Effect of temperature on the spread of Covid-19 in Qatar, Kuwait and other Gulf countries

Maha Abujazar[1], Shafiqah Al-Awadhi [2],
Mustapha Rachdi[3] , Halima Bensmail[4,*]

[1] *CSE Dept., Hamad Bin Khalifa University, Qatar*
[2] *Dept. of Statistics and Operations research, Faculty of Science, Kuwait University*
P.O Box 5969, Safat 13060, Kuwait

[3] *CNRS/UGA/OrangeLabs Telecom4Health, Faculty of Medicine, University Grenoble Alpes (UGA), 38700 La Tronche, France*

[4] *Qatar Computing Research Institute, Hamad Bin Khalifa University*
*Corresponding author: hbensmail@hbku.edu.qa

Abstract

Background: COVID-19 has emerged as a serious pandemic that emerged during since the end of 2019. The dissemination and survival of coronaviruses have been demonstrated to be affected by ambient temperature in epidemiological and laboratory research. The goal of this investigation was to see if temperature plays a role in the infection produced by this novel coronavirus.

Methods: Between March 29, 2020, and September 29, 2020, daily confirmed cases and meteoro-logical parameters in many Gulf countries were collected. Using a generalized additive model, we investigated the nonlinear relationship between mean temperature and COVID-19 confirmed cases.. To further investigate the association, we employed a piecewise linear regression.

Results: According to the exposure-response curves, the association between mean temperature and COVID-19 cases was nearly linear in the window of $21 - 30°C$ while it is almost flat beyond that window. When the number was below $21°C$ (lag 0-14), each $1°C$ increase was associated with a 4.861 percent (95 percent CI: $3.209 - 6.513$) increase in mean temperature (lag 0-14). Our sensitiv-ity analysis confirmed these conclusions.

Conclusions: Our findings show a positive linear association between mean temperature and the number of COVID-19 cases with a threshold of $21°C$. There is little evidence that COVID-19 case numbers would rise as the weather becomes colder, which has important consequences for making health strategy and decision.

Keywords: COVID-19; forecasting; GAM; pandemic; stepwise regression

1. Introduction

During the end of 2019, a new type of Coronaviruses has been discovered called Covid-19. The virus was discovered first in Wuhan city in China, then expanded worldwide. The World Health Organization (WHO) has proclaimed SARS-CoV-2 (Severe Acute Respiratory Syndrome) to be a pandemic (WHO 2021). SARS affects mainly the respiratory system of the patients and the main syndromes which are common between patients are fever, dry cough, dyspnea, headache, and hypoxemia (Koichi 2020). Several studies have been proposed to model this infection without involving other variables and using well known compartment models (Oshinubi et al 2021), (Demongeot et al 2021) and (Sallahi et al 2021). Other models use time series and non-parametric approach such as functional data analysis (Oshinubi et al 2022) and (Demongeot et al 2022) to name a few. It has been well known that temperature, humidity, and ventilation play a role in the persistence, infectivity, dispersal, and elimination of viruses, as well as altering human defensive mechanisms against respiratory diseases (Moriyama 2020). Cold temperature

and low relative humidity are also believed to increase the transmission of other respiratory viruses, including respiratory syncytial virus (RSV), human rhinovirus, and avian influenza virus (Wolkoff 2018) In the process of studying SARS epidemic, researchers have been tested the virus stability on the smooth surfaces at temperature of 22 to 25 Celsius and humidity of 40 to 50 percentage. It was found that the risk of SARS epidemic is increased in the low temperature atmosphere and low humidity. Thus, the incidence of SARS in the lower weather days increased. Our aim is to check whether the relationship between COVID-19 confirmed cases and the volatility in the weather is ascertain.

2. Materials and Methods

2.1 Study area

Our study included several countries from Middle and North of Africa Fig. 1 shows the Middle East map of the number of infections during March-June 2021. We focused on these country since the meteorological data we have obtained was limited.
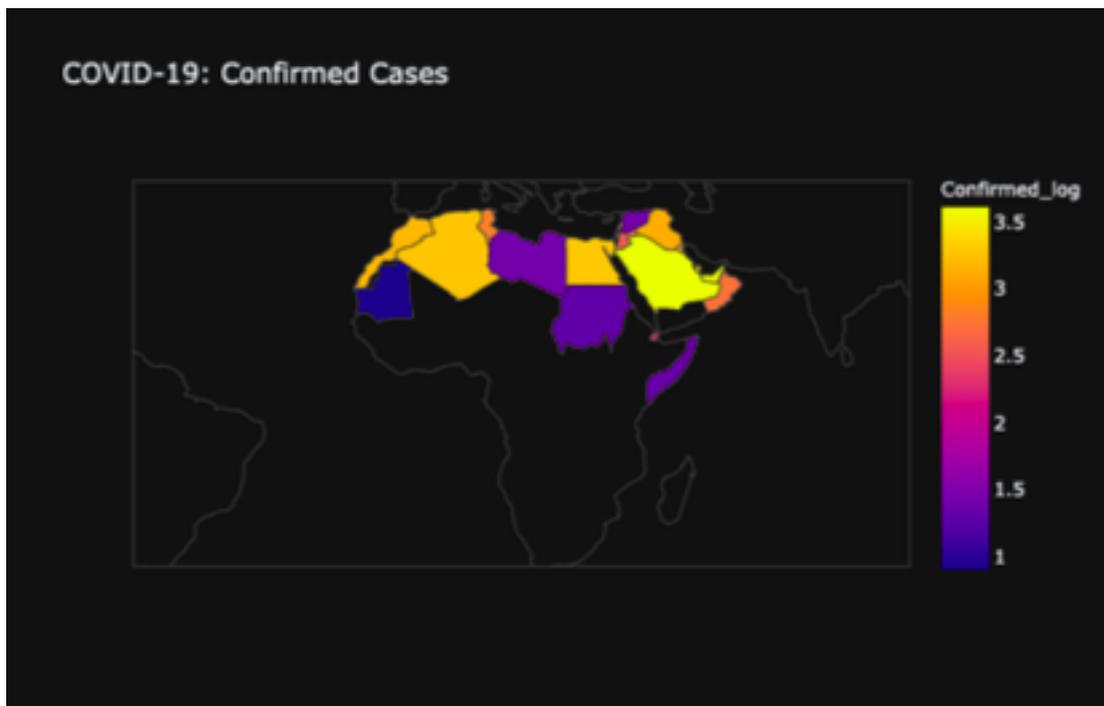
Fig. 1. Middle East map for infection during March-June 2021

