

Original Paper

Short-Range Forecasting of COVID-19 During Early Onset at County, Health District, and State Geographic Levels Using Seven Methods: Comparative Forecasting Study

Christopher J Lynch, PhD; Ross Gore, PhD

Virginia Modeling, Analysis, and Simulation Center, Old Dominion University, Suffolk, VA, United States

Corresponding Author:

Christopher J Lynch, PhD

Virginia Modeling, Analysis, and Simulation Center

Old Dominion University

1030 University Blvd

Suffolk, VA, 23435

United States

Phone: 1 7576866248

Email: cjlynch@odu.edu

Abstract

Background: Forecasting methods rely on trends and averages of prior observations to forecast COVID-19 case counts. COVID-19 forecasts have received much media attention, and numerous platforms have been created to inform the public. However, forecasting effectiveness varies by geographic scope and is affected by changing assumptions in behaviors and preventative measures in response to the pandemic. Due to time requirements for developing a COVID-19 vaccine, evidence is needed to inform short-term forecasting method selection at county, health district, and state levels.

Objective: COVID-19 forecasts keep the public informed and contribute to public policy. As such, proper understanding of forecasting purposes and outcomes is needed to advance knowledge of health statistics for policy makers and the public. Using publicly available real-time data provided online, we aimed to evaluate the performance of seven forecasting methods utilized to forecast cumulative COVID-19 case counts. Forecasts were evaluated based on how well they forecast 1, 3, and 7 days forward when utilizing 1-, 3-, 7-, or all prior-day cumulative case counts during early virus onset. This study provides an objective evaluation of the forecasting methods to identify forecasting model assumptions that contribute to lower error in forecasting COVID-19 cumulative case growth. This information benefits professionals, decision makers, and the public relying on the data provided by short-term case count estimates at varied geographic levels.

Methods: We created 1-, 3-, and 7-day forecasts at the county, health district, and state levels using (1) a naïve approach, (2) Holt-Winters (HW) exponential smoothing, (3) a growth rate approach, (4) a moving average (MA) approach, (5) an autoregressive (AR) approach, (6) an autoregressive moving average (ARMA) approach, and (7) an autoregressive integrated moving average (ARIMA) approach. Forecasts relied on Virginia's 3464 historical county-level cumulative case counts from March 7 to April 22, 2020, as reported by *The New York Times*. Statistically significant results were identified using 95% CIs of median absolute error (MdAE) and median absolute percentage error (MdAPE) metrics of the resulting 216,698 forecasts.

Results: The next-day MA forecast with 3-day look-back length obtained the lowest MdAE (median 0.67, 95% CI 0.49-0.84, $P < .001$) and statistically significantly differed from 39 out of 59 alternatives (66%) to 53 out of 59 alternatives (90%) at each geographic level at a significance level of .01. For short-range forecasting, methods assuming stationary means of prior days' counts outperformed methods with assumptions of weak stationarity or nonstationarity means. MdAPE results revealed statistically significant differences across geographic levels.

Conclusions: For short-range COVID-19 cumulative case count forecasting at the county, health district, and state levels during early onset, the following were found: (1) the MA method was effective for forecasting 1-, 3-, and 7-day cumulative case counts; (2) exponential growth was not the best representation of case growth during early virus onset when the public was aware of the virus; and (3) geographic resolution was a factor in the selection of forecasting methods.

(*J Med Internet Res* 2021;23(3):e24925) doi: [10.2196/24925](https://doi.org/10.2196/24925)

KEYWORDS

coronavirus disease 2019; COVID-19; infectious disease; emerging outbreak; forecasting; modeling and simulation; public health; modeling disease outbreaks

Introduction

The scientific community responded quickly to the global outbreak following COVID-19's identification in December of 2019 [1,2]. Numerous platforms and studies have been created to forecast the spread of the pandemic and meet the need for intervention measures in support of public health and awareness [1,3-5]. Many forecasting efforts focused on the long-term identification of COVID-19 and the flattening of and getting over the curve [3,4,6]. Forecasts assist in identifying and evaluating long-term preventions. Short-range forecasts provide benefit by supporting local understanding for individuals and policy makers and supporting short-range decisions. To inform public health and support awareness of proper forecast interpretation when making decisions, it is important to understand the basics of the generation of forecasts and the boundaries under which their interpretations are valid. To this point, this study explores the error levels of seven common forecasting methods in estimating COVID-19 cumulative case counts at county, health district, and state levels over the upcoming week. Comparing error levels across forecasting methods and geographic granularities provides insight into the assumptions contributing to more accurate forecasts.

Small numbers of COVID-19 cases can lead to large outbreaks [7]. Isolation and preventative measures are recommended practices to reduce the spread of COVID-19 [1-3,8-14]. Forecasts with high error magnitudes can provide expectations that grossly underestimate or overestimate case counts. This can lead to problems, such as the creation of unanticipated hot spots resulting from underestimation, or can cause unnecessary public alarm from overestimation. Interpreting COVID-19 forecasts depends on assumptions such as the geographic area, preventative measures in place, and the population's knowledge of, and behaviors toward, the virus. As assumptions change, the usefulness of the forecasting method should be re-evaluated. The impact of nonpharmaceutical interventions can be delayed 1 to 3 weeks and should factor into policy makers' decisions [15]. Intervention methods can result in secondary effects, such as decreasing levels of physical activity while people practice social distancing [16]. As a result, understanding of the assumptions pertaining to short-range COVID-19 forecasting is needed to properly interpret their findings [17].

This study explores seven commonly utilized forecasting approaches, including the following: naïve [18], moving average (MA) [9,10], autoregressive (AR) [17], growth rate [19], Holt-Winters (HW) exponential smoothing [20,21], autoregressive moving average (ARMA) [22], and autoregressive integrated moving average (ARIMA) [23]. Each forecasting method utilizes different assumptions about how the past values impact the forecast values. The naïve approach is the simplest method and assumes no change from the current value. The MA approach assumes equal weighting of prior values, while exponential smoothing assigns exponentially decreasing weight to older values. The AR approach assumes

linear dependency of prior values but with an added stochastic component. The growth rate approach assumes a linear relationship to its prior values and applies sampling with growth based on the number of increased cases from the prior day. The ARMA approach combines the AR approach to provide a regression based linearly on its past values with the MA approach to account for the error terms within the prior values. The ARIMA approach applies to data that is nonstationary around a mean value and applies a distancing measure one or more times to make the data stationary [24].

Error represents the inability to account for all the variability contributing to changes in COVID-19 case counts. Forecast error represents the under- or overestimation of the actual value [18]. Additionally, assumptions are unlikely to remain constant over time due to shifting public behaviors and implemented public policies. Error magnitude communicates the accuracy of a forecast and can be utilized as a metric to select from a set of potential forecasting methods. Interpreting forecast outcomes relies on the error magnitude as well as situating the assumptions underlining the forecast [25,26]. This means that the effectiveness of a current forecasting method is likely to be impacted as new preventative measures are put into place that alter spread dynamics. Conveying this understanding to the public advances knowledge of health statistics and statistical literacy in public health [21,27-29].

Recommendations for models of infectious diseases in support of public health involve incorporating policy questions, available data, and scientific understanding to yield policy advice, data collection, and scientific insight [30]. By evaluating forecasts by aggregating information from lower levels, such as at the county and health district levels, intervention strategies can be more readily applied based on the relevant demographic characteristics shared by the smaller population samples. Forecasting methods operate under differing assumptions pertaining to how the prior values relate to forecasted values. This study evaluates seven forecasting methods with varied look-back and forecast lengths at the county, health district, and state levels. By evaluating short-range forecasting methods combined with varied look-back and forecast lengths, forecasts can be more effectively used for informing health planning and aiding individuals in evaluating the safety levels of their local and neighboring communities.

