Also, factors causing poor air quality should be addressed, as this factor also contributed to the prediction of COVID-19.

This study used pre-existing environmental variables with a temporal scale of 52 months (preceding the COVID-19 probabilities that are important for prediction or data interpretation (Clyde et al., 1999). Although the covariate area has low PIPs, Clyde et al. (2020) reported that this may be the case if there are extreme correlations among the explanatory variables.

Fig. 3 Spatial distribution of antecedent remote-sensed explanatory variables used in predicting COVID-19 fatalities in the Arab League region showing (a) Average daily NO$_2$ Tropospheric Column (30% Cloud Screened) (1/cm$^2$) (b) Monthly Average Map of Dust Column Mass Density - PM 2.5 (kg/m$^2$) and (c) Monthly Averaged Map of Vertical Pressure Velocity monthly (OMEGA) (Pa/s) at 750 hPa
takes non-stationarity in the variables into account should be applied in future studies.

4 Conclusions

When the coronavirus disease (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) first appeared in China in December 2019, there was little information about the underlying impacts of atmospheric and socioeconomic factors in its transmission and distribution, especially across multiple countries. There have been studies that have attempted to associate important

Table 3  A list of the top five models (in terms of posterior probability) with the zero-one indicators for variable inclusion*

|                | Model #1 | Model #2 | Model #3 | Model #4 | Model #5 |
|----------------|----------|----------|----------|----------|----------|
| Intercept      | 1.00     | 1.00     | 1.00     | 1.00     | 1.00     |
| Area           | 0.47     | 1.00     | 0.00     | 1.00     | 0.00     |
| Population     | 0.61     | 1.00     | 0.00     | 1.00     | 1.00     |
| NO₂            | 0.68     | 1.00     | 1.00     | 0.00     | 1.00     |
| OMEGA          | 0.69     | 1.00     | 1.00     | 1.00     | 0.00     |
| PM25           | 0.94     | 1.00     | 1.00     | 1.00     | 1.00     |
| BF             | NA       | 1.00     | 0.76     | 0.73     | 0.52     |
| PostProbs      | NA       | 0.19     | 0.14     | 0.14     | 0.10     |
| R²             | NA       | 0.68     | 0.56     | 0.62     | 0.60     |
| dim            | NA       | 6.00     | 4.00     | 5.00     | 5.00     |
| logmarg        | NA       | 2.35     | 2.08     | 2.04     | 1.70     |

*“BF” is Bayes Factors; “PostProbs” is posterior probability; R² is the coefficient of determination of the models, “dim” is the dimension of the models (i.e. number of coefficients plus the intercept) and the “logmarg” is the log marginal likelihood under the selected prior distribution (Clyde, 2020).

Fig. 4 Map of COVID-19 deaths in the Arab League Region as of July 28, 2020
**Fig. 5** The image plot of model rankings showing which variables are to be included or excluded. The Y-axis shows the model intercept and explanatory variables (OMEGA = vertical pressure velocity, NO₂ = tropospheric nitrogen dioxide, PM2.5 = dust surface column mass density, area, population). The upper x-axis shows the best first 18 models out of 32 possible models sorted according to posterior probabilities from best (left) to worst (right). Each column represents each model and associated variables that are included. The lower x-axis shows the scaled log posterior probability, which equates to the “log Bayes factors”, which is used to compare among two competing models.

**Fig. 6** The posterior inclusion probabilities (PIPs) using Bayesian model averaging on explanatory variables (on COVID-19 fatalities in the Arab world). The response variable is the COVID-19 fatalities as of July 28, 2020.
air quality factors such as NO\textsubscript{2}, PM2.5, and other polluting gases with COVID-19 transmission and distribution across large areas. This study builds on these findings by applying an integrated modeling approach where uncertainty in model development, parameters, and selection are taken into consideration. This study applied the BMA method to predict COVID-19 distribution using NO\textsubscript{2}, PM2.5, OMEGA, population, and geographic area as covariates or explanatory variables. The following findings and concluding remarks are drawn:

1. There is a positive correlation between COVID-19 deaths and NO\textsubscript{2}, PM2.5, OMEGA, and population in AL countries, meaning that any elevated amount in these important air quality indicators could contribute to COVID-19 fatalities. There is almost no correlation between COVID-19 fatalities and geographical areas of the Arab member countries. This shows that how big or small a country is in terms of an area may not contribute to the pandemic.

2. The extracted PIP from the BMA prediction shows that all the air quality indicators and population have marginal PIPs greater than 0.5, meaning that covariates with PIP > 0.5 correspond to what is known as the median probability model. This shows that these variables are important in the prediction of COVID-19 deaths in the Arabian world. For the geographical area, the PIP was 0.47.

3. From the best five models selected using BMA, the ordinary coefficient of correlation (R\textsuperscript{2}) was highest when all five covariates were included in the model prediction. This shows that the relatively low PIP obtained for the geographical area may be due to cross-correlation among the variables. Therefore, this study shows that all five variables can be included in the prediction of COVID-19 deaths in the Arab world using a BMA method.

Finally, these results have practical applications and would be of significant decision-support aid in combating the transmission, distribution, and fatalities from COVID-19. The results show that lowering the NO\textsubscript{2} and PM2.5 levels in combination with important social measures such as washing hands, practicing social distancing, and wearing a face mask could significantly reduce COVID-19 fatalities in the Arab region and the rest of the world.

Acknowledgements The authors would like to thank the three anonymous reviewers for their constructive suggestions.