2.2 Data collection

Daily confirmed cases were gathered from the Worldometers official websites (Worldometer 2021) https://www.worldometers.info/coronavirus/ corresponding to countries between March 23, 2020 (i.e., the lockdown of most GCC and MENA countries) to September 29, 2020. We obtained the data after the closure of those countries to minimize the potential inclusion of imported cases in this study.

Meteorological data during the same study period for each country were collected from the World Meteorological Organization (https://worldweather.wmo.int/en/home.html) (Worldweather 2021). We also included meteorological factors such as daily mean temperature, humidity, air pressure, and wind speed. To adjust for changes in the age distribution of the southern population, we calculated directly age-standardized infection rate for each day of the study (Clancy et al. 2002). Daily infection was tabulated and 95% percentile of the distribution calculated. Days on which l-day running average of infection was above the 95th percentile were diagnosed as "epidemic" days.

2.3 Generalized additive model.

Our main goal was to determine the shape of the infection-temperature relationship while accounting for potential confounders. One can use the logarithm of infection transformation as dependent variable and then fit a linear function of the temperature and other predictors on the transformation. Unfortunately we cannot use this assumption from what we explained earlier. Therefore, in this paper, we assume that log infection is a nonlinear smooth function. Rather than summarizing the temperature-infection relationship with a single risk for all temperatures, we generated a relative risk estimate that was a smoothly evolving function of temperature. In this case, we propose a generalized additive model (GAM) which uses temperature as a response variable and other environmental variables (Air Pressure, Wind and Humidity) as explanatory variables since a plausible relation is established (see Table 1). Environmental variables may play a role in determining the actual relationship between the number of infections and air temperatures ((Moatar and Gailhard 2006), (Webb et al. 2003) and (Belanger et al. 2005)).

The generalized additive model (GAM) was proposed by Hastie and Tibshirani ((Hastie and Tibshirani 1986)) and is an extension of the generalized linear model. GAM is useful to investigate the nonlinear relationship between weather features and health-related consequences (Liu et al. 2020, Wu et al. 2018, Lin et al. 2018, Peng et al. 2006 and Talmoudi et al. 2017). General linear model assume a strong assumption of parameters linearity while generative additive model does not requires this assumption. Its concept is to characterize the response using an additive nonlinear functions association, which allows for more comprehensive modeling of the impact of the explanatory variables. This specificity makes it a popular instrument in modelling the impact of environmental variables because these effects are often nonlinear and are difficult to specify parametrically (Bruneau and Gregoire 2002) and (Peng and Dominici 2002). The literature review of (Jbilou and El Adlouni 2002) provide a review of generative additive model and its impact in identifying nonlinear relationship in environmental and health investigations. The generalized additive model (GAM) is a semi-parametric refinement of the generalized linear model (GLM), and it could be used to inspect the nonlinear association between weather and health outcomes. This model is written as follows:

$$E(\mathbf{y}) = \beta_0 + \mathbf{f}_1(\mathbf{x}_1) + \mathbf{f}_2(\mathbf{x}_2) + \ldots + \mathbf{f}_p(\mathbf{x}_p) + \epsilon \tag{1}$$

This model's implementation necessitates the estimation of smooth irregular functions $\mathbf{f}_i(\mathbf{x}_i)\, i = 1, \ldots, p$ for each independent variable $\mathbf{x}_i$.

To circumvent the problem of overfitting, the penalized GAM was used where smoothing functions f_i are cubic or penalized smoothing splines These splines are estimated to be the solution to the optimization problem (2). The solution requires a twice continuously differentiable function which minimizes the sum of the penalized squares (penalized residual sum of squares):

$$\arg\min_{\mathbf{f}_i} \left(\|\mathbf{y} - \mathbf{f}_i(\mathbf{x}_i)\|^2 + \sum_i \lambda_i \int \mathbf{f}_i''(\mathbf{x})^2 d\mathbf{x} \right) \tag{2}$$

Equation (2)'s first component measures data fit, whereas the second term penalizes deflection in the function. Smoothing parameters $\lambda_1, \ldots, \lambda_p$ represent the penalties on the irregularity of the fitted function associated with each explanatory variable $\mathbf{x}$. They control the smoothing level of each function f_i and the compromise between the bias and the variance. The main idea is to place far more nodes than are required and then penalize those nodes that would provide insufficient information.