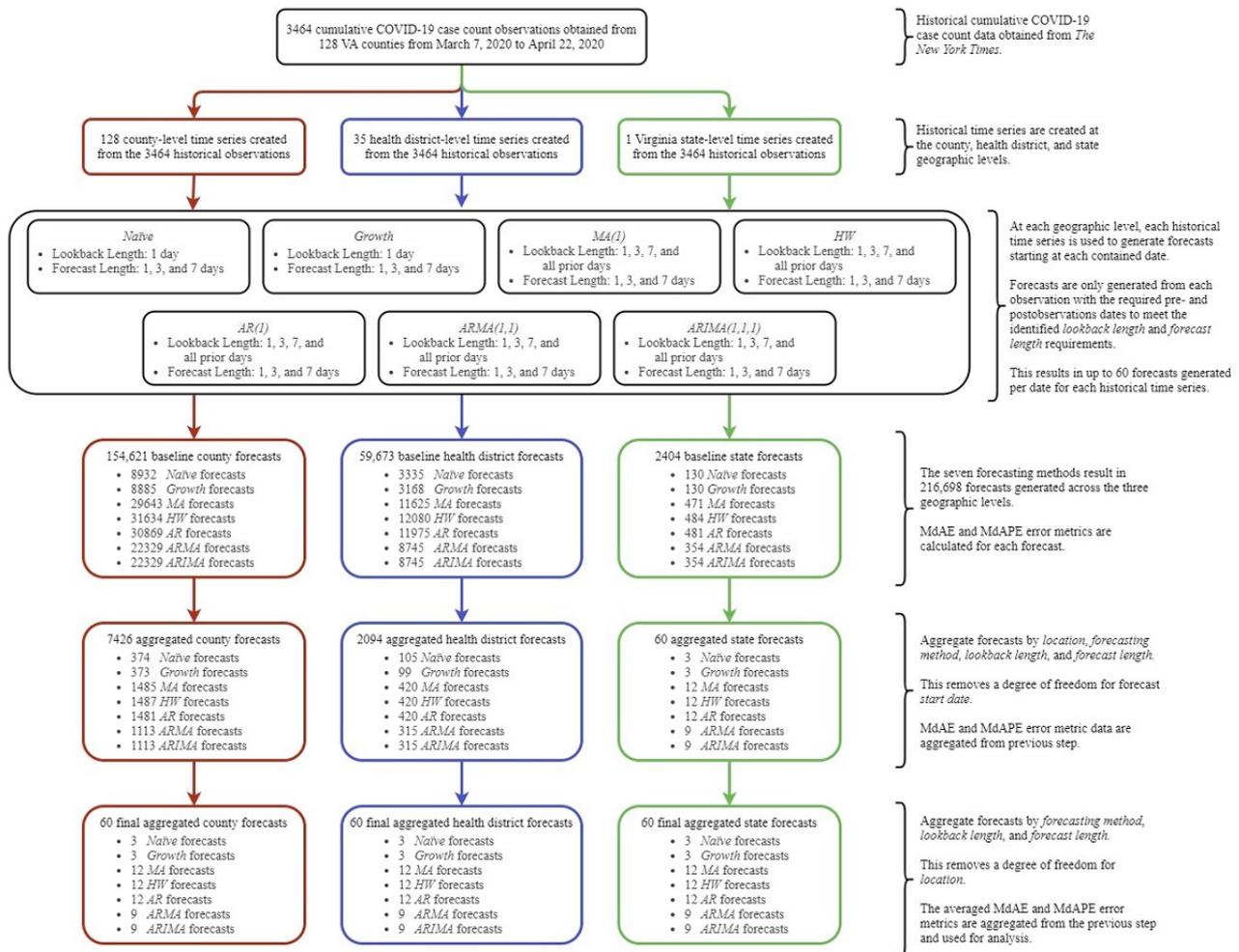
Methods

Data

We obtained 3464 Virginia county-level COVID-19 cumulative case count observations from March 7 to April 22, 2020, using data provided by *The New York Times* and aggregated these observations to the health district and state levels as presented in Figure 1. This period captured the first 3 weeks following the first confirmed COVID-19 case within Virginia and the 3 weeks following the governor's executive order limiting gatherings to groups of less than 10 people. As intervention

measures can take up to 3 weeks to impact the virus spread [15], this time frame was expected to cover Virginia’s case growth prior to experiencing the benefits resulting from the governor’s imposed group size limit.

Figure 1. Experimental design and data overview at the county, health district, and state levels. The generation and aggregation of county-level forecasts are shown on the left path (red), health district-level forecasts on the middle path (blue), and state-level forecasts on the right path (green). The information on the right provides additional detail on each stage in the experimental design. AR: autoregressive; ARIMA: autoregressive integrated moving average; ARMA: autoregressive moving average; HW: Holt-Winters; MA: moving average; MdAE: median absolute error; MdAPE: median absolute percentage error; VA: Virginia.



Forecasting Methods and Assumptions

For the naïve forecasts, the prior day’s value is used for each of the following j forecasted days. For the HW forecasts, exponential smoothing of the prior k day’s values is used to forecast values over the next j days. For the growth rate forecasts, the prior 1 day’s value is used to calculate the current growth rate over the following j days. Then, the prior day’s values for all the counties are used to calculate the growth rate for Virginia for the same j days. A group of n forecasts are generated for the county by uniformly sampling a growth rate between the county’s rates and Virginia’s rates. The average of the n forecasts is utilized as the final forecast for the county. For the MA (1), AR (1), ARMA (1, 1), and ARIMA (1, 1, 1) forecasts, the prior k days are given equal weighting to forecast the next j days.

This study only relies on the daily reported case numbers since the date of first onset within each location and does not incorporate assumptions about the basic reproductive number

of COVID-19. Forecasts are influenced by the reliability of the data, the variables utilized, and the perceptions and reactions to danger and they assume the continuation of past patterns [4,31]. When exploring real-time forecasts of infectious disease models, real-time models have shown higher absolute error values, on average, than full-data models as a result of factors such as significant differences in population sizes between compared areas [32].

Statistical Analysis

We aggregated median absolute error (MdAE) and median absolute percentage error (MdAPE) variables and expressed them as medians, IQRs, and notch ranges representing 95% CIs. Notch ranges are calculated as $\pm 1.58 \times IQR/\sqrt{n}$ [33,34] as implemented using the R function geom_boxplot within the package ggplot2 (The R Foundation) [35]. Nonoverlapping comparisons of confidence intervals represent statistically significant differences [36,37] with P values less than .01 [38-41]. P values conveying significant differences between groups are calculated using Mood’s median test [42,43]. MdAE

compared forecasting outcomes at shared geographic levels due to similarities in scale [44]. MdAPE compared each forecasting method's outcomes across geographic level due to differing scales [45].

We created 226,468 forecasts across the county, health district, and state levels over the period of March 7 through April 22, 2020. Due to the naïve and growth rate methods only utilizing 1-day look-backs and the ARMA and ARIMA methods requiring more than 1-day look-backs, five forecast methods exist for each comparison. Analyses were performed with R software, version 3.6.3 (The R Foundation).

Verification, Validation, and Reproducibility

The data set and code are provided in Lynch and Gore [46] and the experimental methods and steps needed for reproducibility are provided in Lynch and Gore [47]. Code inspections and unit tests were utilized for code verification [48]. MdAE and MdAPE error metrics were used for validation. A comparison of COVID-19 case count data sources found that the differences in reported case counts between *The New York Times*, *Johns Hopkins University*, and *USA Facts* do not indicate inferior or superior sources [49].

Results

Overview

Comparing all forecast methods' MdAE values across the county, health district, and state levels over the first 46 days of infection revealed that MA forecasts using 3-day look-back and 1-day forecast length achieved the lowest MdAE. This MA forecast combination was statistically significantly different in MdAE from 39 of the 59 other combinations at the county level (66%), 53 of 59 (90%) at the health district level, and 51 of 59 (86%) at the state level. This result shows that the use of an equally weighted linear dependency with a stationary mean between the prior 3-day COVID-19 cumulative case counts,

within the MA forecasts, is an effective assumption when forecasting next-day case growths for Virginia at the county, health district, and state levels. Table 1 provides the method with the lowest MdAE and the percentage of other methods from which the difference is determined to be statistically significant at the county level. Table 2 provides this information at the health district level and Table 3 provides this information at the state level.

For the methods using single-day look-back across all levels, the growth rate and naïve methods provided the lowest MdAE at the county, health district, and state levels for all forecast lengths. In general, all five methods achieved similar error confidence intervals when utilizing 1-day look-back. Only at the health district and state levels for 1-day forecast lengths was the growth rate method's difference from the other methods statistically significant, with the growth rate method performing better than all combinations at the state level.

For the methods using 7-day look-backs across all levels, the MA and AR methods were the only ones with MdAE instances that were statistically significantly lower than the other methods. The HW and ARIMA methods achieved the lowest MdAE in two instances but did not perform significantly better than the other methods in either instance. In no instance did the ARMA method obtain the lowest MdAE. The performance of the MA and AR methods supports the assumption of linear dependence between the 7-day prior days' cases and the forecast case counts. However, for the MA method the mean weighting of past values was stationary, while for the AR method it was nonstationary.

For the methods using look-backs of all prior-day case counts across all levels, the MA method achieved the lowest MdAE in all cases. This provides evidence in support of forecasting cumulative case counts using the assumption of a linear dependency and stationary mean among past values to forecast 1, 3, and 7 days when incorporating all prior cumulative cases.

Table 1. County-level median absolute error (MdAE) outcomes by forecasting method, look-back length, and forecast length.

Methods	Look-back length (days), n	Look-ahead length (days), n	df ^a	Forecasting method with lowest MdAE	Median (95% CI) (cumulative cases)	P value ^b	Statistically significantly lower MdAE than other methods ^c , n (%)
All (N=60)	All	All	59	MA (3, 1) ^d	0.67 (0.49-0.84)	<.001	39 (66)
G1 ^e (n=5)	1	1	4	Naïve	0.67 (0.43-0.90)	.09	0 (0)
G1 (n=5)	1	3	4	Naïve	1.30 (0.88-1.73)	.66	0 (0)
G1 (n=5)	1	7	4	Naïve	2.43 (1.69-3.18)	.50	0 (0)
G2 ^f (n=5)	3	1	4	MA	0.67 (0.49-0.84)	.09	0 (0)
G2 (n=5)	3	3	4	MA	0.76 (0.59-0.94)	<.001	4 (100) ^g
G2 (n=5)	3	7	4	MA	1.69 (1.36-2.01)	<.001	3 (75)
G2 (n=5)	7	1	4	HW	0.91 (0.63-1.18)	.03	0 (0)
G2 (n=5)	7	3	4	MA	1.30 (0.95-1.65)	.002	1 (25)
G2 (n=5)	7	7	4	MA	2.32 (1.75-2.90)	.01	0 (0)
G2 (n=5)	All prior	1	4	MA	0.70 (0.53-0.87)	.33	0 (0)
G2 (n=5)	All prior	3	4	MA	0.83 (0.67-1.00)	<.001	4 (100) ^g
G2 (n=5)	All prior	7	4	MA	1.73 (1.36-2.10)	<.001	1 (25)

^aDegrees of freedom represent the number of forecasting combinations minus one.

^bP values were calculated for statistically significant differences in medians across groups.

^cThis was based on comparisons of notch ranges. MdAE was interpreted within geographic levels.

^dMA: moving average; (3, 1) represents a 3-day look-back and a single-day forecast length.

^eG1 includes naïve, MA, autoregressive (AR), growth rate, and Holt-Winters (HW) methods.

^fG2 includes MA, AR, growth rate, HW, autoregressive moving average, and autoregressive integrated moving average methods.

^gMA (3, 3) and MA (all prior, 3) achieved statistically significantly smaller MdAE than all four alternatives.

Table 2. Health district–level median absolute error (Mdae) outcomes by forecasting method, look-back length, and forecast length.

