Georgia Southern University

Digital Commons@Georgia Southern

Department of Biostatistics, Epidemiology, and Environmental Health Sciences Faculty Publications

Department of Biostatistics, Epidemiology, and Environmental Health Sciences

3-9-2021

Lag Effects of Ozone, PM2.5, and Meteorological Factors on COVID-19 New Cases at the Disease Epicenter in Queens, New York

Atin Adhikari

Jingjing Yin

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/bee-facpubs

Part of the Biostatistics Commons, and the Epidemiology Commons

This article is brought to you for free and open access by the Department of Biostatistics, Epidemiology, and Environmental Health Sciences at Digital Commons@Georgia Southern. It has been accepted for inclusion in Department of Biostatistics, Epidemiology, and Environmental Health Sciences Faculty Publications by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

atmosphere

Article

Lag Effects of Ozone, $PM_{2.5}$, and Meteorological Factors on COVID-19 New Cases at the Disease Epicenter in Queens, New York

Atin Adhikari *,† and Jingjing Yin †

Department of Biostatistics, Epidemiology & Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA 30460, USA; jyin@georgiasouthern.edu
* Correspondence: aadhikari@georgiasouthern.edu
† Adhikari and Yin are co-lead authors and equally contributed to this work.

Abstract: The influences of environmental factors on COVID-19 may not be immediate and could be lagged for days to weeks. This study investigated the choice of lag days for calculating cumulative lag effects of ozone, $PM_{2.5}$, and five meteorological factors (wind speed, temperature, relative humidity, absolute humidity, and cloud percentages) on COVID-19 new cases at the epicenter of Queens County, New York, before the governor's executive order on wearing of masks in public places (1 March to 11 April 2020). Daily data for selected air pollutants and meteorological factors were collected from the US EPA Air Quality System, weather observation station of the NOAA National Centers for Environmental Information at John F. Kennedy Airport, and World Weather Online. Negative binomial regression models were applied, including the autocorrelations and trend of the time series, as well as the effective reproductive number as confounders. The effects of ozone, $PM_{2.5}$, and five meteorological factors were significant on COVID-19 new cases with lag9-lag13 days. Incidence rate ratios (IRRs) were consistent for any lag day choice between lag0 and lag14 days and started fluctuating after lag15 days. Considering moving averages >14 days yielded less reliable variables for summarizing the cumulative lag effects of environmental factors on COVID-19 new cases and considering lag days from 9 to 13 would yield significant findings. Future studies should consider this approach of lag day checks concerning the modeling of COVID-19 progression in relation to meteorological factors and ambient air pollutants.

Keywords: COVID-19; SARS-CoV-2; temperature; humidity; air pollution; lag effects

check for updates

Citation: Adhikari, A.; Yin, J. Lag Effects of Ozone, $PM_{2.5}$, and Meteorological Factors on COVID-19 New Cases at the Disease Epicenter in Queens, New York. *Atmosphere* **2021**, *12*, 357. https://doi.org/10.3390/atmos12030357

Received: 16 February 2021
Accepted: 5 March 2021
Published: 9 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Previous studies of Hemmes et al. [1] showed that the transmission of coronaviruses could be affected by climatic factors. The significant effects of weather conditions and several air pollutants on the incidence of COVID-19 have recently been reported from many countries worldwide [2–6]. In our recent study, we also found that the daily max 8 h ozone concentrations, average temperature between 32° and 55° F, relative humidity between 41% and 92%, and cloud percentages were significantly associated with new COVID-19 cases in Queens, New York [7].

Some researchers and the common public anticipated that the arrival of summer would reduce COVID-19 cases with increasing temperature because previous laboratory studies showed that lower temperatures contributed to the high robustness and strong capability to the survival of MERS coronaviruses [8]. A recent laboratory study by Chen et al. [9] examined the stability of SARS-CoV-2 as a function of temperature and other conditions in the lab. They started with a SARS-CoV-2 suspension of 6.7 log $TCID_{50}$ (median Tissue Culture Infection Dose)/mL in a virus transport medium and found that the levels of viruses reduced only a 0.6-log unit at 14 days at 4 °C. In contrast, higher inactivation rates were observed at 22 °C (a 3-log unit reduction at seven days and no detection at 14 days) and 37 °C (a 3-log unit reduction after one day and no viruses detected subsequently). Another new study by Tobías and Molina [5] showed that the number of diagnosed COVID-19 cases was decreasing with the increase in temperature between 12.2 °C and 22.8 °C. The study of Qi et al. [10] from China suggested that every 1 °C increase in ambient temperature led to a decrease in daily confirmed COVID-19 cases by 36–57% when relative humidity was between 67% and 85.5%. However, when the relative humidity increased by 1%, the daily confirmed cases decreased by 11–22% at the average temperature range of 5.04–8.2 °C. A large-scale study (non-peer-reviewed preprint) investigated daily COVID-19 case growth rates in 121 countries or regions and found that the growth rates peaked in countries or regions with a mean temperature of 5 °C and decreased in warmer and colder climates [11]. Another non-peer-reviewed study based on the data collected from 310 geographic regions across 116 countries also reported an inverse relationship between temperature and humidity and COVID-19 cases [12].