The optimization problem (2) has a parametric description based on a cubic splines decay of $\mathbf{f}_i$ functions using the following form:

$$\mathbf{f}_i = \mathbf{B}_i \beta_i \quad \text{where} \quad \mathbf{B}\beta = \sum_j \mathbf{b}_j \beta_j \tag{3}$$

The optimization problem (2) becomes:

$$\arg\min_{\beta_1, \ldots, \beta_p} \left(\|\mathbf{y} - \beta_0 + \sum_i^p \mathbf{B}_i \beta_i\|^2 + \sum_i^p \lambda_i \int \{(\mathbf{B}_i \beta_i)''\}^2 d\mathbf{x} \right) \tag{4}$$

where $\mathbf{f}_i$ is the ith regression function in the additive model, $\mathbf{f}''$ denotes the second derivative of $\mathbf{f}$, $\|\|$ is the Euclidean norm and $\lambda_1, \ldots, \lambda_p$ are penalty terms.

Equation (4) is solved using the iterative reweighted least squares algorithm (IRLS). The problem performed in each iteration is similar to a weighted least squares problem in which the second derivative $\mathbf{f}_i$ is designed to define the smoothing function. Roughness of $\mathbf{f}_i$ depends on the magnitude of $\mathbf{f}_i''$. A straight line is characterized by a second derivative equal to zero. An estimate of λ_i is chosen to optimize a fit with an adjusted bias and variance. The choice of these parameters is therefore done by using generalized cross-validation (GCV) and performance metrics. In fact smoothing parameters estimation is done by evaluating AIC or BIC information criterias (Akaike 1973) and BIC (Schwarz 1978) or GCV (generalized Cross Validation). Here we used Generalized Cross Validation which is obtained by the following:

$$GCV(\lambda) = \frac{1}{p} \sum_{i=1}^{p} \left\{ \frac{\mathbf{y}_i - \mathbf{x}_i \hat{\beta}}{1 - \mathrm{trace}(\mathbf{S}_\lambda)/p} \right\}^2 \tag{5}$$

where $\mathbf{S}_\lambda$ is the smoother where $\mathbf{S}_{ij} = \partial \hat{\beta}_i / \partial \tilde{\beta}_j$ and $\mathrm{trace}(\mathbf{S}) = \sum_i \mathbf{S}_{ii}$ is a measure of effective degree of freedom. We estimated this smooth function by using a GAM, which fits a cubic spline function of temperature GAMs or cubic splines using statistical software packages R-package. GAMs in our analysis were implemented with the mixed computation vehicle with automatic smoothness estimation (mgcv package; version 1.8-28) of R software (version 3.5.2). This package uses the generalized cross validation to find the optimal smoothing function related to each explanatory variable (Wood 2006). The threshold of the smoothing functions of explanatory variables to be retained is 5%.

2.4 Statistical analysis.

Because the temperature effect could last for several days and the period between exposure to an infection span from one day to twenty eight days,

A moving-average procedure is a feasible solution to account for the cumulative lag effect of temperature ((Duan et al. 2019), (Li et al. 2020), (Lu et al. 2015)). As a consequence, a GAM with a Gaussian distribution family was used in this work ((Hastie and Tibshirani 1986)) to examine the moving average lag effect (lag 0-7, lag 0-14, lag 0-21, lag 0-28) of mean temperature on infected cases. We should mention that the temperature effect could last for several days and the period between exposure to an infection span from one day to twenty eight days.

	Mean (SD)	Min	Max
Confirmed cases	1035 (227.082)	146	1632
Mean Temp (°C)	33.71(10.286)	28	45
Humidity (%)	43.00(17.386)	6	80
Air pressure (hPa)	997(75.816)	996	1020
Wind speed (m/s)	7.00(1.199)	11.4	13.5

Table 1. Summary statistics of newly confirmed cases and meteorological variables across Gulf Cooperation Council countries and days.

The model is defined as:

$$\begin{aligned}
\log \mathbf{y}_{it} = {} & \beta_0 + \mathbf{f}_1(\mathbf{Tmp}_{il}) + \mathbf{f}_2(\mathbf{Hum}_{il}) + \mathbf{f}_3(\mathbf{AirP}_{il}) \\
& + \mathbf{f}_4(\mathbf{Win}_{il}) + \mathbf{f}_5(\mathbf{y}_{i,t-1}) + \mathbf{Cty}_i + \mathbf{day}_t + \epsilon_{it}
\end{aligned} \tag{6}$$