Methods	Look-back length (days), n	Look-ahead length (days), n	df ^a	Forecasting method with lowest Mdae	Median (95% CI) (cumulative cases)	P value ^b	Statistically significantly lower Mdae than other methods ^c , n (%)
All (N=60)	All	All	59	MA (3, 1) ^d	3.07 (2.41-3.74)	<.001	53 (90)
G1 ^e (n=5)	1	1	4	Growth rate	4.03 (3.01-5.04)	.31	1 (25)
G1 (n=5)	1	3	4	Growth rate	8.96 (6.56-11.36)	.93	0 (0)
G1 (n=5)	1	7	4	Growth rate	16.48 (11.67-21.28)	.96	0 (0)
G2 ^f (n=5)	3	1	4	MA	3.07 (2.41-3.74)	.01	1 (25)
G2 (n=5)	3	3	4	MA	3.20 (2.50-3.90)	<.001	4 (100) ^g
G2 (n=5)	3	7	4	MA	7.88 (5.71-10.05)	<.001	1 (25)
G2 (n=5)	7	1	4	AR	3.57 (2.67-4.47)	.01	0 (0)
G2 (n=5)	7	3	4	MA	5.52 (3.96-7.08)	<.001	2 (50)
G2 (n=5)	7	7	4	AR	11.83 (8.16-15.49)	<.001	1 (25)
G2 (n=5)	All prior	1	4	MA	3.14 (2.47-3.80)	.04	1 (25)
G2 (n=5)	All prior	3	4	MA	3.16 (2.54-3.78)	<.001	3 (75)
G2 (n=5)	All prior	7	4	MA	7.68 (6.22-9.14)	<.001	3 (75)

^aDegrees of freedom represent the number of forecasting combinations minus one.

^bP values were calculated for statistically significant differences in medians across groups.

^cThis was based on comparisons of notch ranges. Mdae was interpreted within geographic levels.

^dMA: moving average; (3, 1) represents a 3-day look-back and a single-day forecast length.

^eG1 includes naïve, MA, autoregressive (AR), growth rate, and Holt-Winters (HW) methods.

^fG2 includes MA, AR, growth rate, HW, autoregressive moving average, and autoregressive integrated moving average methods.

^gMA (3, 3) achieved statistically significantly smaller Mdae than all four alternatives.

Table 3. State-level median absolute error (MdAE) outcomes by forecasting method, look-back length, and forecast length.

Methods	Look-back length (days), n	Look-ahead length (days), n	df ^a	Forecasting method with lowest MdAE	Median (95% CI) (cumulative cases)	P value ^b	Statistically significantly lower MdAE than other methods ^c , n (%)
All (N=60)	All	All	59	MA (3, 1) ^d	17.43 (7.74-27.11)	<.001	51 (86)
G1 ^e (n=5)	1	1	4	Growth rate	31.50 (6.11-56.89)	<.001	4 (100) ^f
G1 (n=5)	1	3	4	Growth rate	317.50 (163.15-471.85)	.94	0 (0)
G1 (n=5)	1	7	4	Growth rate	325.00 (169.49-480.51)	.18	0 (0)
G2 ^g (n=5)	3	1	4	MA	17.43 (7.74-27.11)	<.001	4 (100) ^f
G2 (n=5)	3	3	4	MA	64.94 (45.93-83.96)	<.001	1 (25)
G2 (n=5)	3	7	4	MA	206.57 (148.57-264.94)	.03	1 (25)
G2 (n=5)	7	1	4	AR	69.37 (34.23-104.51)	.09	1 (25)
G2 (n=5)	7	3	4	MA	82.14 (42.83-121.47)	.02	2 (50)
G2 (n=5)	7	7	4	ARIMA	312.36 (146.54-478.17)	.012	0 (0)
G2 (n=5)	All prior	1	4	MA	25.13 (11.61-38.65)	.27	0 (0)
G2 (n=5)	All prior	3	4	MA	32.67 (21.20-44.14)	.002	2 (50)
G2 (n=5)	All prior	7	4	MA	104.85 (70.67-139.03)	.09	2 (50)

^aDegrees of freedom represent the number of forecasting combinations minus one.

^bP values were calculated for statistically significant differences in medians across groups.

^cThis was based on comparisons of notch ranges. MdAE was interpreted within geographic levels.

^dMA: moving average; (3, 1) represents a 3-day look-back and a single-day forecast length.

^eG1 includes naïve, MA, autoregressive (AR), growth rate, and Holt-Winters (HW) methods.

^fGrowth rate (1, 1) and MA (3, 1) achieved statistically significantly smaller MdAE than all four alternatives.

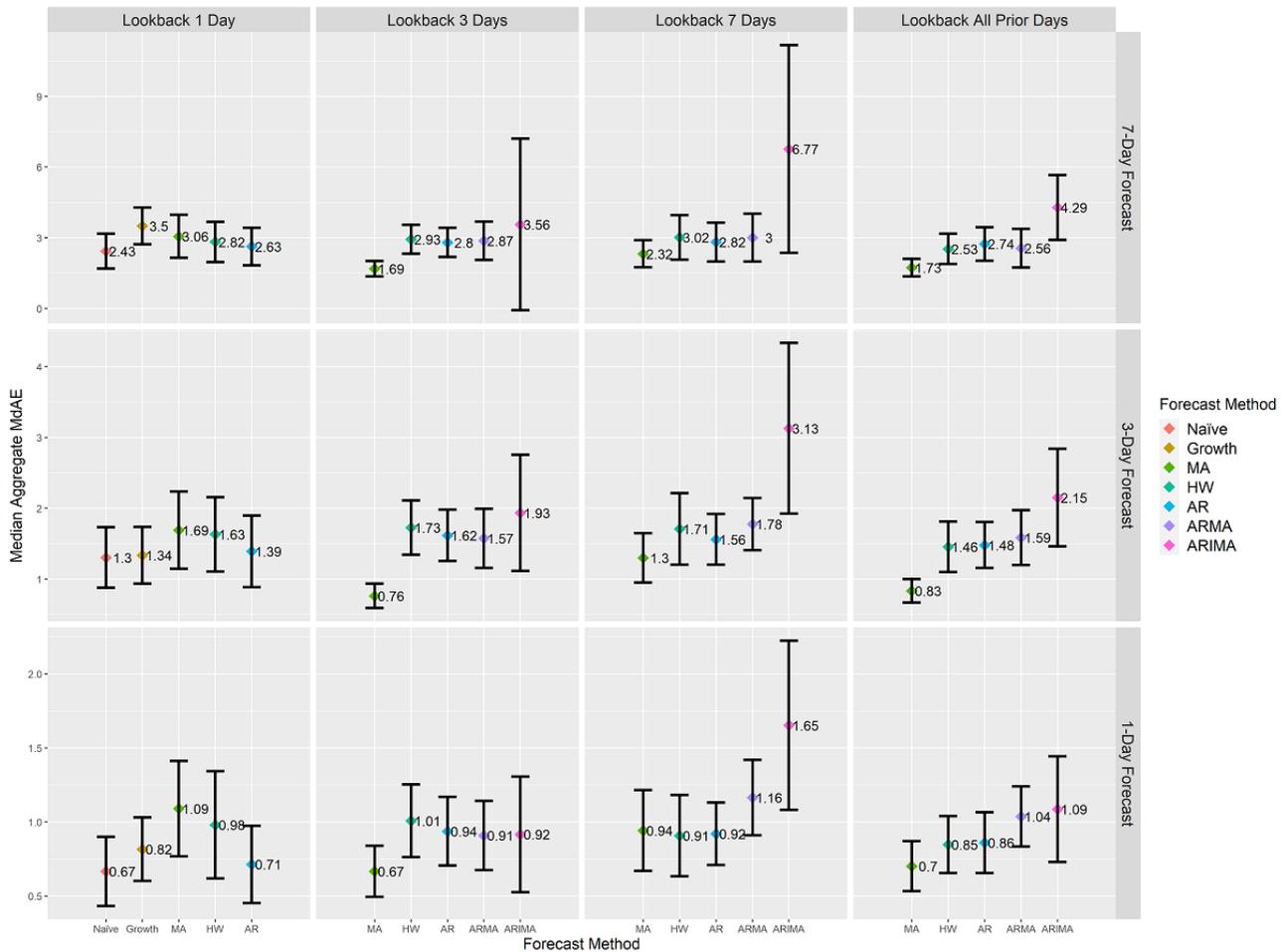
^gG2 includes MA, AR, growth rate, HW, autoregressive moving average, and autoregressive integrated moving average (ARIMA) methods.

County-Level MdAE Results

At the county level, the MA method always achieved a lower MdAE than the ARMA method. Similarly, the ARMA method always achieved a lower MdAE than the ARIMA method. Thus, the ARIMA method’s aggregated error was greater than the ARMA method’s aggregated error, which was greater than the MA method’s aggregated error. This indicates that the assumption of a stationary mean (ie, MA) in prior case counts is more effective than the assumption of a weakly stationary mean (ie, ARMA), which is more effective than a nonstationary

mean (ie, ARIMA) when forecasting at the county level. The ARIMA method had the widest confidence interval for the median error range, indicating the least consistency in COVID-19 forecasts among these methods. Figure 2 provides the county-level MdAE outcomes for each look-back and forecast length combination. The individual results of each of the 60 forecasting combinations at the county level are provided in Multimedia Appendix 1, including median values, confidence intervals, whiskers, sample sizes, and P values. An interactive version of Figure 2 is provided in Multimedia Appendix 2.

Figure 2. County-level forecasts’ aggregated median MdAE values and 95% CI. CI ranges are calculated using box plot notch ranges around the median. Statistically significant differences at a *P* value of .01 are identified by nonoverlapping CI ranges of forecasting methods at each combination of forecast length and look-back length. Units are in terms of COVID-19 cumulative case counts. Y-axis scales differ on each row based on the scale of the contained data. Due to differing assumptions, five of the seven forecasting methods are present for each look-back length as indicated on the x-axis. AR: autoregressive; ARIMA: autoregressive integrated moving average; ARMA: autoregressive moving average; HW: Holt-Winters; MA: moving average; MdAE: median absolute error.