Based on these reports, anyone can expect that COVID-19 cases would reduce in summer; however, these anticipations appeared to be unbecoming, and COVID-19 incidence had been rapidly increasing in most US states even in the hot summer days of July when the temperature was above 37 °C in many places. This observation matches the initial phase of the COVID-19 pandemic in countries such as Iran and Australia, which had warmer climatic conditions, but still, experienced a rapid spread of the disease. Many reasons can be held responsible for these discrepancies. First, lab experimental conditions may not always represent conditions the viruses are encountering in real-life environmental conditions. Viruses released from naturally infected humans may have different survival properties in the surrounding indoor and outdoor environments compared to the growth of the virus in tissue culture media or isolated body fluids tested in the lab. Second, previous studies were conducted often at the country scale or multi-city scale, which might have some measurement error due to high spatial variability in temperature [13], other meteorological factors, and air pollutant levels. Some previous studies mapped the spread of the virus across the world and compared its progress with local climate variables, which is not reasonable. Most studies considered long-term averages of these data for large areas, and therefore, modeling based on these inaccurate data became problematic.

Finally, the lack of reliable information on the lag effects of various meteorological factors and air pollutants on COVID-19 cases may also result in unreliable findings. The consideration of inconsistent lag effects for different environmental factors on COVID-19 cases and deaths appeared to be a common methodological challenge in the recently published articles. Some researchers considered data of 20 to 30 days before the first death [4], some considered 0 to 5 days single and multiple averages [3], and others considered data of the same day or lag-1 to 6 days [5], and 2–14 days [2], or moving averages (MA) of 0–7, 0–14, and 0–21 days [6,7].

Most authors considered the incubation period of COVID-19 as lag days needed, some followed previous reports, and to our knowledge, there is no consensus and no scientific data available yet, which can be utilized for this vital consideration during COVID-19 modeling. We have addressed this knowledge gap in this article and proposed to evaluate the modeling effects across different choices of lag days; thus, the optimal choice of lag days is determined from statistical modeling rather than taking a subjective guess as performed in most other studies. In the investigation, we assumed that all meteorological factors and air pollutants have the same period of lagged effects, and we wanted to determine the proper value of lagged days. In addition, we also wanted to test two hypotheses: (1) The cumulative lag effects of all environmental factors are significantly influencing COVID-19 incidences; and (2) the lag effects considered by previous studies based on the multi-city, multi-state, and multi-country consideration of meteorological variables and air pollutants could be different if these effects are examined in a single location. The rationales for the second hypothesis are (1) more reliable measurements of the environmental factors in a specific location; (2) problems related to data quality and the limitation in times and locations; (3) potential errors generated by combining data from different observational windows, and (4) many other confounding factors related to geographic variations.

2. Materials and Methods

2.1. The Study Area and Data Sources

The study area and sampling sites in Queens county, New York, were described in detail in our recently published article [7]. Queens is the largest borough among the five boroughs of New York City, with an estimated population of 2,253,858 residents and one of the most COVID-19 affected areas in the US. We have collected daily maximum 8-h ozone, average daily $PM_{2.5}$, average temperature, wind speed, relative humidity, and cloud percentages data from the databases of the monitoring stations at Queens College (US EPA Air Quality System; https://www.epa.gov/aqs, accessed on 28 September 2020; Lat: 40.7392; Long: −73.8177), weather observation station of the NOAA National Centers for Environmental Information at John F. Kennedy International Airport, and World Weather Online, reported for the nearby Meadowmere Park area in Queens (Lat: 40.6362; Long: −73.7415). The study period was set from 1 March to 11 April 2020 since it was the period before imposing a stricter measure to control the spread of the coronavirus by the New York governor who ordered all New Yorkers must wear face coverings when social distancing is not possible, including on public transport, in stores, and on crowded sidewalks effective from 17 April (*New York Times*, 15 April 2020). We have used the average of the two $PM_{2.5}$ data sets since $PM_{2.5}$ data were available from two sampling locations at the same site. Absolute humidity was calculated using relative humidity, temperature, and pressure data. Absolute humidity is the measure of water vapor (moisture) in the air, regardless of temperature. Whereas relative humidity also measures water vapor but relative to the temperature of the air. So, the analyses are not redundant. Ma et al. [3] previously analyzed exposure–response curves of meteorological factors and COVID-19 daily mortality counts in Wuhan, China, and found different types of curves for relative and absolute humidities. That is why we decided to use both in our analyses.