In the model, $\log \mathbf{y}_{it}$ is the log-transformed counts in country i on day t (to which we add one since logarithmic function is only defined for $\mathbf{y} > 0$), β_0 is the intercept and $\mathbf{f}_j$ designates a fine layer spline function with the maximum 2 degrees of freedom to circumvent overfitting, $\mathbf{Tmp}_{it}$ is the $\mathbf{l+1}$ -day

rolling mean term $\mathbf{lag}0 - l$ of daily mean temperature in country $\mathbf{i}$. We also controlled the relative humidity $\mathbf{Hum}_{il}$, air pressure $\mathbf{AirP}_{il}$ and wind speed $\mathbf{Win}_{il}$ during the same period for the possible lurking variable. The dependent variable $\log(\mathbf{y})_{i,t-1}$ determines the logarithm transformation of the number of infections lagged one day in country i to account for possible correlation in the data (Liu et al. 2020). The variable $\mathbf{Cty}_i$ is the country fixed effect variable and $\mathbf{day}_t$ apprehends day fixed effect. We then fitted the piecewise log-linear model on the top of fitting GAM with polynomial spline as the following:

$$
\begin{aligned}
\log \mathbf{Ey}_{it} \ = \ & \beta_0 + \beta_1 \mathbf{Tmp}_{il} + \beta_2 (\mathbf{Tmp}_{il} - \xi)_+ + \mathbf{f_{Hum}}(\mathbf{Hum}_{il}) \\
& + \mathbf{f_{AirP}}(\mathbf{AirP}_{il}) + \mathbf{f_{Win}}(\mathbf{Win}_{il})
\end{aligned} \tag{7}
$$

Here $(\mathbf{Tmp}_{il} - \xi)_+ = \max(\mathbf{Tmp}_{il} - \xi, 0)$.
This model is the same as

$$
\begin{aligned}
\log \mathbf{Ey}_{it} \ = \ & \beta_0 + \beta_1 \mathbf{Tmp}_{il} + \mathbf{f_{Hum}}(\mathbf{Hum}_{il}) \\
& + \mathbf{f_{AirP}}(\mathbf{AirP}_{il}) + \mathbf{f_{Win}}(\mathbf{Win}_{il})
\end{aligned} \tag{8}
$$

if the temperature is less than the threshold value ξ. It is also same as

$$
\begin{aligned}
\log \mathbf{Ey}_{it} \ = \ & \beta_0 - \beta_2 \xi + (\beta_1 + \beta_2) \mathbf{Tmp}_{il} + \mathbf{f_{Hum}}(\mathbf{Hum}_{il}) \\
& + \mathbf{f_{AirP}}(\mathbf{AirP}_{il}) + \mathbf{f_{Win}}(\mathbf{Win}_{il})
\end{aligned} \tag{9}
$$

if the temperature is greater than the threshold value ξ.
Equations (8) and (9) are constrained to meet at the threshold value ξ. The aforementioned model is also known as a B-mode splined linear regression model. The covariates in our study include Country ($\mathbf{Cty}$), $\mathbf{day}$ of week (day), daily average temperature ($\mathbf{Tmp}$) as representative variables of daily temperature. We identified the threshold values rendering the model with the lowest AIC (Akaike 1973) and BIC (Schwarz 1978).

3. Result

3.1 Descriptive statistics
The degree of smoothness of the estimated infection-temperature relative risk curve is controlled by its number of degrees of freedom. When fitting a linear regression, one uses one degree of freedom for accounting for the slope and a two degree of freedom when fitting a quadratic curve (one for its slope and one for its curvature). To allow for extremely nonlinear shapes, we used six degrees of freedom to describe the association of infection with each weather variable.
We began with a substantial preliminary analysis of GAMs with natural cubic spline in which several types of representative values, including daily maximum temperature, average temperature, air pressure, wind and humidity were considered as predictors of daily infection. We applied Akaike's Information Criterion (Akaike 1973) and (Schwarz 1978) for the model-fitting criteria. We selected daily average temperature, daily average relative humidity, air pressure, wind speed all with a 7-day lag (lag 0-7), 14-day lag (lag 0-14), 21-day lag (lag 0-21), 28-day lag (lag 0-28).
Table 1 summarizes the descriptive statistics of all the variables. In this study, we used cases during the period (May 22, 2021 to June 30, 2021) and the average number was 1035 per day. Average daily mean temperature, relative humidity, air pressure, and wind speed were 33.71°C, 43.00%, 997.00 hPa, and 7.00 m/s, respectively.
Table 2 shows the pearson coefficient of correlation between several variables. Mean temperature is moderately significantly correlated (positively) with relative humidity ($r = -0.354$) and with air pressure ($r = 0.163$), using a level of significance $\alpha = 0.05$. Mean temperature was negatively correlated with wind speed ($r = -0.053, p < 0.05$).

Var	Tmp	Wind	AirP	Hum
Tmp	1.00	−0.053*	0.163**	0.053*
Wind	−0.053*	1.00	0.274*	−0.117
AirP	0.163**	0.274	1.00	0.127
Hum	0.053	−0.117	0.127	1.00

Table 2. Pearson coefficient of correlation between environmental variables. *, ** and *** indicate significance …