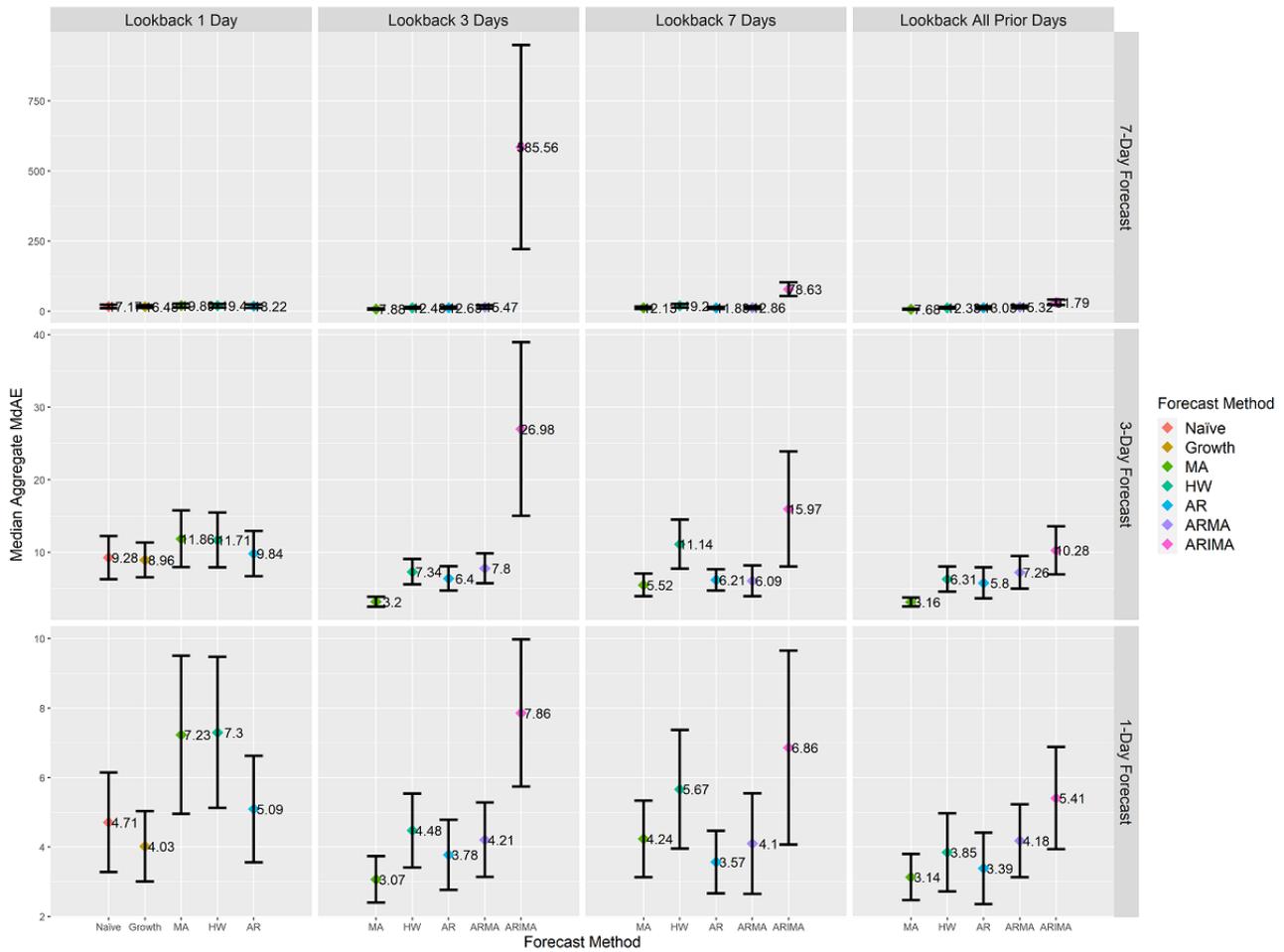


Health District–Level MdAE Results

At the health district level, the MA method always achieved lower MdAE than the ARMA method, which achieved lower MdAE than the ARIMA method. This further provided evidence that effective forecasting of cumulative COVID-19 case counts contains an assumption of stationary means in past observations. For 3-day look-back lengths with 3-day forecasts, the MA

method achieved statistically significantly lower MdAE than all other methods. Figure 3 provides the MdAE at the intersection of look-back length and forecast length at the health district level. The individual results of each of the 60 forecasting combinations at the health district level are provided in Multimedia Appendix 3, including median values, confidence intervals, whiskers, sample sizes, and *P* values. An interactive version of Figure 3 is provided in Multimedia Appendix 4.

Figure 3. Health district–level forecasts’ aggregated median MdAE values and 95% CI. CI ranges are calculated using box plot notch ranges around the median. Statistically significant differences at a *P* value of .01 are identified by nonoverlapping CI ranges of forecasting methods at each combination of forecast length and look-back length. Units are in terms of COVID-19 cumulative case counts. Y-axis scales differ on each row based on the scale of the contained data. Due to differing assumptions, five of the seven forecasting methods are present for each look-back length as indicated on the x-axis. AR: autoregressive; ARIMA: autoregressive integrated moving average; ARMA: autoregressive moving average; HW: Holt-Winters; MA: moving average; MdAE: median absolute error.

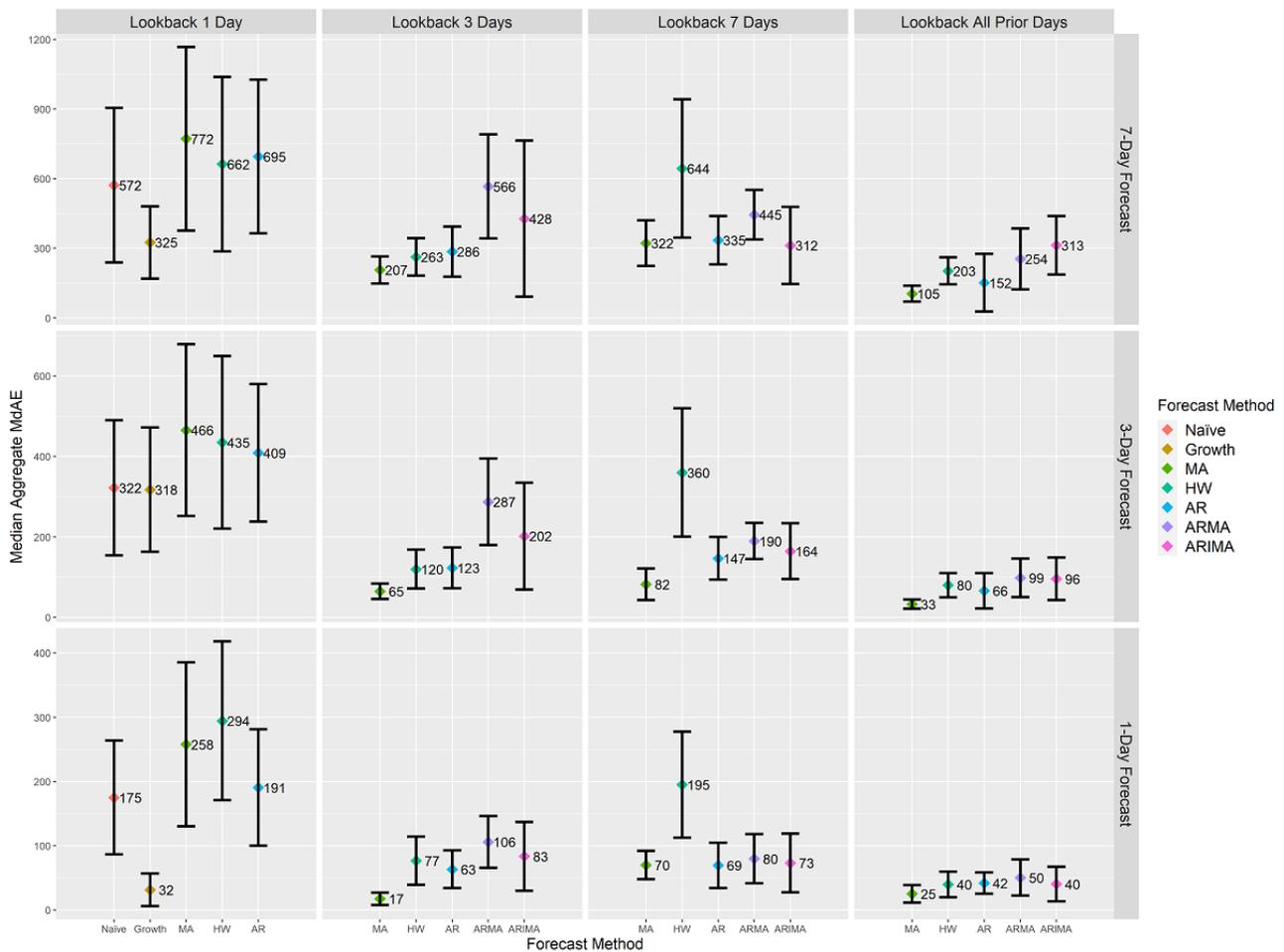


State-Level MdAE Results

At the state level, the growth rate method was the most effective method. In every case, it either (1) attained the lowest MdAE value compared to the other methods or (2) had the smallest notch range. The ARMA and ARIMA methods both maintained MdAE notch bands that were similar to the other methods when utilizing all prior day and 7-day look-back lengths. However, the HW method’s MdAE notch bands increased as that of the ARMA and ARIMA methods decreased. These results make it

unclear as to which of the assumptions related to stationary means were most effective for forecasting with the ARMA and ARIMA methods at the state level. Figure 4 provides MdAE values at the intersection of look-back length and forecast length at the state level. The individual results of each of the 60 forecasting combinations at the state level are provided in Multimedia Appendix 5, including median values, confidence intervals, whiskers, sample sizes, and *P* values. An interactive version of Figure 4 is provided in Multimedia Appendix 6.