Data on COVID-19 new confirmed cases for Queens county were collected from USAFacts (https://usafacts.org/ accessed on 28 September 2020), which is a not-for-profit, nonpartisan civic initiative providing the most comprehensive and understandable source of government data available in the US. We have collected the data on R_t of SARS-CoV-2 for the New York region from https://rt.live/, accessed on 28 September 2020 (model by Kevin Systrom and Thomas Vladeck). The R_t or the effective reproductive number is a key measure of how fast the virus is growing. It is the average number of people who become infected by an infectious person. If Rt is above 1.0, the virus will spread quickly. When Rt is below 1.0, the virus will stop spreading. Regarding the calculation of R_t, the researchers involved in rt.live stated that they had assumed a seed number of people and a curve of R_t over the history of the pandemic and then distributed those cases into the future using a known delay distribution between infection and positive report. Then they scaled this and added noise based on known testing volumes via a negative binomial with an exposure parameter for a given day to recover an observed series (from https://rt.live/faq, accessed on 28 September 2020). We have considered R_t as a confounding factor in this study because R_t could be influenced by different intervention policies, including social distancing advice from the government, human behavioral patterns and movements, and the availability of diagnostics for COVID-19 testing.

Missing data in NOAA databases for the JFK airport sampling station were replaced by data from World Weather Online because their sampling station was located a few miles away from the JFK airport. For the remaining missing data, we used different statistical approaches for different situations of missing. Ozone values had five missing observations, we used the average of values on the day before and the day after to replace for the three single-day missing observations at day t as $o(t) = \text{mean}(t - 1, t + 1)$; for the other two consecutive missing ozone observations on t and t + 1 day, we used the approaches of Last Observation Carry Forward (LOCF) as $o(t) = o(t - 1)$ and Next Observation

Carried Backward (NOCB) as o(t + 1) = o(t + 2) [14,15]. $PM_{2.5}$ data had seven consecutive observations missing, and we imputed these missing values by the fitted time series with seasonal adjustment and linear interpolation [16]. Similarly, as Rt data has been available since 24 February 2020, to calculate various time frame of moving averages (MA) up to 21 days, these data were used for the follow-up regression analysis from 1 March to 11 April 2020, and we needed to impute missing Rt values between 9 February to 23 February 2020. We applied time series imputation with spline interpolation [17] and without seasonal adjustment to replace the consecutive missing Rt observations with the fitted time series data. Overall, the data used in this study were collected mostly from governmental agencies, which used reliable air sampling and environmental monitoring methods.

2.2. Data Analysis

Negative binomial regression models were applied for modeling the effects of two air pollutants ($PM_{2.5}$ and ozone) and five meteorological factors (wind speed, temperature, cloud percentage, relative humidity, and absolute humidity) on new daily COVID-19 cases from 1 March to 11 April 2020. Many existing COVID-19 studies [18,19] used log-linear models, which cannot fit zero values as it is not valid to take logarithms on zero. A more appropriate family of statistical models called generalized linear models (GLM) are available to fit the count data with zeros, and the Poisson model is the most well-known GLM model and has been adopted by some researchers for the COVID-19 studies [20]. In our study, we applied another less common member of the GLM family, the Negative Binomial model, because it fits over dispersed count data (i.e., the variance of outcome is larger than its mean, which is usually the case for skewed count outcomes) better than the Poisson model. For details and formulation of the Negative Binomial model, please refer to the article by Greene (2008) [21].