Author contributions Conceptualization: [Alaba Boluwade], Methodology: [Alaba Boluwade, Amna, M., Al Ruheili], Formal analysis and investigation: [Alaba Boluwade], Writing - original draft preparation: [Alaba Boluwade, Amna, M., Al Ruheili]; Writing - review and editing: [Alaba Boluwade, Amna, M., Al Ruheili], Supervision: [Alaba Boluwade]

Funding This research received no external funding.

Availability of data and material Data will be made available on reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Al-Maamary HM, Kazem HA, Chaichan MT (2017) Climate change: the game changer in the Gulf Cooperation Council region. Renew Sustain Energy Rev 76:555–576

Amini SM, Parmeetr CF (2011) Bayesian Model Averaging in R. J Econ Soc Meas 36(4):253–287

Arvai MZ, Prasad A, Katayama MK (2014) Macroprudential policy in the GCC countries. International Monetary Fund

Bansal M (2020) Cardiovascular disease and COVID-19. Diabetes Metab Syndr 14(3):247–250. doi: https://doi.org/10.1016/j.dsx.2020.03.013

Boluwade A (2020) Regionalizing & Partitioning Africa’s Coronavirus (COVID-19) Fatalities Using Environmental Factors and Underlying Health Conditions for Social-economic Impacts. 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), Giza, Egypt, 2020, pp. 439–443, doi: https://doi.org/10.1109/NILES0944.2020.9257875

Callen MT, Cherif R, Hasanov F, Hegazy MA, Khandelwal P (2014) Economic diversification in the GCC: Past, present, and future. International Monetary Fund

Chan KHJS, Malik Peiris SY, Lam LLM, Poon KY 1, Yuen (2011) and W. H. Seto. The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. https://doi.org/10.1155/2011/734690

Chauhan A, Singh RP (2020) Decline in PM2.5 concentrations over major cities around the world associated with COVID-19. Environmental Research 109634, Volume: 187, DOI: https://doi.org/10.1016/j.envres.2020.109634

Chauhan AK, Singh RP (2021) Effect of Lockdown on HCHO and Trace Gases over India during March 2020, AAQR, Volume 21, Issue 4, Article Number 200445, DOI https://doi.org/10.4209/aaqr.2020.07.0445

Clyde M (2021) BAS Bayesian Adaptive Sampling for Bayesian Model Averaging. Recent R package version. https://cran.r-project.org/web/packages/BAS/vignettes/BAS-vignette.html

(Accessed on 20 July 2021)

Clyde M (1999) Bayesian Model Averaging and Model Search Strategies (with discussion). In: Bernardo JM, Dawid AP, Berger JO, Smith AFM (eds) Bayesian Statistics 6. Oxford University Press, pp 157–185

Contini D, Costabile F (2020) Does Air Pollution Influence COVID-19 Outbreaks? Atmosphere, 11, 377

Conticini E, Frediani, Bruno, Caro D (2020) Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in northern Italy?Environ. Pollut 114465:0269–7491. https://doi.org/10.1016/j.envpol.2020.114465

Copernicus (2020) Air Quality. Available at: https://atmosphere.copernicus.eu/air-quality. Accessed on December, 2020
of The Total Environment, Volume 756, 144020, https://doi.org/10.1016/j.scitotenv.2020.144020
Singh RP, Chauhan AK (2020) Impact of lockdown on air quality in India during COVID-19 pandemic, Air Quality, Atmosphere and Health, https://doi.org/10.1007/s11869-020-00863-1, Volume: 13, Issue: 8, Pages: 921–928, DOI: 10.1007/s11869-020-00863-1
Sowers J, Vengosh A, Weinthal E (2011) Climate change, water resources, and the politics of adaptation in the Middle East and North Africa. Clim Change 104:599–627. https://doi.org/10.1007/s10584-010-9835-4
Vlahovic B, Mugesi R, Radojevic V (2018) The Arab League—an opportunity to improve the EU cheese export (No. 2038-2018-3495)
Weichmann D et al (2020) Autopsy Findings and Venous Thromboembolism in Patients With COVID-19. A Prospective Cohort Study. Ann Intern Med. https://doi.org/10.7326/M20-2003
Woodby B, Arnold MM, Valacchi G (2020) SARS-CoV-2 infection, COVID-19 pathogenesis, and exposure to air pollution: What is the connection? [published online ahead of print, 2020 Oct 6]. Ann N Y Acad Sci. https://doi.org/10.1111/nyas.14512. doi:10.1111/nyas.14512
World Bank. Population, total. https://data.worldbank.org/indicator/SP.POP.TOTL (Accessed on 20 July 2020)

World Health Organization. Cardiovascular Diseases. Available: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1. (Accessed on 20 July 2020)
Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, Huang H, Zhang L, Zhou X, Du C et al (2020a) Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med. https://doi.org/10.1001/jama.2020.0994
Wu X, Nethery, Rachel C, Sabath BM, BraunDanielle, Dominici, Francesca (2020b) Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. Science Advance. Vol 6, Issue 45. https://doi.org/10.1126/sciadv.abc4049
Zeugner S (2020) Bayesian Model Averaging with BMS for BMS Version 0.3.0. http://bms.zeugner.eu/tutorials/bms.pdf. (Accessed on 20
Zhang X, Tang M, Guo F, Wei F, Yu Z, Gao K, Jin M, Wang J, Chen K (2021) Associations between air pollution and COVID-19 epidemic during quarantine period in China, Environmental Pollution, Volume 268, Part A, 115897, https://doi.org/10.1016/j.envpol.2020.115897

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.