Figure 4. State-level forecasts' aggregated median MdAE values and 95% CI. CI ranges are calculated using box plot notch ranges around the median. Statistically significant differences at a *P* value of .01 are identified by nonoverlapping CI ranges of forecasting methods at each combination of forecast length and look-back length. Units are in terms of COVID-19 cumulative case counts. Y-axis scales differ on each row based on the scale of the contained data. Due to differing assumptions, five of the seven forecasting methods are present for each look-back length as indicated on the x-axis. AR: autoregressive; ARIMA: autoregressive integrated moving average; ARMA: autoregressive moving average; HW: Holt-Winters; MA: moving average; MdAE: median absolute error.

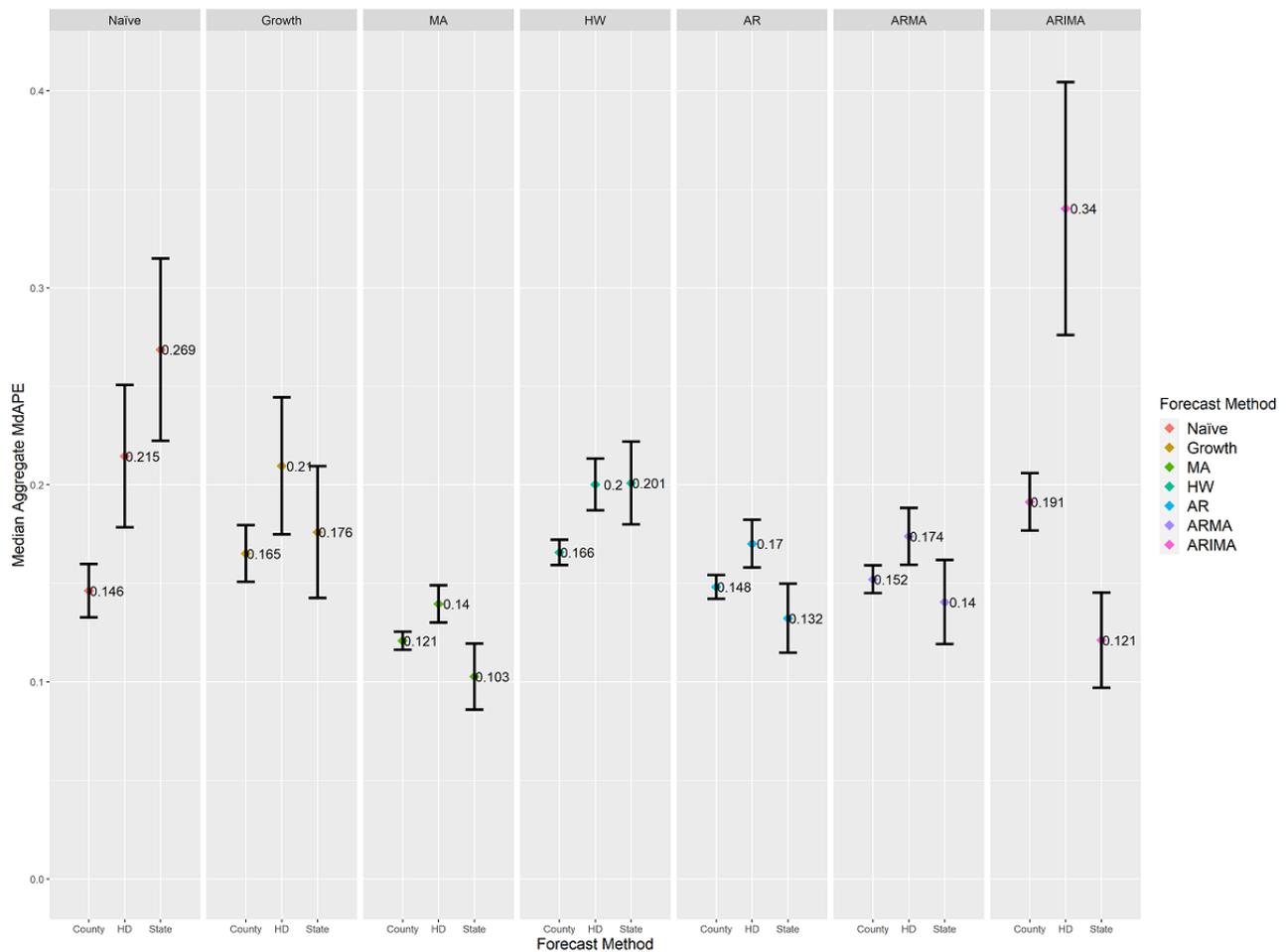


Cross-Geographic-Level MdAPE Results

MdAE reflects the scale of the data and is not appropriate for making inferences about changes in confirmed case counts between county, health district, and state levels [26,44]. Figures 2-4 convey differing scales of error values across the three levels. As a result, it was not possible to evaluate results featured in these figures against each other. To remedy this shortcoming, we applied MdAPE to identify statistically significant

differences for each forecasting method individually when applied to county, health district, and state levels as provided in Figure 5. The individual results of each of the 60 forecasting combinations at the county, health district, and state levels are provided in Multimedia Appendices 7-9, including median values, confidence intervals, whiskers, sample sizes, and *P* values. An interactive version of Figure 5 is provided in Multimedia Appendix 10.

Figure 5. Aggregated median MdAPE values and 95% CI ranges at the county, health district (HD), and state levels differentiated by forecasting method. Comparing CI ranges for a forecast method across each geographic level reveals statistically significant differences in median values for the forecasting method due to geographic scale. Nonoverlapping CI ranges indicate statistically significant differences at a *P* value of .01. MdAPE provides a comparison within each forecast method separately, not a comparison across different methods. AR: autoregressive; ARIMA: autoregressive integrated moving average; ARMA: autoregressive moving average; HW: Holt-Winters; MA: moving average; MdAPE: median absolute percentage error.



Statistically significant differences were observable within a forecasting method across the county, health district, and state levels during the initial 46 days of confirmed COVID-19 case spread within Virginia. The growth rate method was the only one whose performance did not statistically significantly differ across levels; thereby, it was the only method unaffected by geographic level. The naïve method achieved a statistically significantly lower MdAPE at the county level than at the health district and state levels. The MA, HW, AR, ARMA, and ARIMA methods all contained instances of achieving statistically significantly lower MdAPE scores at the county and/or state levels than at the health district level.

Discussion

Principal Findings

Our results show the effectiveness of seven forecasting methods for the first 46 days of virus spread within Virginia at the county, health district, and state levels. In addition, a daily view of the growth rate forecast at the county level from March 7, 2020, to the present is publicly available online [50]. Tracking case and death counts yield insight into the virus’s impact on a geographic region at a given point in time. Forecasts utilize the trends and averages of prior case count observations to provide expectations

of case counts into the future. These forecasts keep the public informed on the state of the virus across the world and on virus levels within their own geographic areas of interest. Additionally, forecasts inform public policy for combatting the spread of the virus, supporting public health, and helping to anticipate the impacts of medical burdens across regions [51]. However, interpreting forecast outcomes requires understanding the assumptions behind the forecasting method as well as the assumptions pertaining to the geographic area and the presence of intervention strategies. Therefore, we compared the error levels pertaining to 60 forecasting combinations using the MA, AR, naïve, growth rate, HW, ARMA, and ARIMA forecasting methods. Our findings support public health with respect to forecasting by reinforcing health statistics and statistical literacy of forecasted COVID-19 outcomes.

COVID-19 cumulative case growth is such that the growth curve is exponential in the absence of preventative measures. The larger error observed in HW forecasts over MA forecasts provides support that an exponential model is not the best fit at the start of the virus spread for Virginia. The preventative measures taken by the population appear to have shifted the virus’s growth behavior from exponential to linear. This finding is also supported by Lammers et al [21]. This finding supports

the idea that population interventions are effective at impacting the spread of the virus. However, as the virus continues to spread and reoccur, the inability to manufacture a vaccine for COVID-19 quickly enough to immunize the population remains a concern [52]. As such, combining short-range forecasts at the county and health district levels with targeted intervention strategies can improve planning, support, and response time. The use of rigorous government interventions may slow the rate of infections, but early detection, isolation, treatment, and adequate medical supplies are required for continued intervention against the virus [53,54].

Our cross-geographic validation checks using MdAPE indicate that the level of geographic resolution should be considered when creating forecasts of expected case counts. A forecast utilized at the state level is not likely to be as useful for determining expected growth when disaggregated across its counties during early virus onset. This results from the differing geographic assumptions present within counties or health districts when compared to the state. This finding is consistent with the literature reporting that case growths vary across countries and across states [14]. Variations result from factors such as population behaviors in response to the pandemic, implemented policy interventions, and population densities [7,9,10,15]. Furthermore, since the growth rate method did not produce statistically significant MdAPE differences across geographic levels, it may be a good choice for decision makers whose region does not match the county, health district, or state levels.

To identify a best option among our tested combinations, we compared MdAE ranges against each other within each geographic tier. The MA method using a 3-day look-back length and a single-day forecast length provided the smallest error (ie, lowest MdAE) at the county level (median 0.67, 95% CI 0.49-0.84; $P < .001$), the health district level (median 3.07, 95% CI 2.41-3.74; $P < .001$), and the state level (median 17.43, 95% CI 7.74-27.11; $P < .001$). Compared to the other forecasting combinations, the MA method's confidence intervals statistically significantly differed from 39 out of 59 alternatives (66%; county level) to 51 out of 59 alternatives (86%; state level) to 53 out of 59 alternatives (90%; health district level) at a P value level of .01. When relying on only the prior day's case counts, the growth rate method stood out as the best option at the health district and state levels; however, the naïve, growth rate, HW, MA, and AR methods performed similarly well at the county level.

When utilizing 3 or more days of prior observations, a diverse range of options is available. For next-day forecasts, there was no method that performed better at a level that was statistically significant among the five options. For 3-day forecasts, the MA method was statistically significantly better than 25% to 100% of the other four options in all cases. For 7-day forecasts, the MA method performed statistically significantly better than 25% to 75% of the other four options when using a 3-day look-back or an all prior-day look-back. When using a 7-day look-back, the AR method performed the best and its difference from the ARIMA method was statistically significant. The ARMA and ARIMA methods achieved the lowest error in any of the combinations.

