

DOI: https://doi.org/10.48009/2_iis_2021_122-135

Analyzing COVID-19 time-series data and government policy responses for forecasting deaths

Joshua Hammond, *University of St. Thomas*, hamm9537@stthomas.edu

Melinda Caouette, *University of St. Thomas*, caou1088@stthomas.edu

Manjeet Rege, *University of St. Thomas*, rege@stthomas.edu

Abstract

Governments around the world have implemented various public health and policy countermeasures intended to slow the continued outbreak of COVID-19 disease in 2020 and into 2021. How effective are these countermeasures, and is there a causal relationship between government policy response, confirmed case counts, and deaths that would allow for better forecasting? This study aims to identify relationships between various indexes of government policy, confirmed case counts, and confirmed deaths and quantify their predictive power using vector autoregression (VAR) methods and SAS/ETS[®] software. Our findings show that multivariate vector autoregression and vector error correction models generally outperform univariate models for most countries. We did not find substantial evidence that including government policy response metrics in the models improves forecasting accuracy.

Keywords: COVID-19, Time-Series Forecasting, Vector Autoregression, Granger Causality

Introduction

In December 2019 the first cases of COVID-19 disease caused by a new virus known as SARS-CoV-2 emerged in Wuhan China. This new virus spread rapidly worldwide with an alarming fatality rate. The World Health Organization (WHO) declared the outbreak a global pandemic on March 11, 2020. As of May 2021, over 180 countries have reported cases of COVID-19 and governments are responding with various policy responses to slow and control the spread of the virus. Countermeasures include stay-at-home orders; school closures; work closures; restrictions on gatherings; restrictions on internal and international travels; testing policy; contact tracing; face covering policy; and vaccine policy with the primary intent of reinforcing physical distancing between people to control viral spread.

The stringency and scope of policy responses varies between different countries, and even countries that have similar policies can have a wide range of outcomes. Some countries like Taiwan and New Zealand have kept confirmed case counts and deaths down; whereas other countries like Brazil continue to suffer from high death rates.

In this study, we explore the causal relationship between government response policy, confirmed case counts, and deaths and compare the forecasting accuracy for confirmed deaths using vector autoregression (VAR) and vector error correction (VEC) models.

Vector autoregression is a powerful tool and has been used extensively in economics for forecasting and evaluating the economic impacts of monetary policy as championed by Sims (Sims, 1972; Sims, 1980; Christiano 2012). VAR models have been gaining broader acceptance in fields outside of economics including epidemiology. We feel that the application of VAR models has wide-ranging applicability for the readers of *Issues in Information Systems* – particularly in the modeling and analysis of policy responses – which can be extended to many fields in the social sciences.

Literature Review

Stewart et al. (2020) addressed a problem very similar to this paper - *can we use traditional univariate time-series forecasting methods to predict the number of positive COVID-19 cases in New York city over a fixed period?* (Stewart et al., 2020) They found that a simple 7-day moving average forecast was the most accurate compared with exponential smoothing and polynomial time-series approaches.

Nguyen et al. (2021) use the Vector Error-Correction Model (VECM) to predict the number of hospital beds treating COVID-19 patients using multi-variate data. They used both the positive case counts for a specific geographical region as well as the daily number of beds occupied by COVID-19 patients in hospitals that serve that region and found encouraging results.

Research by Hoang et al. (2016) attempts to use Granger causality to identify key climactic factors which cause Dengue Fever in provinces in Vietnam. Their findings showed that temperature and hours of sunshine strongly influenced the occurrence of Dengue Fever in specific provinces and is a strong example of how Granger causality can be used in epidemiology to identify variables that influence the spread of a disease.

Serkan et al. (2019) compare various methods of multivariate time-series forecasting, including newer deep learning methods, and compare to a Vector Auto-Regression (VAR) model. Root Mean Square Error of the models was used to evaluate the prediction accuracy. They found that the deep learning methods outperformed the VAR model highlighting the fact that in a VAR cannot model non-linear relationships as well as an Artificial Neural Network.

Cai & Chen (2016) use Granger causality to identify the predictive power linking the income growth of urban residents to automobile sales. They were able to prove that with a lag of 4 years income growth is predictive of overall automobile sales volume. This provided a clear exposition of how to use Granger causality with econometric time-series forecasting.

Research Methodology

The goal of this research is to perform data analysis on the impact of government policies on the number of COVID-19 cases and deaths. Using time-series data analysis will help generate more accurate forecasts and allow governments to make more informed decisions when deciding whether to tighten or relax policies.

The research hypotheses to be tested are as follows:

H₁: More stringent government public health policy measures have some predictive power – not necessarily causal – which can improve forecast accuracy of COVID-19 deaths.

H₂: COVID-19 deaths can be forecast accurately using vector autoregression and vector error correction models.