3.2 Effect of temperature on infections

The exposure-response curves in Figure 2 show that the relationship between temperature and infection cases was significantly nonlinear (lag 0-14, with p < 0.001). Specifically, the relationship was approximately linear in the window of 21°C (70°F) for Qatar and became flat above, suggesting that the single threshold of the temperature effect on COVID-19 was 21°C. For Kuwait and Saudi Arabia, we found that the temperature-infection associations using average temperature with a lag of 7 and 14 day exhibited a J- shaped formation; that is, there was no association or a small negative association for the temperature range below the threshold and a positive association for the temperature around 26°C (80°F).
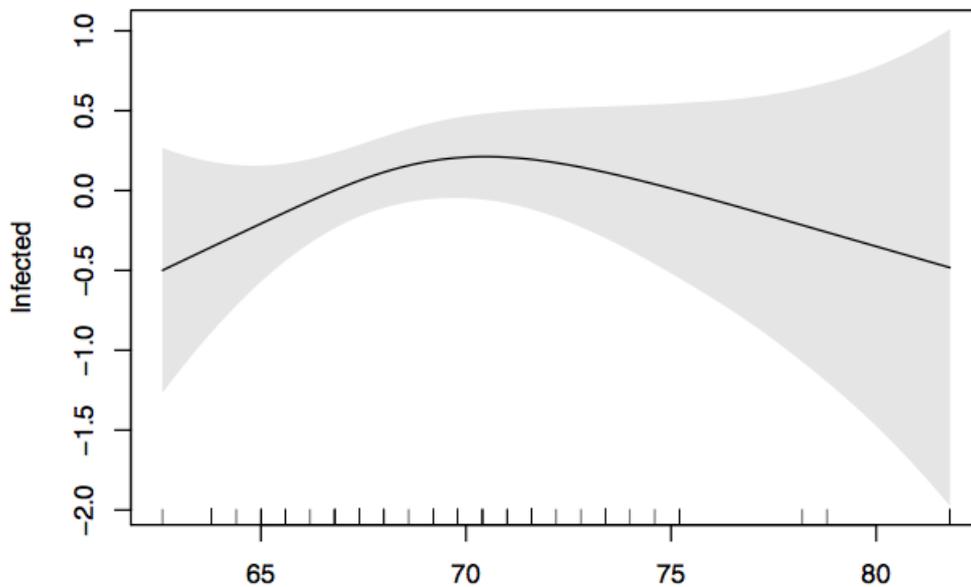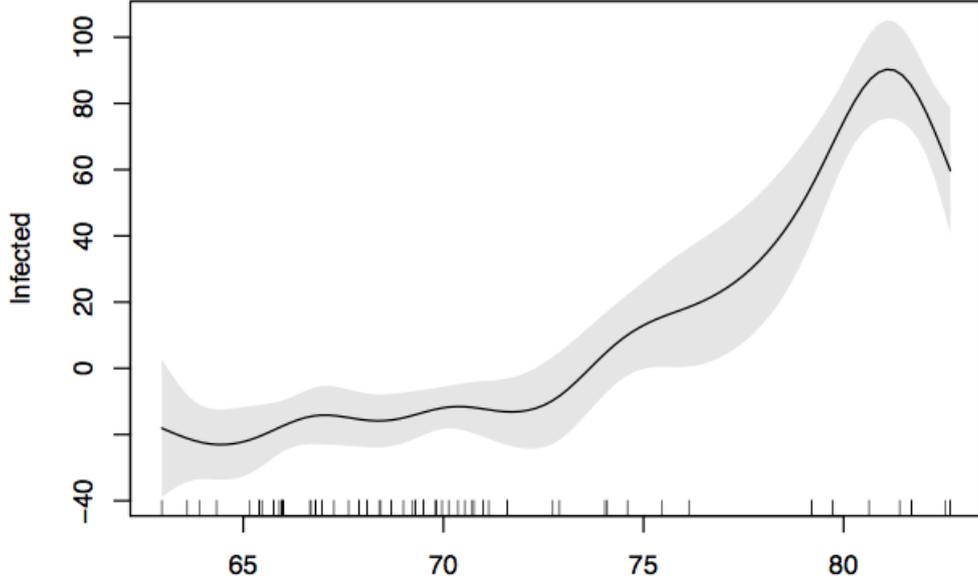

Fig. 2. Exposure-response curves for the effects of temperature on COVID-19 confirmed cases for Qatar. The x axis is the mean temperature (F) (14-day rolling average chosen by AIC and BIC). The y axis expresses the effect of the smoother to the fitted values

Based on GAM's outcome, a piecewise linear regression was adjusted with a specific value at a 21°C to quantify the impact of temperature above and below the threshold for Qatar and Oman (26 for Kuwait and KSA). As showed in Table 4, each 1°C rise in mean temperature ed to a 3.471% (95% CI: 2.277-4.586) increase in the daily number of COVID-19 confirmed cases when mean temperature was below 21°C for Kuwait. This positive effect is largest for Oman (percentage change = 6.949%, 95% CI: 2.292-9.205).

4. discussion

Using a generalized additive model, we investigated the nonlinear connection between ambient temperature and COVID-19 confirmed instances in this work. When the mean temperature was in the interval [20-26]°C, the exposure-response relationship was positive linear, but it became flat above 26°C, indicating that greater temperatures may not impede the transmission of this unique coronavirus.

Fig. 3. Response curves linking temperature with confirmed cases for Kuwait. The x axis is the mean temperature (F) (14-day moving average chosen by AIC and BIC). The y axis indicates the effect of the smoother function to the fitted values.

We compared GAM with linear and non-linear approaches such as Random Forest (RF), Support Vector Machine (SVM), Neural Network(NN), Tree, Gradient Boosting (GBM) and XGradient Boosting (XGBM) and Linear Regression (GLM). Table 3 summarizes the findings for several models that used the 10-fold cross-validation (GCV) process to predict the log-transformed COVID-19 counts, with performance metrics like MSE, RMSE, used to assess model performance. The Random Forest regressor produced the best result after GAM using the metrics MSE and RMSE ($1.328 \times 10^3, 823.22$) respectively. The models with the least performance are found to be Gradient Boosting Regressor (GBM) and XGradient Boosting (XGBM).