The Negative Binomial regression models the logarithm of the mean of COVID-19 case counts at day t as a linear function of the predictor of interest. The regression models were fitted for each predictor separately due to high correlations between predictors. The model equation is:

$$\log(E(Y_t)) = \beta_0 + \beta_1 MA_k(X_t) + \beta_2 \log(Y_{t-1} + 1) + \beta_3 \log(Y_{t-2} + 1) + \beta_4 t + \beta_5 MA_k(R_t),$$

where the primary variable of interest is the cumulative lag effect of predictor X at day t, which is denoted as variable $MA_k(X_t)$. This cumulative lag effect is calculated by the moving average of lag-0 to lag-k days for the given environmental predictor X on day t, which is $MA_k(X_t) = \text{average}(X_{t-k}, X_{t-k+1}, \ldots, X_t)$ and k is set as an integer value between 0 and 21. The primary focus of this research was to evaluate the impact of the choice of lag days (k) in the calculation of the cumulative lag effect and investigate whether a practical common choice of k exists across different predictors. We chose 21 days as the maximum lag days for calculating the cumulative lag effect since the incubation period of COVID-19 under conservative assumptions is about 14 days [22], and the worst-case maximum incubation period for COVID-19 can be 19 and 27 days based on available literature [23,24]. In addition, we wanted to capture mostly the lag effect by setting the anticipated effect time one week earlier than the 14-day incubation time for typical COVID-19 cases. By setting the maximum lag day (k) value as 21, which is a very conservative estimate of k, we believe we have evaluated the entire range of all possible values of k in the study where k = [0, 21]. In addition, we adjusted the single-predictor models by four confounders, i.e., lagged outcomes on the previous two days, trend, and the R_t values. The lagged outcomes were considered to account for the potential autocorrelation of the time series of new cases on the $t-1$ and $t-2$ days (i.e., $\log(Y_{t-1} + 1)$ and $\log(Y_{t-2} + 1)$). We added 1 before taking the log to avoid the situation of log (0) because, at the earlier month of March 2020, there were days with zero new cases. The trend was considered to control for the linear trend of the time series at each day (i.e., at time t) due to other unobserved factors. R_t values (the effective reproductive number) were considered because R_t has been found to be declining with interventions in different countries, including lockdowns, school closures, etc. And we used the same approach of calculating the MA of lag-0 to lag-k values as we did for the environmental factors for the cumulative lag effects of R_t for the same reason of the incubation period of the disease. Furthermore, we checked if there was any potential seasonal pattern of new cases that we may control for the seasonality of time series in the regression models. For example, there could be a weekly pattern due to varying testing availability and case registrations between days during the week and on weekends. However, our data did not show such a pattern, which indicated that the disease progression and management of disease testing by health departments were incessant on the weekends. Effect estimates were calculated as the exponential form of the regression coefficients (i.e., $\exp(\beta_1)$ for each of the environmental factors), which reported as the incidence rate ratios (IRRs), along with the corresponding *p*-values to show statistical significance.

All analyses in this study were conducted using R statistical software (R Foundation for Statistical Computing, Vienna, Austria). A *p*-value of <0.05 was considered to be statistically significant. We used the "MASS" R—package [25] to fit the negative binomial regression models and the R-package "imputeTS" [26] for missing data imputation of $PM_{2.5}$ and R_t data.

3. Results

Changes of IRRs and *p*-values in relation to lag-day values are presented in Figure 1 and Table 1. We found that none of the meteorological factors and the two selected air pollutants had significant effects on COVID-19 until day 4 ($p > 0.05$), and relative humidity and absolute humidity started showing significance from day 5 ($p < 0.001$). Interestingly, all effects were significant with lag9-lag13 (green area in Figure 1) days. IRR values were more stable and showed small changes (close to 1) with lag0-lag14 days and started to fluctuate after the lag15 day. We found that ozone, wind speed, and temperature were positively associated with COVID-19 new cases, while PM2.5, cloud, relative humidity, and absolute humidity were negatively associated with new COVID-19 cases in most models.

Table 1. Actual incidence rate ratios (IRRs) and *p* numerical values for lags of days from 0 to 21 days.