To test these hypotheses, we trained vector autoregression and vector error correction models on one year of data running from 1/1/2020 to 12/31/2020. Models were evaluated by calculating the root mean square error (RMSE) of the residuals (forecasted deaths using the model) when compared to the actual confirmed deaths for 52 days running from 1/1/2021 to 2/21/2021. This date range was chosen because it was contemporaneous with the writing and submission of this paper highlighting our interest in near real-time forecasting.

Data and Preparation

The data set used for this analysis was obtained from the Oxford University COVID-19 Government Response Tracker (OxCGRT) and is publicly available at: <https://github.com/OxCGRT/covid-policy-tracker> (Hale et al., 2020). The OxCGRT systematically collects 19 indicators of government responses including containment and closure policies, economic policies, and health system policies for over 180 countries and many sub-national regions. Most indicators are represented on an ordinal scale to show the strictness of each policy. The data from the indicators is aggregated into the government response index, containment and health index, economic support index, and stringency index to reflect the composite effect of multiple indicators. In addition to the nineteen individual indicators and four indexes, the data set contains time-series data for the number of confirmed cases and confirmed deaths for each of the over 180 countries from 2020-01-01 to 2021-03-01. The original data set, downloaded from GitHub in March 2021, had 49 columns and 120,268 rows. The data set is updated frequently, and the layout and structure have changed since that time. A detailed working paper is available which explains the indexes and how they are calculated (Hale et al., 2021).

Data Preparation

Data preparation and cleaning were performed using a Jupyter notebook. All the rows that correspond to the period after 2021-02-21 were removed because some countries had a lag in reporting COVID-19 data. Missing values for “ConfirmedCases” and “ConfirmedDeaths” were populated with zero because these columns were left as blank until the first case was reported. Other missing data were imputed with the mean value. Unused columns like “CountryCode” and various response indexes for display were removed to clean up the data set. Long column names were truncated, and white space was removed from column names to meet the SAS naming convention. The data set has 40 columns and 113,305 rows.

Table 1 below is the data dictionary for the data set including an example row. Detailed explanations of how the indexes were collected and the significance of the scales is in the OxCGRT documentation on GitHub: <https://github.com/OxCGRT/covid-policy-tracker/tree/master/documentation>.

Our modified and cleaned data set is available for review in full on Git-Hub: https://github.com/joshhammond/IIS_COVID19_Data.

Table 1: Data Dictionary

Name	Description	Data Type	Example
CountryName	Country Name	String	United States
RegionName	Sub-National Regions (State, etc.)	String	
Jurisdiction	NAT_TOTAL or STATE_TOTAL	String	NAT_TOTAL
Date	Reported date	Date	3/21/2020
C1_SchoolClosing	Record closings of schools and universities; ordinal scale of 0 to 3	Integer	3
C2_WorkspaceClosing	Record closings of workspace; ordinal scale of 0 to 3	Integer	3
C3_CancelPublicEvents	Record cancelling public events; ordinal scale of 0 to 2	Integer	2
C4_RestrictionsGatherings	Record limits on gatherings; ordinal scale of 0 to 4	Integer	4
C5_ClosePublicTransport	Record closings of public transport; ordinal scale of 0 to 2	Integer	1
C6_StayAtHome	Record orders to 'shelter-in-place'; ordinal scale of 0 to 3	Integer	2
C7_RestrictionsIntMvt	Record restrictions on internal movement between cities/regions; ordinal scale of 0 to 2	Integer	2
C8_InternationalTravel	Record restrictions on international travel; ordinal scale of 0 to 4	Integer	3
E1_Income support	Record if cash payments to people who lose jobs; ordinal scale of 0 to 2	Integer	0
E2_Debt/contract relief	Record if freezing financial obligations for households; ordinal scale of 0 to 2	Integer	0
E3_Fiscal measures	Announced economic stimulus spending	Float	0
E4_International support	Announced offers of aid spending to other countries	Float	0
H1_PublicInfo	Record of public info campaigns; ordinal scale of 0 to 2	Integer	2
H2_TestingPolicy	Record who has access to testing; ordinal scale of 0 to 3	Integer	3
H3_Contact Tracing	Record gov policy on contact tracing after a positive diagnosis; ordinal scale of 0 to 2	Integer	1
H4_EmergencyHealthcareInvest	Announced short term spending on healthcare	Float	0
H5_InvestmentInVaccines	Announced spending on vaccine development	Float	0
H6_FacialCoverings	Record policies on facial coverings outside home; ordinal scale of 0 to 4	Integer	1
H7_VaccinationPolicy	Record policies on vaccine delivery for diff groups; ordinal scale of 0 to 5	Integer	0
C1/C2/C3/C4/C5/C6/C7 E1/H1/H6/H7 Flags	See OxCGRT documentation for details	Binary Flag	0
ConfirmedCases	Cumulative confirmed cases	Integer	26025
ConfirmedDeaths	Cumulative confirmed deaths	Integer	475
StringencyIndex	C1+C2+C3+C4+C5+C6+C7+C8+H1	Float	72.69
GovernmentResponseIndex	C1+C2+C3+C4+C5+C6+C7+C8+H1+H2+H3+H6+H7	Float	54.44
ContainmentHealthIndex	C1+C2+C3+C4+C5+C6+C7+C8+H1+H2+H3+H6+H7+E1+E2	Float	62.82
EconomicSupportIndex	E1+E2	Float	0

Analysis

Our hypothesis is that government public health policy measures have some predictive power – not necessarily causal – which can improve forecast accuracy of COVID-19 cases and deaths. Initial analysis looked at relationships between confirmed deaths, confirmed cases, and the OxCGRT Containment Health, Government Response, and Stringency composite indexes.