Since random forest performed best after GAM, we looked at the features that play major roles when compared to the other features using random forest feature selector algorithm and Figure 5 summarizes features importance in absolute value. We can see that temperature ranked at the top followed by humidity, wind and air pressure. Previous studies found that temperature plays a role in the survival and

Model.	MAE	MSE	RMSE
GAM	**205.19**	**1.011 $\times 10^3$**	798.10
RF	**268.82**	1.328×10^3	823.22
SVM	274.92	1.544×10^3	854.44
GLM	283.16	1.569×10^3	885.28
Tree	301.56	2.001×10^3	1012.12
NN	271.49	1.981×10^3	890.58
GBM	468.15	3.522×10^3	974.97
XGBM	461.07	2.966×10^3	937.21

Table 3. Results of the state-of-the-art machine learning algorithms for prediction of infection counts of COVID-19. MAE is the mean absolute error, MSE is the mean square error and RMSE is the root mean square error.

transmission of other coronaviruses, such as SARS-CoV and MERS-CoV. Based on data from Hong Kong, Guangzhou, Beijing, and Taiyuan, Tan et al. (2005) discovered that the best environmental tem-
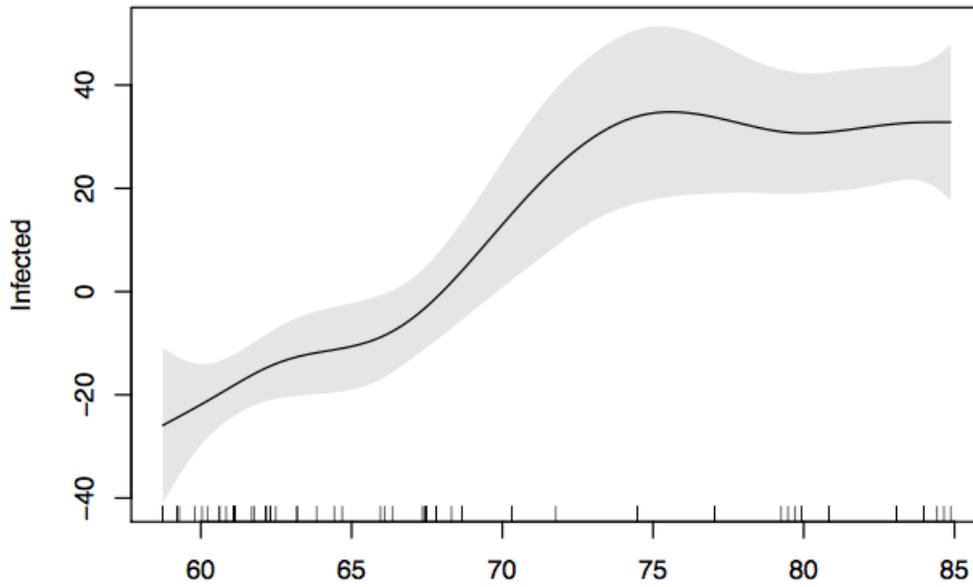
Fig. 4. Exposure-response curves for the effects of temperature on COVID-19 confirmed cases for Kuwait. The x axis is the mean temperature (F) (14-day moving average chosen by AIC and BIC).
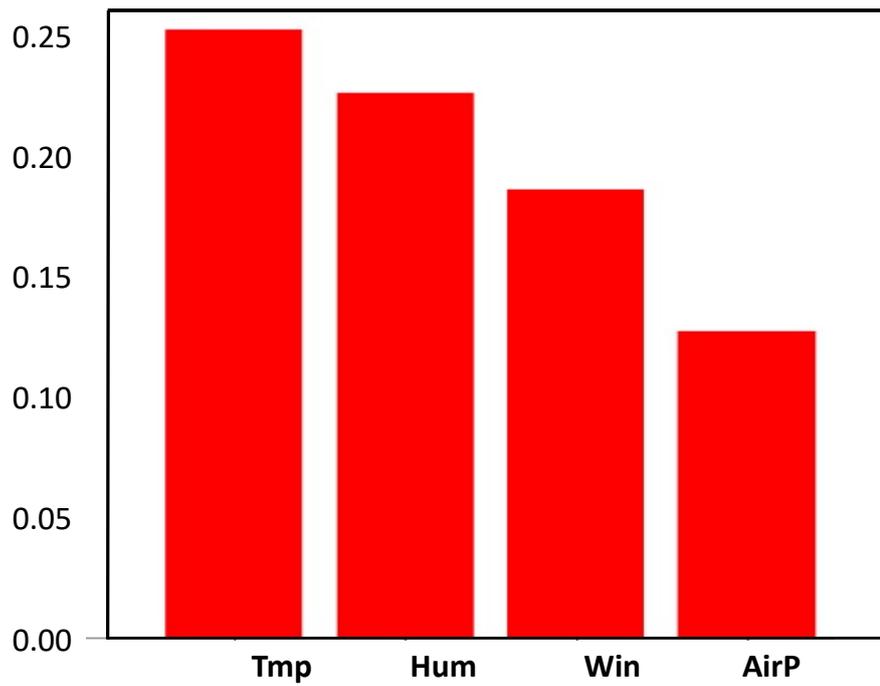
Fig. 5. Features importance.