	$PM_{2.5}$ (μg/m^3)		Ozone (ppb)		Wind Speed (m/s)		Temperature (°F)		Cloud (%)		Relative Humidity (%)		Absolute Humidity (g/cm^3)	
Lag days	**IRR**	***p***	**IRR**	***p***	**IRR**	***p***	**IRR**	***p***	**IRR**	***p***	**IRR**	***p***	**IRR**	***p***
0	1.00	0.976	0.99	0.407	1.01	0.365	1.01	0.540	1.00	0.952	1.00	0.962	1.01	0.739
1	1.03	0.509	0.97	0.300	1.01	0.602	1.01	0.685	1.00	0.642	1.00	0.953	1.02	0.657
2	1.05	0.433	0.95	0.109	1.00	0.914	0.99	0.720	1.00	0.712	1.01	0.356	1.03	0.704
3	1.11	0.138	0.96	0.384	0.98	0.530	1.00	0.991	1.00	0.791	0.99	0.571	0.99	0.919
4	1.09	0.319	0.95	0.461	1.02	0.000	0.99	0.003	0.99	0.143	0.96	0.025	0.83	0.000
5	1.01	0.128	1.01	0.034	1.03	0.000	1.04	0.000	0.99	0.000	0.93	0.000	0.91	0.000
6	1.04	0.000	1.00	0.562	1.01	0.217	1.05	0.000	1.00	0.000	0.98	0.000	0.99	0.435
7	1.03	0.035	1.01	0.265	1.00	0.812	1.12	0.000	0.99	0.000	0.97	0.000	0.85	0.000
8	1.08	0.000	0.93	0.000	1.02	0.055	1.11	0.000	0.99	0.000	0.98	0.000	0.95	0.151
9	0.95	0.002	0.94	0.000	1.08	0.000	1.06	0.000	0.99	0.000	0.96	0.000	0.75	0.000
10	0.95	0.006	1.14	0.000	1.05	0.000	1.09	0.000	0.98	0.000	0.94	0.000	0.80	0.000
11	0.86	0.000	1.06	0.000	1.07	0.000	1.08	0.000	0.99	0.000	0.96	0.000	0.88	0.000
12	0.81	0.000	1.11	0.000	1.11	0.000	1.07	0.000	0.98	0.000	0.94	0.000	0.83	0.000
13	0.80	0.000	1.09	0.000	1.03	0.000	1.06	0.000	0.99	0.000	0.96	0.000	0.83	0.000
14	0.74	0.000	0.94	0.001	1.00	0.722	1.09	0.000	0.99	0.005	0.97	0.000	0.82	0.000
15	0.70	0.000	1.11	0.000	1.00	0.842	1.24	0.000	1.02	0.000	1.00	0.547	1.83	0.000
16	1.28	0.000	1.16	0.000	1.00	0.766	1.31	0.000	0.98	0.000	0.94	0.000	0.90	0.097
17	0.76	0.000	1.83	0.004	1.07	0.000	1.28	0.000	0.97	0.000	0.93	0.000	0.76	0.000
18	0.65	0.192	1.99	0.001	1.17	0.000	0.96	0.003	0.99	0.000	0.98	0.000	0.63	0.000
19	0.73	0.318	1.98	0.003	1.06	0.009	0.97	0.014	0.99	0.008	0.97	0.000	0.65	0.000
20	0.52	0.032	1.12	0.000	1.07	0.000	1.06	0.000	0.98	0.000	0.95	0.000	0.75	0.000
21	0.37	0.001	2.24	0.001	1.26	0.000	1.16	0.000	0.98	0.000	0.97	0.000	0.86	0.001

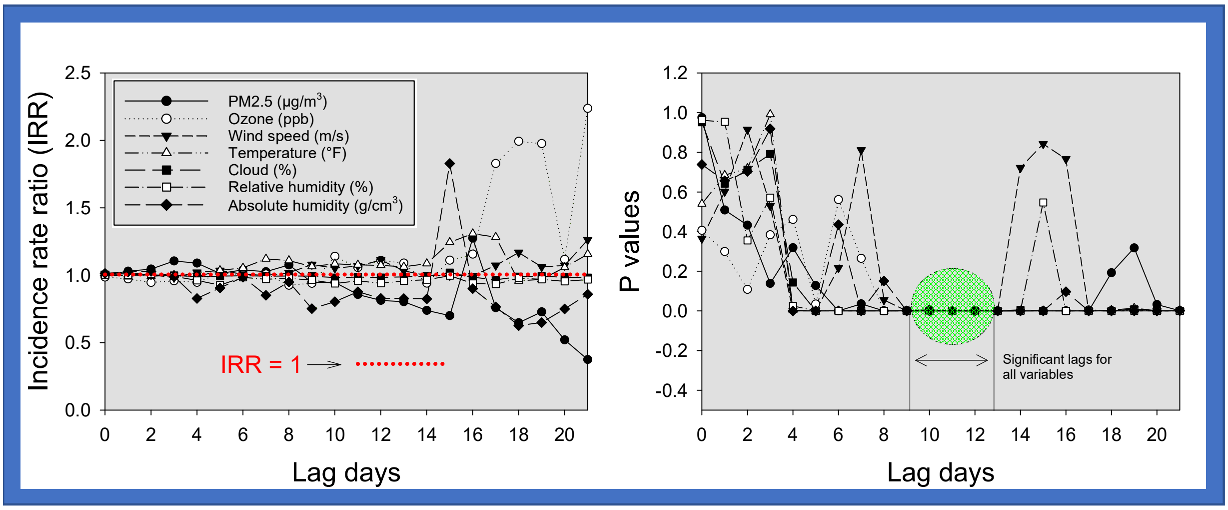

Figure 1. Changes in incidence rate ratios (IRRs) (**left**) and *p*-values (**right**) in relation to lags of days.