Analysis was performed using SAS/ETS® (Econometrics and Time-Series) software version 15.1 running on the SAS OnDemand for Academics system using Base SAS software version 9.04.01. The On-Demand for Academics (ODA) session runs inside a Jupyter notebook via the SASPy Python interface. All data preparation, outlined above, was performed using pandas and then transformed into a SAS data store for processing using the SAS/ETS® VARMAX procedure (SAS Institute Inc., 2018). Technical details on how SAS implements vector autoregression, vector error correction, Granger causality, the Dickey-Fuller test for stationarity, and Akaike’s information corrected criteria (AICC) are provided in the online documentation for the VARMAX procedure located here: <https://support.sas.com/documentation/onlinedoc/ets/151/varmax.pdf> (SAS Institute Inc., 2018). Figure 3 below outlines the sequence of steps taken in the analysis of the data.

Another package – such as [statsmodels](#) for Python – could be used instead of SAS if desired.

Vector Auto-Regression

An Auto-Regressive (AR) model attempts to forecast future data points of a time-series using its own past steps. These models are often used in econometric forecasting. Vector Auto-Regression (VAR) extends the concept of AR to include multiple time-series in the model. The relationships between each time-series can be modeled as dependent (endogenous) or independent (exogenous) variables.

The simplest VAR model is a bivariate model with two time-series, x and y , that we assume are interdependent. Extending the concept of auto-regression, instead of predicting the future values of a single time-series based on its own previous values, we use the past values of *both* time-series to predict the future values. This leads to two linear forecast equations as shown below in (1) and (2) for a model with two lags ($t-1$ and $t-2$). For a model with p lags and n time-series there are $p*n^2$ parameters b that need to be estimated along with the constant c_n (intercept) and error term series $e_{n,t}$. A discussion of the methods used to estimate these parameters is outside the scope of this paper; in practice ordinary least squares is often used.

$$x_t = c_1 + b_{1,1,1} * x_{t-1} + b_{1,1,2} * y_{t-1} + b_{2,1,1} * x_{t-2} + b_{2,1,2} * y_{t-2} + e_{1,t} \quad (1)$$

$$y_t = c_2 + b_{1,2,1} * x_{t-1} + b_{1,2,2} * y_{t-1} + b_{2,2,1} * x_{t-2} + b_{2,2,2} * y_{t-2} + e_{2,t} \quad (2)$$

An indexing method is necessary because the number of parameters is rather large. We propose the method shown in figure 1.

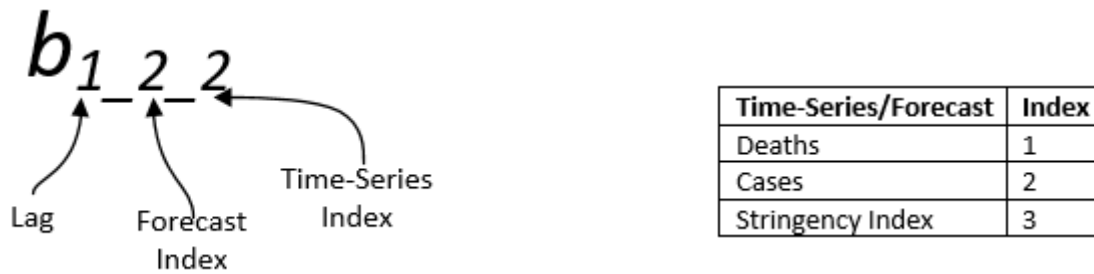


Figure 1: Parameter Indexing Method

In this paper we are concerned with three time-series: *Deaths*, *Cases*, and a *Response Policy* (Stringency Index). For a VAR model with p lags, the general form for our forecast equations is shown in equations (3), (4), and (5).