These findings support the assumption of stationarity within the mean of the prior days' cumulative case counts. This is reflected in how well the MA method performed and how poorly the ARIMA method performed at forecasting cumulative case counts. Rarely do the ARIMA or ARMA methods achieve lower error values than any other combination. This reflects the idea that the assumption of stationary means of past observations is a more effective representation of cumulative COVID-19 growth than assumptions of weak stationarity or nonstationarity. The need to apply a differencing step to remove nonstationarity using the ARIMA method is not present within the data during this period. Additionally, placing extra weight on the recent past does not improve forecasting during this period, as the HW and AR methods were consistently less effective than the MA method. These findings suggest that the ARIMA and ARMA methods are unlikely to be good fits and should not be used to forecast case counts during early onset within areas that have only a few weeks of historical data collected, whose residents are aware of the existence of the virus and are engaging in preventative behaviors, and that contain similar population densities to Virginia.

Several studies utilized forecasts to estimate case fatality and recovery ratios, epidemiological parameters, and transmission dynamics based on data from the start of the outbreak [55,56]. Studies also support the idea that epidemiological differences contribute to variations in the severity of the contracted disease [2,57]. Based on historical similarities to previous influenza strains, social distancing can potentially reduce transmission of the virus; however, the effectiveness may vary alongside changes in seasonal factors in travel as well as between tropical and temperate climates [58,59]. Distancing may be especially beneficial in rural areas, where fewer hospitals and health care facilities exist, by emphasizing strategies oriented toward specific population age groups [60].

The results of this study can be expanded to include areas' demographic characteristics, geographic characteristics, and preventative measures to strive for more accurate forecast models. A recent study found COVID-19 growth to strongly correlate with population density, percent of the population living in rural areas, and yearly flu vaccination rate [14]. Exploring forecast behaviors of areas sharing these traits may further reduce forecasting error and reveal subgroupings of viable forecasting method options. Additionally, forecasting models can be paired with mortality models [61] to gain better estimates of infection forecasts per demographic characteristics. Predictive methods derived from search engines' data [62] can also be incorporated within forecasting methods. This would provide a way to connect forecasts with human search behaviors based on the frequency of searched terms identified in relation to COVID-19 prevention and recovery. Forecasting models can be paired with models of local medical burden and pandemic preparedness [51] for more detailed representations of expected medical strains and greater flexibility in testing preventative measures.

Stay-at-home orders have been successful as intervention strategies to slow the spread of the virus. However, the impact caused by a neighboring area's removal of mobility restrictions needs greater exploration [10]. The secondary effects of the

starting or ending of proximity-based prevention methods on neighboring areas can help assess the potential impact of a mitigation strategy. For identified hot spots, the Centers for Disease Control and Prevention provides outreach to local officials and helps in identifying adapted interventions for the local area [7]. Rapid identification and timeliness of response are critical, especially if the impact of interventions can take up to 3 weeks to be effective [15]. Reliable forecasting can aid in the identification of emerging hot spots and support timely response. To this end, increased knowledge of forecasting characteristics based on geographic level, demographic characteristics, population density characteristics, and population behaviors can help reveal the primary drivers of upcoming cases. This knowledge can be leveraged to inform early, targeted interventions or to provide risk updates to targeted populations within an area. People could then modify their mobility and social decisions themselves in a timely manner separate from population mandated measures.

Limitations

Here we discuss internal and external validity threats as well as other limitations that affected our work. Internal validity threats arise when factors affect the dependent variables without the researchers' knowledge. It is possible that some implementation flaws could have affected our modeling results or the ensuing data analysis. However, the algorithms in our source code were (1) built on established libraries, (2) passed several internal code reviews, and (3) are publicly accessible, along with the data and results. Threats to external validity occur when the results of our analysis and our simulation cannot be generalized. Our results are limited to Virginia, from March to April 2020 with respect to COVID-19 cases reported by *The New York Times*. Our results are not immediately generalizable to (1) different infectious diseases, (2) other COVID-19 data sets, (3) different periods of time, or (4) different geographic areas.

Several other assumptions and limitations pertain to this study. Seven forecasting methods with differing baseline assumptions were evaluated with respect to how well they forecast the early growth of COVID-19 cases within Virginia; however, numerous additional forecasting methods exist with different combinations of assumptions that can also be explored with respect to this pandemic. Conclusions should not be drawn about the effectiveness of these findings for forecast lengths greater than 7 days, as the appropriateness of underlying assumptions, such as stationarity of prior days' values, would need to be re-evaluated. Larger median error values of the 7-day forecasts, versus their 1- and 3-day counterparts, were observable, further supporting the need for evaluation of forecast assumptions pertaining to the characteristics of COVID-19 beyond 7-day forecast lengths. The selected forecasting methods assumed that policies and population behaviors remained unchanged during the forecast periods; therefore, the forecasts do not account for future starts or ends of policies, such as stay-at-home orders or return-to-work dates.

Finally, our results do not reflect how the spread of the virus would occur for locations experiencing first contact with the virus without yet having public awareness of the existence of the virus. These findings are applicable under the assumption that the general population was already aware of the presence of COVID-19. At the starting point of Virginia cases, the local population was already aware that cases had reached the United States, the virus had been classified as a pandemic, and the virus was receiving major media attention. Additionally, the governor had issued an executive order declaring a state of emergency due to COVID-19 on March 12, 2020 [10,63]. This provided 47 days for Virginians to prepare and modify their standard movement and interaction behaviors as they deemed necessary for their own safety. As a result, these results were not captured in the same context as the period of time when areas within the United States were first impacted by the virus (ie, areas of Washington, California, and New York).

Authors' Contributions

All authors contributed to the conceptualization of the study and the data collection. CL conducted the literature review, constructed the experimental design, and drafted the manuscript. RG developed the procedure for the growth rate forecasts and led code development. All authors contributed to data analysis, data interpretation, figure creation, and revisions and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

County-level median absolute error (MdAE) outcomes for each forecasting combination.

[\[XLSX File \(Microsoft Excel File\), 16 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Interactive plot conveying county-level median absolute error (MdAE) outcomes for each forecasting combination.

[\[ZIP File \(Zip Archive\), 1038 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Health district-level median absolute error (MdAE) outcomes for each forecasting combination.

[\[XLSX File \(Microsoft Excel File\), 16 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Interactive plot conveying health district-level median absolute error (MdAE) outcomes for each forecasting combination.

[\[ZIP File \(Zip Archive\), 1038 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

State-level median absolute error (MdAE) outcomes for each forecasting combination.

[\[XLSX File \(Microsoft Excel File\), 16 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Interactive plot conveying state-level median absolute error (MdAE) outcomes for each forecasting combination.

[\[ZIP File \(Zip Archive\), 1038 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

County-level median absolute percentage error (MdAPE) outcomes for each forecasting combination.

[\[XLSX File \(Microsoft Excel File\), 16 KB-Multimedia Appendix 7\]](#)

Multimedia Appendix 8

Health district-level median absolute percentage error (MdAPE) outcomes for each forecasting combination.

[\[XLSX File \(Microsoft Excel File\), 17 KB-Multimedia Appendix 8\]](#)

Multimedia Appendix 9

State-level median absolute percentage error (MdAPE) outcomes for each forecasting combination.

[\[XLSX File \(Microsoft Excel File\), 17 KB-Multimedia Appendix 9\]](#)

Multimedia Appendix 10

Interactive figure conveying county-, health district-, and state-level median absolute percentage error (MdAPE) outcomes for each forecasting combination.

[\[ZIP File \(Zip Archive\), 1034 KB-Multimedia Appendix 10\]](#)

References

1. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020 Mar 26;382(13):1199-1207 [[FREE Full text](#)] [doi: [10.1056/NEJMoa2001316](https://doi.org/10.1056/NEJMoa2001316)] [Medline: [31995857](https://pubmed.ncbi.nlm.nih.gov/31995857/)]
2. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020 Apr 07;323(13):1239-1242. [doi: [10.1001/jama.2020.2648](https://doi.org/10.1001/jama.2020.2648)] [Medline: [32091533](https://pubmed.ncbi.nlm.nih.gov/32091533/)]
3. Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman J, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect Dis Model* 2020;5:256-263 [[FREE Full text](#)] [doi: [10.1016/j.idm.2020.02.002](https://doi.org/10.1016/j.idm.2020.02.002)] [Medline: [32110742](https://pubmed.ncbi.nlm.nih.gov/32110742/)]
4. Petropoulos F, Makridakis S. Forecasting the novel coronavirus COVID-19. *PLoS One* 2020;15(3):e0231236 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0231236](https://doi.org/10.1371/journal.pone.0231236)] [Medline: [32231392](https://pubmed.ncbi.nlm.nih.gov/32231392/)]
5. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020 May;20(5):533-534 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)] [Medline: [32087114](https://pubmed.ncbi.nlm.nih.gov/32087114/)]
6. Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, et al. Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR Public Health Surveill* 2020 May 13;6(2):e19115 [[FREE Full text](#)] [doi: [10.2196/19115](https://doi.org/10.2196/19115)] [Medline: [32391801](https://pubmed.ncbi.nlm.nih.gov/32391801/)]
7. Oster AM, Kang GJ, Cha AE, Beresovsky V, Rose CE, Rainisch G, et al. Trends in number and distribution of COVID-19 hotspot counties - United States, March 8-July 15, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 21;69(33):1127-1132 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6933e2](https://doi.org/10.15585/mmwr.mm6933e2)] [Medline: [32817606](https://pubmed.ncbi.nlm.nih.gov/32817606/)]
8. Yu P, Zhu J, Zhang Z, Han Y. A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period. *J Infect Dis* 2020 May 11;221(11):1757-1761 [[FREE Full text](#)] [doi: [10.1093/infdis/jiaa077](https://doi.org/10.1093/infdis/jiaa077)] [Medline: [32067043](https://pubmed.ncbi.nlm.nih.gov/32067043/)]