Country	AIC lag 0-7	AIC lag 0-14	AIC lag 0-21	AIC lag 0-28	BIC lag 0-7	BIC lag 0-14	BIC lag 0-21	BIC lag 0-28	Decision
Qatar	609.963	**605.281**	690.013	620.856	623.113	630.462	**604.095**	636.216	
SA	649.325	**638.687**	656.452	669.445	664.366	**653.987**	670.995	682.326	
UAE	**655.380**	689.727	691.520	475.462	671.003	**493.333**	574.656	587.097	
Kuwait	483.855	**445.407**	445.550	648.500	501.206	**462.353**	494.309	562.855	
Oman	422.036	**388.542**	442.521	391.734	436.457	404.020	355.675	**302.301**	
Bahrain	492.177	**444.057**	484.939	476.580	506.277	488.570	**448.144**	489.083	

Table 4. AIC and BIC scores for lag 0-7, lag 0-14, lag 0-21 and lag 0-28.

	Mean temperature $\leq \xi°C$		Mean temperature $> \xi°C$	
	Percentage change	95% Confidential interval	Percentage change	95% Confidence Interval
Kuwait	3.471*	(2.322 − 4.619)	−1.151	(−2.747 − 0.446)
Saudi Arabia	4.145**	(3.204 − 4.486)	−1.305	(−2.572 − 0.961)
Oman	6.949*	(2.292 − 9.205)	−2.606	(−5.245 − 0.032)
Qatar	4.471*	(3.322 − 4.919)	−1.216	(−2.142 − 0.544)

Table 5. The effects of a 1° C increase in mean temperature on COVID-19 confirmed cases for Qatar.

References

Wiener, N. (1949).Extrapolation, Interpolation, and Smoothing of Stationary Time Series. The MIT Press: Cambridge (Mass.), USA.

World Health Organization. (2021) Available online: https://www.who.int/.

Koichi Yuki, Miho Fujiogi, and Sophia Koutsogiannaki. (2020). COVID-19 pathophysiology: A review. *Clinical Immunology*, **215**, 108425.

Worldometer (2021) Available online: https://www.worldometers.info/coronavirus/ (accessed on 10th April 2021).

Worldweather (2021) Available online: https://www.worldweatheronline.com/country.

Jahangiri, M., Jahangiri, M., & Najafgholipour, M. (2020). The sensitivity and specificity analyses of ambient temperature and population size on the transmission rate of the novel coronavirus (COVID-19) in different provinces of Iran. *The Science of the total environment*, **728**, 138872.

Chan KH, Peiris JS, Lam SY, et al. The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus (2011). *Advances in virology*. 2011:734690.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguière, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Müller, S., ... Doerr, H. W. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *The New England Journal of Medicine*, **348(20)**, 1967-76.

Xie, J., and Zhu, Y. (2020). Association between ambient temperature and COVID-19 infection in 122 cities from China. *The Science of the total environment*, **724, 138201**. https://doi.org/10.1016/j.scitotenv.2020.138201

Sajadi, M. M., Habibzadeh, P., Vintzileos, A., Shokouhi, S., Miralles-Wilhelm, F., and Amoroso, A. (2020). Temperature, Humidity, and Latitude Analysis to Estimate Potential Spread and Seasonality of Coronavirus Disease 2019 (COVID-19). *JAMA network open*, **3(6)**, e2011834. https://doi.org/10.1001/jamanetworkopen.2020.11834

Kang, J. Y., Michels, A., Lyu, F., Wang, S., Agbodo, N., Freeman, V. L., and Wang, S.(2020). Rapidly measuring spatial accessibility of COVID-19 healthcare resources: a case study of Illinois, USA. *International Journal of Health Geographics,* **19, 36**. https://doi.org/10.1186/s12942-020-00229-

Moriyama M., Hugentobler W.J., Iwasaki A.(2020) Seasonality of respiratory viral infections (2020). *Annu Rev Virol.* doi: 10.1146/annurev-virology-012420-022445. [PubMed] [CrossRef] [Google Scholar]

Wolkoff P. Indoor air humidity, air quality, and health: an overview (2018). *Int J Hyg Environ Health*; **221(3)**:376?390. [PubMed].

Webb, B.W., Clack, P.D., and Walling, D.E., 2003. Water-air temperature relationships in a Devon river system and the role of flow (2003). *Hydrological Processes*, **17 (15)**, 3069?3084. doi:10.1002/(ISSN)1099-1085

Belanger, M., et al., (2005). Estimation de la température de l?eau de rivière en utilisant les réseaux de neurones et la régression linéaire multiple. Revue des sciences de l'eau. *Journal of Water Science*, **18 (3)**, 403-421.

Water temperature behaviour in the River Loire since 1976 and 1881 (2006). *C. R. Geoscience*, **338**; 319?328.

Lin, H., Tao, J., Kan, H., Qian, Z., Chen, A., Du, Y., Liu, T., Zhang, Y., Qi, Y., Ye, J., (2018). Ambient particulate matter air pollution associated with acute respiratory distress syndrome in Guangzhou, China. *J. Expo. Sci. Environ. Epidemiol.* **28**, 392. https://doi.org/10.1038/s41370-018-0034-0.

Liu, K., Hou, X., Ren, Z., Lowe, R., Wang, Y., Li, R., Liu, X., Sun, J., Lu, L., Song, X., (2020). Climate factors and the East Asian summer monsoon may drive large outbreaks of dengue in China. *Environ. Res.*, 109190 https://doi.org/10.1016/j.envres.2020.109190.