$$Deaths_t = c_1 + b_{1,1,1} * Deaths_{t-1} + b_{1,1,2} * Cases_{t-1} + b_{1,1,3} * StringencyIndex_{t-1} + \dots + b_{p,1,1} * Deaths_{t-p} + b_{p,1,2} * Cases_{t-p} + b_{p,1,3} * StringencyIndex_{t-p} + e_{1,t} \quad (3)$$

$$Cases_t = c_2 + b_{1,2,1} * Deaths_{t-1} + b_{1,2,2} * Cases_{t-1} + b_{1,2,3} * StringencyIndex_{t-1} + \dots + b_{p,2,1} * Deaths_{t-p} + b_{p,2,2} * Cases_{t-p} + b_{p,2,3} * StringencyIndex_{t-p} + e_{2,t} \quad (4)$$

$$StringencyIndex_t = c_3 + b_{1,3,1} * Deaths_{t-1} + b_{1,3,2} * Cases_{t-1} + b_{1,3,3} * StringencyIndex_{t-1} + \dots + b_{p,3,1} * Deaths_{t-p} + b_{p,3,2} * Cases_{t-p} + b_{p,3,3} * StringencyIndex_{t-p} + e_{3,t} \quad (5)$$

The SAS output for a VAR with 9 lags is shown below in figure 2. P-values are provided, and it is possible to exclude parameters that are above a specific p-value threshold by setting the parameter to zero to create a restricted model. We investigated both restricted and unrestricted models but found little improvement. All models used for the results are unrestricted models.

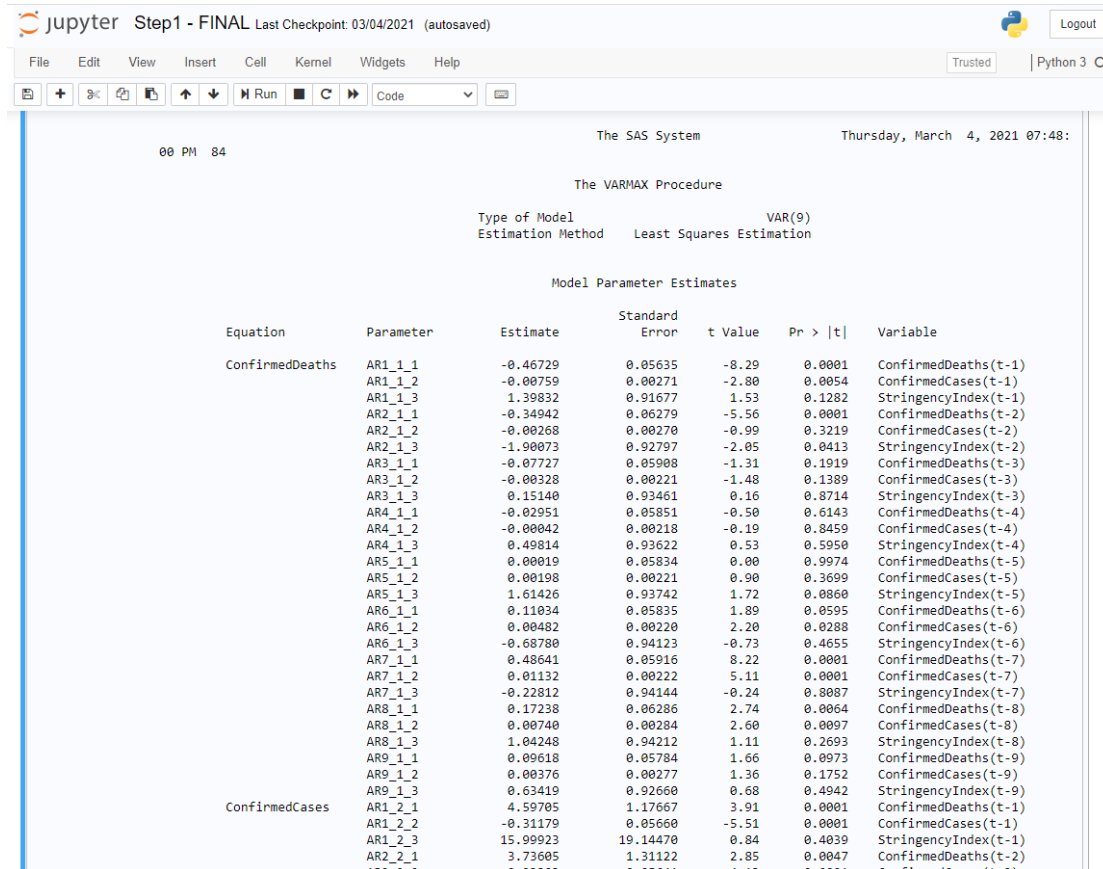


Figure 2: SAS VARMAX Output

The SAS code used to generate the output is listed below:

```
proc varmax data=work.covid_region_sds plots=forecasts;
id Date interval=day;
model ConfirmedDeaths ConfirmedCases StringencyIndex / noint p=9 dif=(ConfirmedDeaths(1,1)
ConfirmedCases(1,1) StringencyIndex(1));
output out=forecasts lead=52;
```

Testing for Stationarity

Data must be stationary (i.e., data has no trend, mean and variance are constant over time) to use most VAR models. We used the Dickey-Fuller test for stationarity to check if the time-series met this requirement. All the time-series in the data set required correction which was accomplished by calculating first and second differences directly in the model as highlighted in the code below:

```
proc varmax data=work.covid_region_sds plots=forecasts;
id Date interval=day;
model ConfirmedDeaths ConfirmedCases StringencyIndex / noint p=9 dif=(ConfirmedDeaths(1,1)
ConfirmedCases(1,1) StringencyIndex(1));
```

Lag Selection for Auto-Regression

The selection of the number of steps back (lags) to consider when calculating the next step in the model is important. In this analysis, the interval we are concerned with is one day. We used Akaike's information corrected criteria (AICC) looking out to a maximum of 21 days (lags). The model with the smallest AICC value was selected and this number of lags was used in the forecast model.