9. Gallaway MS, Rigler J, Robinson S, Herrick K, Livar E, Komatsu KK, et al. Trends in COVID-19 incidence after implementation of mitigation measures - Arizona, January 22-August 7, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Oct 09;69(40):1460-1463 [FREE Full text] [doi: [10.15585/mmwr.mm6940e3](https://doi.org/10.15585/mmwr.mm6940e3)] [Medline: [33031366](https://pubmed.ncbi.nlm.nih.gov/33031366/)]
10. Moreland A, Herlihy C, Tynan MA, Sunshine G, McCord RF, Hilton C, CDC Public Health Law Program, CDC COVID-19 Response Team, Mitigation Policy Analysis Unit. Timing of state and territorial COVID-19 stay-at-home orders and changes in population movement - United States, March 1-May 31, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Sep 04;69(35):1198-1203 [FREE Full text] [doi: [10.15585/mmwr.mm6935a2](https://doi.org/10.15585/mmwr.mm6935a2)] [Medline: [32881851](https://pubmed.ncbi.nlm.nih.gov/32881851/)]
11. Siedner MJ, Harling G, Reynolds Z, Gilbert RF, Haneuse S, Venkataramani AS, et al. Social distancing to slow the US COVID-19 epidemic: Longitudinal pretest–posttest comparison group study. *PLoS Med* 2020 Aug 11;17(8):e1003244. [doi: [10.1371/journal.pmed.1003244](https://doi.org/10.1371/journal.pmed.1003244)]
12. Lyu W, Wehby GL. Shelter-in-place orders reduced COVID-19 mortality and reduced the rate of growth in hospitalizations. *Health Aff (Millwood)* 2020 Sep;39(9):1615-1623. [doi: [10.1377/hlthaff.2020.00719](https://doi.org/10.1377/hlthaff.2020.00719)] [Medline: [32644825](https://pubmed.ncbi.nlm.nih.gov/32644825/)]
13. Turk PJ, Chou S, Kowalkowski MA, Palmer PP, Priem JS, Spencer MD, et al. Modeling COVID-19 latent prevalence to assess a public health intervention at a state and regional scale: Retrospective cohort study. *JMIR Public Health Surveill* 2020 Jun 19;6(2):e19353 [FREE Full text] [doi: [10.2196/19353](https://doi.org/10.2196/19353)] [Medline: [32427104](https://pubmed.ncbi.nlm.nih.gov/32427104/)]
14. White ER, Hébert-Dufresne L. State-level variation of initial COVID-19 dynamics in the United States. *PLoS One* 2020;15(10):e0240648 [FREE Full text] [doi: [10.1371/journal.pone.0240648](https://doi.org/10.1371/journal.pone.0240648)] [Medline: [33048967](https://pubmed.ncbi.nlm.nih.gov/33048967/)]
15. Li Y, Campbell H, Kulkarni D, Harpur A, Nundy M, Wang X, et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: A modelling study across 131 countries. *Lancet Infect Dis* 2021 Feb;21(2):193-202. [doi: [10.1016/s1473-3099\(20\)30785-4](https://doi.org/10.1016/s1473-3099(20)30785-4)]
16. Wang Y, Zhang Y, Bennell K, White DK, Wei J, Wu Z, et al. Physical distancing measures and walking activity in middle-aged and older residents in Changsha, China, during the COVID-19 epidemic period: Longitudinal observational study. *J Med Internet Res* 2020 Oct 26;22(10):e21632 [FREE Full text] [doi: [10.2196/21632](https://doi.org/10.2196/21632)] [Medline: [33027035](https://pubmed.ncbi.nlm.nih.gov/33027035/)]
17. Liu D, Clemente L, Poirier C, Ding X, Chinazzi M, Davis J, et al. Real-time forecasting of the COVID-19 outbreak in Chinese provinces: Machine learning approach using novel digital data and estimates from mechanistic models. *J Med Internet Res* 2020 Aug 17;22(8):e20285 [FREE Full text] [doi: [10.2196/20285](https://doi.org/10.2196/20285)] [Medline: [32730217](https://pubmed.ncbi.nlm.nih.gov/32730217/)]
18. Armstrong JS. The forecasting dictionary. In: Armstrong JS, editor. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Boston, MA: Springer; 2001:761-819.
19. Tovissodé CF, Lokonon BE, Glèlè Kakai R. On the use of growth models to understand epidemic outbreaks with application to COVID-19 data. *PLoS One* 2020;15(10):e0240578 [FREE Full text] [doi: [10.1371/journal.pone.0240578](https://doi.org/10.1371/journal.pone.0240578)] [Medline: [33079964](https://pubmed.ncbi.nlm.nih.gov/33079964/)]
20. Singh RK, Drews M, De La Sen M, Kumar M, Singh SS, Pandey AK, et al. Short-term statistical forecasts of COVID-19 infections in India. *IEEE Access* 2020;8:186932-186938. [doi: [10.1109/access.2020.3029614](https://doi.org/10.1109/access.2020.3029614)]
21. Lammers J, Crusius J, Gast A. Correcting misperceptions of exponential coronavirus growth increases support for social distancing. *Proc Natl Acad Sci U S A* 2020 Jul 14;117(28):16264-16266 [FREE Full text] [doi: [10.1073/pnas.2006048117](https://doi.org/10.1073/pnas.2006048117)] [Medline: [32581118](https://pubmed.ncbi.nlm.nih.gov/32581118/)]
22. Pourghasemi HR, Pouyan S, Farajzadeh Z, Sadhasivam N, Heidari B, Babaei S, et al. Assessment of the outbreak risk, mapping and infection behavior of COVID-19: Application of the autoregressive integrated-moving average (ARIMA) and polynomial models. *PLoS One* 2020;15(7):e0236238 [FREE Full text] [doi: [10.1371/journal.pone.0236238](https://doi.org/10.1371/journal.pone.0236238)] [Medline: [32722716](https://pubmed.ncbi.nlm.nih.gov/32722716/)]
23. Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, et al. Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR Public Health Surveill* 2020 May 13;6(2):e19115 [FREE Full text] [doi: [10.2196/19115](https://doi.org/10.2196/19115)] [Medline: [32391801](https://pubmed.ncbi.nlm.nih.gov/32391801/)]
24. Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*. 2nd edition. Melbourne, Australia: OTexts; 2018.
25. Armstrong JS. Evaluating forecasting methods. In: Armstrong JS, editor. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Boston, MA: Springer; 2001:443-472.
26. Armstrong J, Collopy F. Error measures for generalizing about forecasting methods: Empirical comparisons. *Int J Forecast* 1992 Jun;8(1):69-80. [doi: [10.1016/0169-2070\(92\)90008-w](https://doi.org/10.1016/0169-2070(92)90008-w)]
27. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest* 2007 Nov;8(2):53-96. [doi: [10.1111/j.1539-6053.2008.00033.x](https://doi.org/10.1111/j.1539-6053.2008.00033.x)] [Medline: [26161749](https://pubmed.ncbi.nlm.nih.gov/26161749/)]
28. Wallman KK. Enhancing statistical literacy: Enriching our society. *J Am Stat Assoc* 1993 Mar;88(421):1. [doi: [10.2307/2290686](https://doi.org/10.2307/2290686)]
29. Amidon TR, Nielsen AC, Pflugfelder EH, Richards DP, Stephens SH. Visual risk literacy in “flatten the curve” COVID-19 visualizations. *J Bus Tech Commun* 2020 Oct 01;35(1):101-109. [doi: [10.1177/1050651920963439](https://doi.org/10.1177/1050651920963439)]
30. Heesterbeek H, Anderson RM, Andreasen V, Bansal S, De Angelis D, Dye C, Isaac Newton Institute IDD Collaboration. Modeling infectious disease dynamics in the complex landscape of global health. *Science* 2015 Mar 13;347(6227):aaa4339 [FREE Full text] [doi: [10.1126/science.aaa4339](https://doi.org/10.1126/science.aaa4339)] [Medline: [25766240](https://pubmed.ncbi.nlm.nih.gov/25766240/)]