Autocorrelation with the previous two days and the trend effect of the outcome time series were found to be consistently positively associated with new COVID-19 daily new cases, which indicates that during the time frame of 3 March to 11 April 2020, the new cases generally progressed as an increasing trajectory. The cumulative lag effect of R_t during previous k days was found to be negative in most models; however, it does not necessarily mean R_t was negatively associated with COVID-19 new cases. It is notable that R_t had a strong decreasing trend over the study period, while the COVID-19 cases had an increasing trend. The time series trend and autocorrelation effects were more significant with *p*-values < 0.0001 and thus were dominant for the change in the COVID-19 new cases. Therefore, we concluded the change in cases was influenced primarily by the time series itself, not R_t or environmental factors. This finding is consistent with another non-peer-reviewed article by Adeyemi et al. [27].

4. Discussion

This study was conducted to assess the structure of the lag effects of two major air pollutants and a few meteorological factors on COVID-19 new cases for the first time in the disease epicenter involving the New York area. Although there are some reports on lag effects from other countries, a detailed study underscoring the New York area is important because the attack was heavier there, and climatic conditions were significantly different from disease epicenters in other countries.

We found that all of the selected variables presented a short-time lagged effect, which was at least four days for relative and absolute humidity levels. Overall, we concluded that selected meteorological factors and pollutants all have significantly influenced COVID-19 incidence when MA values of day t to day t-k where k = [9, 13] are considered, and the effects become less reliable if k > 14 days. Thus, significant cumulative lagged effects of meteorological factors and air pollutants were observed, as we anticipated in our first hypothesis. Concurring our second hypothesis, we found that the choice of lags for calculating the MA values are different from the lags considered by previous researchers [2–7]. This new finding generated by the present study will be vital in modeling COVID-19 concerning environmental factors, particularly in places with a similar condition, such as New York. It is notable that choosing the moving average of more than two weeks (i.e., lag day k > 14) would yield a less reliable variable for the cumulative lag effects of environmental factors on COVID-19 new cases. Additionally, we found lag day k chosen from 9 to 13 would yield significant findings. Therefore, we recommend using the moving average of daily environmental measurements during the past one to two weeks to capture the cumulative lag effect for future studies to be conducted in other global locations. However, different data from various sources and with different study designs are likely to have different optimal choice for the lag day k, and therefore, to find reliable estimates, we suggest future studies consider performing lag day checks similar to our study for the modeling of COVID-19 progression concerning selected meteorological factors and air pollutants.

Our estimates are reliable because of official data sources: used data were primarily reported by US governmental agencies [7], and the controlled study design that limits the source of unobserved confounding factors: a) single location. Instead of combining data from many cities with different geographic and demographic variations we have used data from a county of New York, which is a US COVID-19 epicenter; and (b) controlled time frame. Our analysis is limited to the first six weeks of COVID-19 before the executive order issuance on wearing masks in public places. Another unavoidable limitation is the uncertainty of exposure levels. We have used the data collected by stationary monitors. These data may not represent actual personal level exposures for the infected people and those residing away from the monitoring stations.

This study also showed that selected air pollutants and meteorological factors might not have a similar type of effect on COVID-19, and both positive and negative effects should be considered in the disease progression modeling. Therefore, future studies should be designed where the interaction effects of meteorological factors and air pollutants on this pandemic should be investigated. In fact, the interaction effects of air pollutants and meteorological factors were previously identified in noninfectious diseases [28,29] and hospital emergency department visits [9]. Exploring these interactions might resolve inconsistencies between reports on the impacts of meteorological factors and air pollutants conducted in different countries and regions of the world.

The associations between ambient air pollutants and respiratory infections, including COVID-19, could be influenced differently by the times of exposure and exposed populations. The underlying cause of the negative correlation for $PM_{2.5}$ was probably due to the lower transportation of vehicles in New York in the early phase of COVID-19 when case numbers were quickly increasing. Because COVID-19 is a new pandemic and pathology is not adequately explored yet with respect to different air pollutant exposures, it is difficult to provide a biologically and clinically feasible explanation at this point for the observed lags of various environmental factors on COVID-19. Synergistic and antagonistic interactions between different air pollutants in causing COVID-19 require further investigations in controlled laboratory settings. Preexisting respiratory health of the exposed population could be an important determining factor as well. However, we found some reports where respiratory diseases, such as acute and chronic bronchitis, were significantly influenced by lagged ozone exposures. Acute bronchitis was influenced by same day (lag 0), lag 2, lag 3–7, 9 days of ozone exposures, whereas chronic bronchitis was influenced by lagged ozone exposures of 6, 7, and 9 days [30]. For COVID-19, lagged exposure effects were reported by a few earlier studies, and we have already referred to them in the Introduction [2–7].