Granger Causality

We investigated the causal relationships between different variables by using the Granger Causality test. The simple idea behind this powerful concept is that given two time-series, a **policy stringency index** and **confirmed deaths**, you can state that the stringency index *granger-causes* confirmed deaths if including the past values of the stringency index improves the predictive accuracy of confirmed deaths in a model more than considering confirmed deaths alone. This fits well with VAR models as there can be multiple time-series influencing each other. The test states a null hypothesis that stringency index does not granger-cause confirmed deaths, with a significantly small p-value indicating that the null hypothesis can be rejected. Our threshold for p-value cut off was 0.05. Time-series that were not below the threshold were excluded from the model. The SAS code for calculating Granger causality with Stringency Index granger-causing deaths, and independently, Confirmed Cases granger-causing deaths is shown below:

```
proc varmax data=work.covid_region_sds;
id Date interval=day;
model ConfirmedDeaths ConfirmedCases StringencyIndex / noint p=9 dif=(ConfirmedDeaths(1,1)
ConfirmedCases(1,1) StringencyIndex(1));
causal group1=(ConfirmedDeaths) group2=(StringencyIndex);
causal group1=(ConfirmedDeaths) group2=(ConfirmedCases);
```

The p-values for most of the OxCGRT composite indexes (government response index, containment and health index, economic support index, and stringency index) were very high ($\gg 0.05$) indicating that we cannot reject the null hypothesis that these indexes *do not* granger-cause confirmed deaths. In a handful of countries, we did find that the stringency index granger-caused confirmed deaths and we included it in the model. The indexes are fairly static over time, and we hypothesize that they alone do not tell the entire story of how the policies are implemented in each country and how the public responds to them.

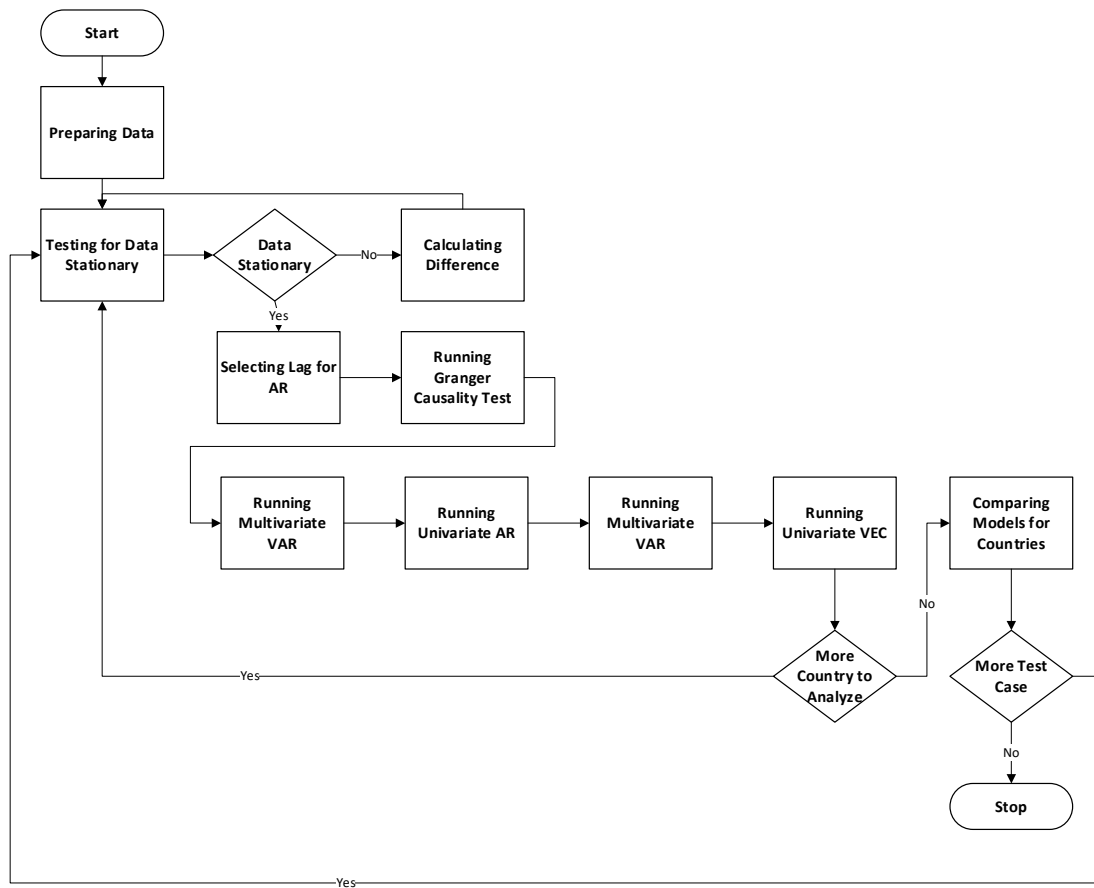


Figure 3: Analysis Methodology

Forecast Modeling

Four models were used for each country:

1. Multivariate Vector Auto Regression (VAR)
 - a. Consisting of at least confirmed cases and confirmed deaths, and including other time-series if granger-causation can be confirmed
2. Univariate Auto Regression (AR)
 - a. Using only confirmed deaths time-series
3. Multivariate Vector Error Correction (VEC)
 - a. Consisting of at least confirmed cases and confirmed deaths, and including other time-series if granger-causation can be confirmed
4. Univariate Vector Error Correction (VEC)
 - a. Using only confirmed deaths time-series

Models were trained on one year of data running from 1/1/2020 to 12/31/2020. Models were then evaluated by calculating the root mean square error (RMSE) as shown in equation (6) for 52 days running from 1/1/2021 to 2/21/2021.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \quad (6)$$

Fourteen countries from six continents were selected to maximize diversity when exploring the relationships between the government policy response index values and COVID-19 deaths. It is important to include countries that have successfully controlled the spread of COVID-19 as well as countries that are struggling with high confirmed cases and deaths. China and South Korea were chosen because of their early outbreak, lockdown, or implementation of government response policy. In contrast, Sweden was chosen because of its relaxed voluntary response at the start of the pandemic.

Results

Table 2 below summarizes the findings for forecasting deaths using our models during the 52-day period from 1/1/2021 to 2/21/2021. Lower RMSE values represent better forecast accuracy. The best model with the lowest RMSE for each country is highlighted. The multivariate VAC and VEC models outperformed the univariate models in all countries except China, India, the United Kingdom, United States, Australia, and New Zealand. China has a very flat curve and can be considered somewhat of an outlier. We interpret this to mean that for most countries the multi-variate models outperform simple univariate AR/VEC models. The OxCGRT Stringency Index was found to granger-cause confirmed deaths in the models for Brazil, Italy and Japan and was included in the VAR and VEC models, and we saw the RMSE improved for Italy and Brazil, but not Japan. The model with Stringency Index included for Japan did improve the error rate.

The “Error Rate” is calculated by comparing the forecasted cumulative deaths to the actual cumulative deaths on 2/21/2021. Most models were very accurate providing good predictive power out to 52 days.

Visualizations showing the forecast versus actual cumulative deaths with 95% confidence bounds are shown in figures 4-6. For each country two plots are shown, the first includes the entire timeline from 1/1/20 – 2/21/21, whereas the second shows the forecasted timeline from 1/1/21 – 2/21/21.

Table 2: Summary of 52-Day Forecast Results for (best model highlighted for each country)

Country	Multivariate VAR's RMSE	Univariate AR's RMSE	Multivariate VEC's RMSE	Univariate VEC's RMSE	Forecast on 02/21/2021	Cumulative Deaths on 02/21/2021	Error Rate
China	1.7	1.5	81.6	146.4	4,634	4,636	-0.04%
United States	6494.7	5985.6	9833.9	11733.3	500028	499009	0.20%
Australia	10.0	2.4	19.2	21.5	913	909	0.44%
Brazil	2149.5	7368.5	3793.4	7957.4	244955	246504	-0.63%
Japan	1036.7	988.9	266.7	718.4	7,421	7,485	-0.86%
Israel	117.6	605.2	191.3	634.6	5643	5577	1.18%
Sweden	670.3	840.8	606.6	952.1	12415	12649	-1.85%
India	6358.7	1300.8	2304.7	7262.0	159,386	156,385	1.92%
Italy	4245.2	3304.5	1183.4	3270.0	98,521	95,718	2.93%
New Zealand	2.3	0.4	2.6	0.5	25	26	-3.85%
South Africa	1768.3	2284.1	7695.7	2178.3	51,013	49,053	4.00%
United Kingdom	5977.5	5644.1	7092.9	8593.3	128803	120580	6.82%
Spain	2930.8	3788.8	3750.5	2956.9	61,260	67,101	-8.70%
South Korea	206.6	259.2	169.6	680.8	1,901	1,562	21.70%

Generalization

Results for the forecasts were accurate across a diverse set of countries. All four models we tested did not require any formal epidemiological theory of disease spread. Using the approach outlined in the paper we feel these models will generalize to other countries and strains of SARS-CoV-2.