31. Tizzoni M, Bajardi P, Poletto C, Ramasco JJ, Balcan D, Gonçalves B, et al. Real-time numerical forecast of global epidemic spreading: Case study of 2009 A/H1N1pdm. *BMC Med* 2012 Dec 13;10:165 [FREE Full text] [doi: [10.1186/1741-7015-10-165](https://doi.org/10.1186/1741-7015-10-165)] [Medline: [23237460](https://pubmed.ncbi.nlm.nih.gov/23237460/)]
32. Reich NG, Lauer SA, Sakrejda K, Iamsirithaworn S, Hinjoy S, Suangtho P, et al. Challenges in real-time prediction of infectious disease: A case study of dengue in Thailand. *PLoS Negl Trop Dis* 2016 Jun;10(6):e0004761 [FREE Full text] [doi: [10.1371/journal.pntd.0004761](https://doi.org/10.1371/journal.pntd.0004761)] [Medline: [27304062](https://pubmed.ncbi.nlm.nih.gov/27304062/)]
33. Nuzzo RL. The box plots alternative for visualizing quantitative data. *PM R* 2016 Mar;8(3):268-272. [doi: [10.1016/j.pmrj.2016.02.001](https://doi.org/10.1016/j.pmrj.2016.02.001)] [Medline: [26892802](https://pubmed.ncbi.nlm.nih.gov/26892802/)]
34. Chambers JM, Cleveland WS, Kleiner B, Tukey PA. *Graphical Methods for Data Analysis*. London, UK: Chapman and Hall/Cole Publishing Co; 1983.
35. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. 2nd edition. Cham, Switzerland: Springer; 2016.
36. Krzywinski M, Altman N. Visualizing samples with box plots. *Nat Methods* 2014 Feb;11(2):119-120. [doi: [10.1038/nmeth.2813](https://doi.org/10.1038/nmeth.2813)] [Medline: [24645192](https://pubmed.ncbi.nlm.nih.gov/24645192/)]
37. McGill R, Tukey JW, Larsen WA. Variations of box plots. *Am Stat* 1978 Feb;32(1):12. [doi: [10.2307/2683468](https://doi.org/10.2307/2683468)]
38. Altman DG, Bland JM. How to obtain the P value from a confidence interval. *BMJ* 2011;343:d2304. [doi: [10.1136/bmj.d2304](https://doi.org/10.1136/bmj.d2304)] [Medline: [22803193](https://pubmed.ncbi.nlm.nih.gov/22803193/)]
39. Cumming G, Finch S. Inference by eye: Confidence intervals and how to read pictures of data. *Am Psychol* 2005;60(2):170-180. [doi: [10.1037/0003-066x.60.2.170](https://doi.org/10.1037/0003-066x.60.2.170)]
40. Cumming G. Inference by eye: Reading the overlap of independent confidence intervals. *Stat Med* 2009 Jan 30;28(2):205-220. [doi: [10.1002/sim.3471](https://doi.org/10.1002/sim.3471)] [Medline: [18991332](https://pubmed.ncbi.nlm.nih.gov/18991332/)]
41. Cumming G, Finch S. A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions. *Educ Psychol Meas* 2001;61(4):532-574. [doi: [10.1177/0013164401614002](https://doi.org/10.1177/0013164401614002)]
42. Divine GW, Norton HJ, Barón AE, Juárez-Colunga E. The Wilcoxon–Mann–Whitney procedure fails as a test of medians. *Am Stat* 2018 Mar 15;72(3):278-286. [doi: [10.1080/00031305.2017.1305291](https://doi.org/10.1080/00031305.2017.1305291)]
43. Brown GW, Mood AM. On median tests for linear hypotheses. In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press; 1951 Presented at: Second Berkeley Symposium on Mathematical Statistics and Probability; July 31-August 12, 1950; Berkeley, CA p. 159-166 URL: <https://projecteuclid.org/ebooks/bsmsp/Proceedings%20of%20the%20Second%20Berkeley%20Symposium%20on%20Mathematical%20Statistics%20and%20Probability/chapter/On%20Median%20Tests%20for%20Linear%20Hypotheses/bsmsp/1200500226>
44. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. *Int J Forecast* 2006 Oct;22(4):679-688. [doi: [10.1016/j.ijforecast.2006.03.001](https://doi.org/10.1016/j.ijforecast.2006.03.001)]
45. Shcherbakov MV, Brebels A, Shcherbakova NL, Tyukov AP, Janovsky TA, Kamaev V. A survey of forecast error measures. *World Appl Sci J* 2013;24:171-176 [FREE Full text] [doi: [10.5829/idosi.wasj.2013.24.itmies.80032](https://doi.org/10.5829/idosi.wasj.2013.24.itmies.80032)]
46. Lynch CJ, Gore R. Short-range early phase COVID-19 forecasting R-project and data, V2. Mendeley Data. 2020 Dec 15. URL: <https://data.mendeley.com/datasets/cytrb8p42g/2> [accessed 2021-03-02]
47. Lynch CJ, Gore R. Application of one-, three-, and seven-day forecasts during early onset on the COVID-19 epidemic dataset using moving average, autoregressive, autoregressive moving average, autoregressive integrated moving average, and naïve forecasting methods. *Data Brief* 2021 Apr;35:106759 [FREE Full text] [doi: [10.1016/j.dib.2021.106759](https://doi.org/10.1016/j.dib.2021.106759)] [Medline: [33521186](https://pubmed.ncbi.nlm.nih.gov/33521186/)]
48. Lynch CJ, Diallo SY, Kavak H, Padilla JJ. A content analysis-based approach to explore simulation verification and identify its current challenges. *PLoS One* 2020;15(5):e0232929 [FREE Full text] [doi: [10.1371/journal.pone.0232929](https://doi.org/10.1371/journal.pone.0232929)] [Medline: [32401795](https://pubmed.ncbi.nlm.nih.gov/32401795/)]
49. Wang G, Gu Z, Li X, Yu S, Kim M, Wang Y, et al. Comparing and integrating US COVID-19 data from multiple sources with anomaly detection and repairing. *arXiv*. Preprint posted online on November 28, 2020 [FREE Full text]
50. Gore R, Lynch CJ. Virginia County COVID-19 Daily Case Total Forecaster. Suffolk, VA: Virginia Modeling, Analysis and Simulation Center (VMASC), Old Dominion University; 2020. URL: <https://vmasc.shinyapps.io/va-county-covid-forecast/> [accessed 2021-03-02]
51. Sha D, Miao X, Lan H, Stewart K, Ruan S, Tian Y, et al. Spatiotemporal analysis of medical resource deficiencies in the US under COVID-19 pandemic. *PLoS One* 2020;15(10):e0240348 [FREE Full text] [doi: [10.1371/journal.pone.0240348](https://doi.org/10.1371/journal.pone.0240348)] [Medline: [33052956](https://pubmed.ncbi.nlm.nih.gov/33052956/)]
52. Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd SC, 2009 H1N1 Surveillance Group. Improving the evidence base for decision making during a pandemic: The example of 2009 influenza A/H1N1. *Biosecur Bioterror* 2011 Jun;9(2):89-115 [FREE Full text] [doi: [10.1089/bsp.2011.0007](https://doi.org/10.1089/bsp.2011.0007)] [Medline: [21612363](https://pubmed.ncbi.nlm.nih.gov/21612363/)]
53. Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *J Med Virol* 2020 Jun;92(6):645-659 [FREE Full text] [doi: [10.1002/jmv.25750](https://doi.org/10.1002/jmv.25750)] [Medline: [32141624](https://pubmed.ncbi.nlm.nih.gov/32141624/)]
54. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 2020 Apr 24;368(6489):395-400 [FREE Full text] [doi: [10.1126/science.aba9757](https://doi.org/10.1126/science.aba9757)] [Medline: [32144116](https://pubmed.ncbi.nlm.nih.gov/32144116/)]

55. Anastassopoulou C, Russo L, Tsakris A, Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One* 2020;15(3):e0230405 [FREE Full text] [doi: [10.1371/journal.pone.0230405](https://doi.org/10.1371/journal.pone.0230405)] [Medline: [32231374](https://pubmed.ncbi.nlm.nih.gov/32231374/)]
56. Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: A descriptive and modelling study. *Lancet Infect Dis* 2020 Jul;20(7):793-802 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30230-9](https://doi.org/10.1016/S1473-3099(20)30230-9)] [Medline: [32247326](https://pubmed.ncbi.nlm.nih.gov/32247326/)]
57. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: A model-based analysis. *Lancet Infect Dis* 2020 Jun;20(6):669-677 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7)] [Medline: [32240634](https://pubmed.ncbi.nlm.nih.gov/32240634/)]
58. Cobey S. Modeling infectious disease dynamics. *Science* 2020 May 15;368(6492):713-714. [doi: [10.1126/science.abb5659](https://doi.org/10.1126/science.abb5659)] [Medline: [32332062](https://pubmed.ncbi.nlm.nih.gov/32332062/)]
59. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci U S A* 2009 Mar 03;106(9):3243-3248 [FREE Full text] [doi: [10.1073/pnas.0806852106](https://doi.org/10.1073/pnas.0806852106)] [Medline: [19204283](https://pubmed.ncbi.nlm.nih.gov/19204283/)]
60. Ranscombe P. Rural areas at risk during COVID-19 pandemic. *Lancet Infect Dis* 2020 May;20(5):545 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30301-7](https://doi.org/10.1016/S1473-3099(20)30301-7)] [Medline: [32311327](https://pubmed.ncbi.nlm.nih.gov/32311327/)]
61. Yang W, Kandula S, Huynh M, Greene S, Van Wye G, Li W, et al. Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: A model-based analysis. *Lancet Infect Dis* 2021 Feb;21(2):203-212 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30769-6](https://doi.org/10.1016/S1473-3099(20)30769-6)] [Medline: [33091374](https://pubmed.ncbi.nlm.nih.gov/33091374/)]
62. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Kalhori SRN. Predicting COVID-19 incidence through analysis of Google Trends data in Iran: Data mining and deep learning pilot study. *JMIR Public Health Surveill* 2020 Apr 14;6(2):e18828 [FREE Full text] [doi: [10.2196/18828](https://doi.org/10.2196/18828)] [Medline: [32234709](https://pubmed.ncbi.nlm.nih.gov/32234709/)]
63. Northam RS, Thomasson K. Executive Order Number Fifty-One: Declaration of a State of Emergency Due to Novel Coronavirus (COVID-19). Richmond, VA: Commonwealth of Virginia, Office of the Governor; 2020 Mar 12. URL: <https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/2020-03/VA%20Governor%20Northam%203%2012%202020%20Declaration-of-a-State-of-Emergency-Due-to-Novel-Coronavirus-%28COVID-19%29.pdf> [accessed 2021-03-02]

Abbreviations

- AR:** autoregressive
ARIMA: autoregressive integrated moving average
ARMA: autoregressive moving average
HW: Holt-Winters
MA: moving average
MdAE: median absolute error
MdAPE: median absolute percentage error

Edited by G Eysenbach, C Basch; submitted 10.10.20; peer-reviewed by D Valdez, Z Ren; comments to author 27.10.20; revised version received 07.12.20; accepted 22.02.21; published 23.03.21

Please cite as:

Lynch CJ, Gore R

Short-Range Forecasting of COVID-19 During Early Onset at County, Health District, and State Geographic Levels Using Seven Methods: Comparative Forecasting Study

J Med Internet Res 2021;23(3):e24925

URL: <https://www.jmir.org/2021/3/e24925>

doi: [10.2196/24925](https://doi.org/10.2196/24925)

PMID: [33621186](https://pubmed.ncbi.nlm.nih.gov/33621186/)

©Christopher J Lynch, Ross Gore. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 23.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.