5. Conclusions

The findings from this study conclude that the effects of ozone, $PM_{2.5}$, and five meteorological factors in Queens, New York were significant on COVID-19 new cases with lag9-lag13 days, which were supported by the consistent IRRs for any lag day choice between lag0-lag14 days and fluctuations observed after lag15 days. We also conclude that consideration of moving averages > 14 days would yield less reliable variables for summarizing the cumulative lag effects of environmental factors on COVID-19 new cases, and consideration of lag days from 9 to 13 is recommended to obtain consistent and thus more reliable findings.

Author Contributions: Conceptualization, A.A.; methodology, J.Y. and A.A.; formal analysis, J.Y.; investigation, A.A. and J.Y.; writing—original draft preparation, A.A.; writing—review and editing, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Ethical Statement: The data on the daily new COVID-19 cases, air pollutants, and meteorological parameters were obtained from the publicly available databases. Therefore, ethical approval was not required.

Data Availability Statement: The data used can be obtained from the databases of USAFacts, the US EPA Air Quality System, the weather observation station of the NOAA National Centers for Environmental Information at the John F. Kennedy International Airport, and World Weather Online.

Acknowledgments: We are grateful to USAFacts, the US EPA Air Quality System, the weather observation station of the NOAA National Centers for Environmental Information at the John F. Kennedy International Airport, and World Weather Online for providing access to their collected data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hemmes, J.H.; Winkler, K.C.; Kool, S.M. Virus survival as a seasonal factor in influenza and poliomyelitis. *Antonie Leeuwenhoek* **1962**, *28*, 221–233. [CrossRef]
2. Briz-Redón, Á.; Serrano-Aroca, Á. A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. *Sci. Total Environ.* **2020**, *728*, 138811. [CrossRef] [PubMed]
3. Ma, Y.; Zhao, Y.; Liu, J.; He, X.; Wang, B.; Fu, S.; Yan, J.; Niu, J.; Zhou, J.; Luo, B. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci. Total Environ.* **2020**, *724*, 138226. [CrossRef] [PubMed]
4. Sajadi, M.M.; Habibzadeh, P.; Vintzileos, A.; Shokouhi, S.; Miralles-Wilhelm, F.; Amoroso, A. Temperature, Humidity, and Latitude Analysis to Estimate Potential Spread and Seasonality of Coronavirus Disease 2019 (COVID-19). *JAMA Netw. Open* **2020**, *3*, e2011834. [CrossRef]
5. Tobías, A.; Molina, T. Is temperature reducing the transmission of COVID-19? *Environ. Res.* **2020**, *186*, 109553. [CrossRef] [PubMed]
6. Xie, J.; Zhu, Y. Association between ambient temperature and COVID-19 infection in 122 cities from China. *Sci. Total Environ.* **2020**, *724*, 138201. [CrossRef]
7. Adhikari, A.; Yin, J. Short-Term Effects of Ambient Ozone, PM2.5, and Meteorological Factors on COVID-19 Confirmed Cases and Deaths in Queens, New York. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4047. [CrossRef]
8. Pyankov, O.V.; Bodnev, S.A.; Pyankova, O.G.; Agranovski, I.E. Survival of aerosolized coronavirus in the ambient air. *J. Aerosol Sci.* **2018**, *115*, 158–163. [CrossRef]
9. Chen, Y.; Zheng, M.; Lv, J.; Shi, T.; Liu, P.; Wu, Y.; Feng, W.; He, W.; Guo, P. Interactions between ambient air pollutants and temperature on emergency department visits: Analysis of varying-coefficient model in Guangzhou, China. *Sci. Total Environ.* **2019**, *668*, 825–834. [CrossRef]
10. Qi, H.; Xiao, S.; Shi, R.; Ward, M.P.; Chen, Y.; Tu, W.; Su, Q.; Wang, W.; Wang, X.; Zhang, Z. COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. *Sci. Total Environ.* **2020**, *728*, 138778. [CrossRef] [PubMed]
11. Ficetola, G.F.; Rubolini, D. Climate affects global patterns of CoViD-19 early outbreak dynamics. *medRxiv* **2020**. [CrossRef]
12. Islam, N.; Shabnam, S.; Erzurumluoglu, A.M. Temperature, humidity, and wind speed are associated with lower Covid-19 incidence. *medRxiv* **2020**. [CrossRef]
13. Soares, P.M.M.; Cardoso, R.M.; Miranda, P.M.A.; De Medeiros, J.; Belo-Pereira, M.; Espirito-Santo, F. WRF high resolution dynamical downscaling of ERA-Interim for Portugal. *Clim. Dyn.* **2012**, *39*, 2497–2522. [CrossRef]
14. Engels, J.M.; Diehr, P. Imputation of missing longitudinal data: A comparison of methods. *J. Clin. Epidemiol.* **2003**, *56*, 968–976. [CrossRef]
15. Shao, J.; Zhong, B. Last observation carry-forward and last observation analysis. *Stat. Med.* **2003**, *22*, 2429–2441. [CrossRef] [PubMed]
16. Noor, N.M.; Al Bakri Abdullah, M.M.; Yahaya, A.S.; Ramli, N.A. Comparison of Linear Interpolation Method and Mean Method to Replace the Missing Values in Environmental Data Set. In *Materials Science Forum*; Trans Tech Publications Ltd.: Bäch, Switzerland, 2015; Volume 803, pp. 278–281.
17. Zhang, Z. Missing data imputation: Focusing on single imputation. *Ann. Transl. Med.* **2016**, *4*, 9.
18. Zhu, Y.; Xie, J.; Huang, F.; Cao, L. Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. *Sci. Total Environ.* **2020**, *727*, 138704. [CrossRef]
19. Schuit, M.; Ratnesar-Shumate, S.; Yolitz, J.; Williams, G.; Weaver, W.; Green, B.; Miller, D.; Krause, M.; Beck, K.; Wood, S.; et al. Airborne SARS-CoV-2 Is Rapidly Inactivated by Simulated Sunlight. *J. Infect. Dis.* **2020**, *222*, 564–571. [CrossRef] [PubMed]
20. Jiang, Y.; Wu, X.-J.; Guan, Y.-J. Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infect. Control. Hosp. Epidemiol.* **2020**, *41*, 1011–1015. [CrossRef] [PubMed]
21. Greene, W. Functional forms for the negative binomial model for count data. *Econ. Lett.* **2008**, *99*, 585–590. [CrossRef]