Suggestions for Further Research

Exploring how other factors, such as vaccination rate, impact cumulative cases and deaths would be a logical next step. Looking at sub-national regions – such as individual US states – would also be interesting if granular policy data can be obtained.

Conclusion

In this study we attempted to identify relationships between various indexes of government policy, confirmed case count, and confirmed deaths and quantify their predictive power using Vector Auto Regression (VAR) models.

Both the timing and the stringency of these policies varied greatly across countries and even internal administrative areas *within* the same countries. The measures vary over time, with some countries tightening down early and relaxing over time, while others did not implement certain measures at all.

We found that the multivariate VAC and VEC models outperformed the univariate models in eight out of fourteen countries studied. We interpret this to mean that for most countries the multivariate models outperform simple univariate AR/VEC models. In a handful of countries including the stringency index also improved the forecast, however most of the other policy indexes did not improve the forecast accuracy. Therefore, we must reject hypothesis H1, but we do find evidence to support H2.

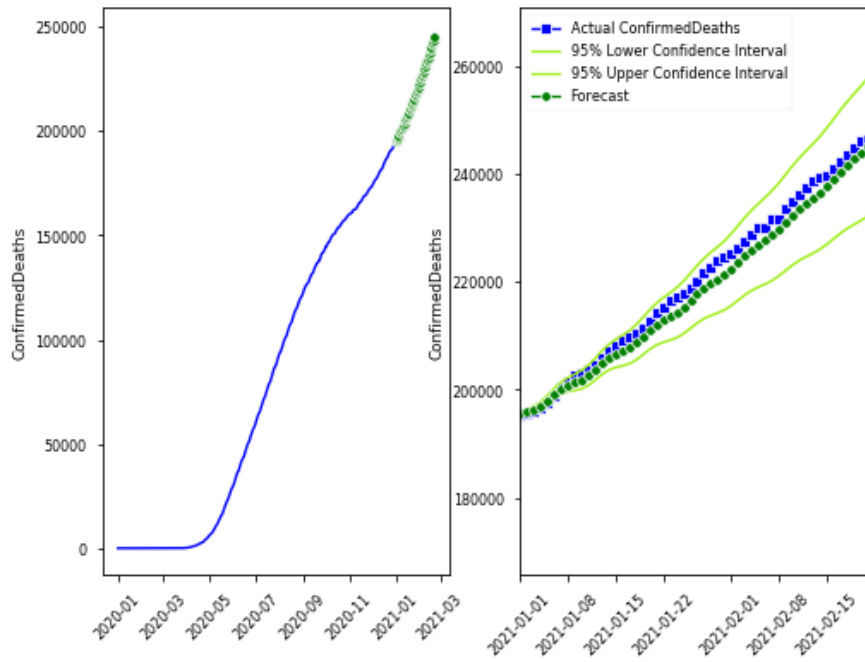


Figure 4: Brazil (Multivariate VAR w/ Stringency Index)

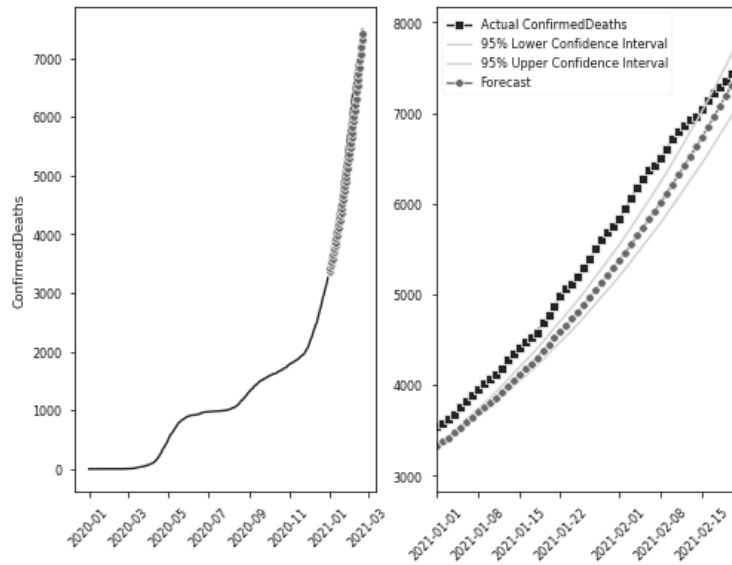


Figure 5: Japan (Multivariate VEC w/ Stringency Index)