Peng, R.D., Dominici, F., Louis, T.A., (2006). Model choice in time series studies of air pollution and mortality. *J. Roy. Stat. Soc. Ser. A.* (Stat. Soc.) **169**, 179-203. https://doi.org/10.1111/j.1467-985X.2006.00410.x

Talmoudi, K., Bellali, H., Ben-Alaya, N., Saez, M., Malouche, D., Chahed, M.K., (2017). Modeling zoonotic cutaneous leishmaniasis incidence in central Tunisia from 2009?2015: forecasting models using climate variables as predictors. *PLoS Negl. Trop. Dis.* **11**, e0005844. https://doi.org/10.1371/journal.pntd.0005844.

Wu, X., Lang, L., Ma, W., Song, T., Kang, M., He, J., Zhang, Y., Lu, L., Lin, H., Ling, L., (2018). Non-linear effects of mean temperature and relative humidity on dengue incidence in Guangzhou, China. *Sci. Total Environ.* **628**, 766-771. https://doi.org/10.1016/j. scitotenv.2018.02.136.

Duan, Y., Liao, Y., Li, H., Yan, S., Zhao, Z., Yu, S., Fu, Y., Wang, Z., Yin, P., Cheng, J., (2019). Effect of changes in season and temperature on cardiovascular mortality associated with nitrogen dioxide air pollution in Shenzhen, China. *Sci. Total Environ.*, **697, 134051**. https://doi.org/10.1016/j.scitotenv.2019.134051.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K.S., Lau, E.H., Wong, J.Y., (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus infected pneumonia. *New Engl. J. Med.* https://doi.org/10.1056/NEJMoa2001316.

Lu, F., Zhou, L., Xu, Y., Zheng, T., Guo, Y., Wellenius, G.A., Bassig, B.A., Chen, X., Wang, H., Zheng, X., (2015). Short-term effects of air pollution on daily mortality and years of life lost in Nanjing, China. *Sci. Total Environ.* **536**, 123?129. https://doi.org/10.1016/ j.scitotenv.2015.07.048.

Hastie, T. and Tibshirani, R., 1986. Generalized additive models. *Statistical Science*, **1**, 297?310. doi:10.1214/ss/ 1177013604.

Akaike, H., (1973). Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest, Hungary, 267-281.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**:461-464.

Clancy L, Goodman P, Sinclair H, Dockery DW (2002). Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet*.19; 360(9341):1210-4. doi: 10.1016/S0140-6736(02)11281-5. PMID: 12401247.

Jbilou, J. and El Adlouni, S. (2012). Generalized Additive Models in Environmental Health: A Literature Review, Novel Approaches and Their Applications in Risk Assessment, Dr. Yuzhou Luo (Ed.), InTech. doi:10.5772/ 38811.

Peng, R.D. and Dominici, F. (2008). Statistical methods for environmental epidemiology with R. R: *A Case Study in Air Pollution and Health (Springer)*. doi:10.1007/978-0-387-78167-9

Bruneau, B. and Grégoire, F. (2011). Étude de la distribution spatiale des données d'abondance de maquereau bleu (Scomber scombrus) et de capelan (Mallotus villosus) des relevés d'hiver aux poissons de fond des Divisions 4VW de l'OPANO à l'aide de modèles additifs généralisés. Rapport technique canadien des sciences halieutiques et aquatiques.

Team, R.C., (2014). R: A language and environment for statistical computing. Vienna: *R Foundation for Statistical Computing*. (ISBN 3-900051-07-0).

TWood, S. (2006). Generalized additive models: an introduction with R. Boca Raton, FL: CRC press.

Oshinubi K, Al-Awadhi F, Rachdi M, and Demongeot J (2021). Data analysis and forecasting of COVID-19 pandemic in Kuwait. *Kuwaitian Journal of Science (available on medRxiv)*. Doi: https://doi.org/10.48129/KJS.SPLCOV.14501

Demongeot J, Oshinubi K, Rachdi M, Seligman H, Thuderoz F and Waku J (2021). Estimation of daily reproduction numbers in Covid-19 outbreak. *Computation*, Vol. **9**, Issue 109. Doi:https://doi.org/10.3390/computation9100109.

Sallahi N, Park H, El Mellouhi F, Rachdi M, Ouassou I, Belhaouari S, Arredouani A and Bensmail H (2021). Using unstated cases to correct for COVID-19 pandemic outbreak and its impact on easing the intervention for Qatar. *Biology*, Vol. **10**, Issue 6, pp 463. https://doi.org/10.3390/biology10060463 - 24 May 2021.

Oshinubi K, Ibrahim F, Rachdi M, and Demongeot J (2022). Functional Data Analysis: Application to daily observation of COVID-10 prevalence in France. *AIMS Mathematics*, Vol. **7**, Issue 4, pp: 5347-5385. Doi: 10.3934/math.2022298

Demongeot J, Oshinubi K, Rachdi M, Hobbad L, Alahiane M, Iggui S, Gaudart J and Ouassou I (2022) The application of ARIMA model to analyze Covid-19 incidence pattern in several countries. *Journal of Mathematics and Computer Sciences (JMCS)*, Doi: https://doi.org/10.28919/jmcs/6541.