22. Lauer, S.A.; Grantz, K.H.; Bi, Q.; Jones, F.K.; Zheng, Q.; Meredith, H.R.; Azman, A.S.; Reich, N.G.; Lessler, J. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* **2020**, *172*, 577–582. [CrossRef]
23. Bai, Y.; Yao, L.; Wei, T.; Tian, F.; Jin, D.Y.; Chen, L.; Wang, M. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA* **2020**, *323*, 1406–1407. [CrossRef]
24. Reuters. Coronavirus Incubation Could Be as Long as 27 Days, Chinese Provincial Government Says. Reuters. 2020. Available online: https://www.reuters.com/article/us-china-health-incubation/coronavirus-incubation-could-be-as-long-as-27-days-chinese-provincial-government-says-idUSKCN20G06W (accessed on 21 July 2020).
25. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*, 4th ed.; Springer: New York, NY, USA, 2020.
26. Moritz, S.; Bartz-Beielstein, T. imputeTS: Time Series Missing Value Imputation in R. *R J.* **2017**, *9*, 207–218. [CrossRef]
27. Adeyemi, S.; Yakutcan, U.; Demir, E. A statistical Assessment of Association between Meteorological Parameters and COVID-19 Pandemic in 10 Countries: The Devil Is in the Detail. Available online: https://www.researchgate.net/publication/342313160 (accessed on 21 July 2020).
28. Huang, C.-H.; Lin, H.-C.; Tsai, C.-D.; Huang, H.-K.; Lian, I.-B.; Chang, C.-C. The Interaction Effects of Meteorological Factors and Air Pollution on the Development of Acute Coronary Syndrome. *Sci. Rep.* **2017**, *7*, srep44004. [CrossRef] [PubMed]
29. Guo, Q.; Xiong, X.; Liang, F.; Tian, L.; Liu, W.; Wang, Z.; Pan, X. The interactive effects between air pollution and meteorological factors on the hospital outpatient visits for atopic dermatitis in Beijing, China: A time-series analysis. *J. Eur. Acad. Dermatol. Venereol.* **2019**, *33*, 2362–2370. [CrossRef] [PubMed]
30. Kousha, T.; Rowe, B.H. Ambient ozone and emergency department visits due to lower respiratory condition. *Int. J. Occup. Med. Environ. Health* **2014**, *27*, 50–59. [CrossRef]