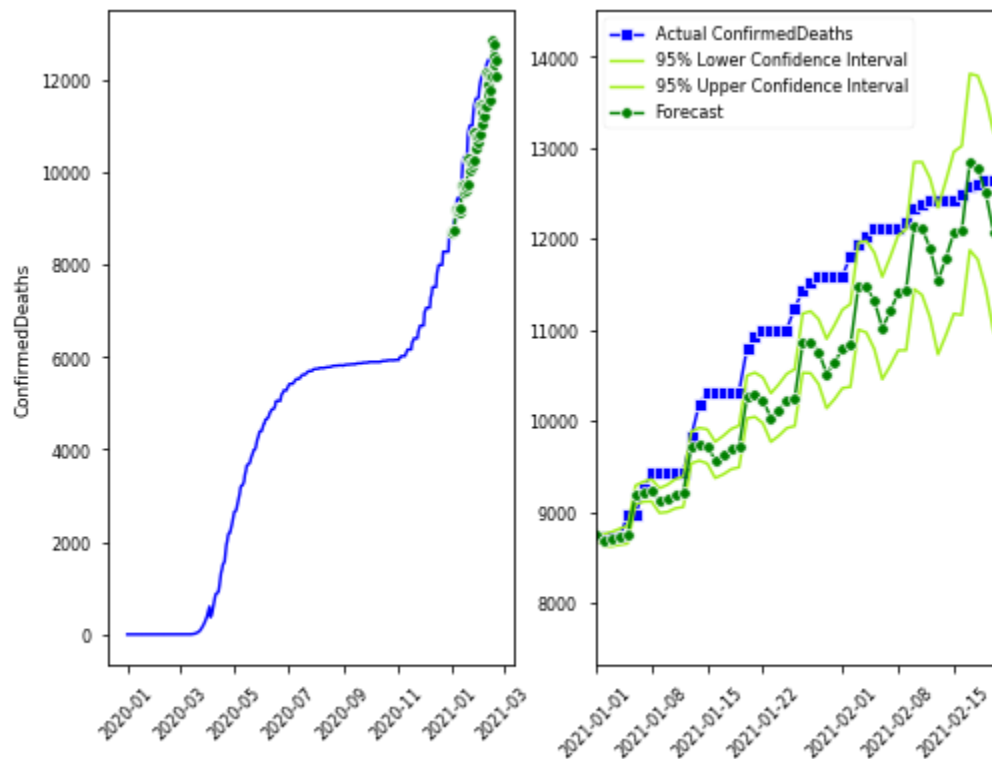


Figure 6: Sweden (Multivariate VEC)

References

- Cai, C., & Chen, X. (2016). Granger causality relationship between automobile sales volume and income of urban residents. *2016 International Conference on Logistics, Informatics and Service Sciences (LISS)*, Sydney, NSW, Australia, 1-4. <https://doi.org/10.1109/LISS.2016.7854534>
- Christiano, L. (2012). Christopher A. Sims and Vector Autoregressions. *The Scandinavian Journal of Economics*, 114(4), 1082-1104.
- Hale, T., Angrist, N., Cameron-Blake, E., Hallas, L., Kira, B., Majumdar, S., Petherick, A., Phillips, T., Tatlow, H., Webster, S. (2020). Oxford COVID-19 Government Response Tracker, Blavatnik School of Government.
- Hale, T, Anania, J., Angrist, N., Bobby, T., Cameron-Blake, E., Ellen, L., Goldszmidt, R., Hallas, L., Kira, B., Luciano, M., Majumdar, S., Nagesh, R., Petherick, A., Phillips, T., Tatlow, H., Webster, S., Wood, A., Zhang, Y. (2021, March 23). *Variation in Government Responses to COVID-19* (Working Paper version 11.0). Blavatnik School of Government, University of Oxford. Retrieved from <https://www.bsg.ox.ac.uk/research/publications/variation-government-responses-covid-19>
- Hoang, P., Zucker, J., Choisy, M., & Vinh, H. (2016). Causality analysis between climatic factors and dengue fever using the Granger causality. *2016 IEEE RIVF International Conference on Computing & Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*, Hanoi, Vietnam, 49-54. <https://doi.org/10.1109/RIVF.2016.7800268>

Nguyen, H. M., Turk, P., & McWilliams, A. (2021). A Multivariate Forecasting Model for the COVID-19 Hospital Census Based on Local Infection Incidence. medRxiv. <https://doi.org/10.1101/2021.02.18.21251243>

SAS Institute Inc. (2018). *SAS/ETS® 15.1 User's Guide*. Retrieved from <https://support.sas.com/documentation/onlinedoc/ets/151/varmax.pdf>

Serkan, Ö., Volkan, A., & Adnan, Y. (2019). Comparison of Predictive Models for Forecasting Time-series Data. In Proceedings of the 2019 3rd International Conference on Big Data Research (ICBDR 2019). Association for Computing Machinery, New York, NY, USA, 172–176. <https://doi.org/10.1145/3372454.3372482>

Sims, Christopher A. (1972). Money, Income, and Causality. *The American Economic Review*, 62(4), 540-552.

Sims, Christopher A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1), 1-48. doi:10.2307/1912017

Stewart, J., Igoche, D., & Davis, G. (2020). Accuracy of Simple Forecasting Methods in Predicting COVID-19 Incidence in New York City. *Issues in Information Systems*, 21(4), 187-192. https://doi.org/10.48009/4_iis_2020_187-192