Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/110116

Vector Error Correction Modeling of COVID-19 Infected Cases and Deaths

Rajarathinam Arunachalam and Tamilselvan Pakkirisamy*

Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli – 627012, Tamil Nadu State, India

Received: 7 Mar. 2021, Revised: 22 May 2021, Accepted: 28 May 2021
Published online: 1 Jan. 2022

Abstract: The main objective of this paper is to investigate the dynamic relationship between the COVID-19 infected cases and the number of deaths due to COVID-19 using the Johnsen-Fisher co-integration test, vector error correction model and Granger causality test. The daily COVID-19-infected new cases and daily deaths due to COVID-19 in the United States, Canada, Ukraine and India were collected from the website for the period from 01-04-2020 to 26-12-2020. The summary statistics revealed that the highest numbers of COVID-19-infected cases were registered in the United States, followed by India, Canada and Ukraine; the highest numbers of deaths due to COVID-19 were registered in the United States, followed by India, Ukraine and Canada. The death percentage is exceedingly high in Canada, followed by the United States, Ukraine and India. The Johnsen-Fisher co-integration test results reveal the existence of one co-integration equation. The vector error correction model and Granger causality test reveal that long-term and short-term causality exists between COVID-19 infection and death cases. The speed of adjustment is found to be 9.9%.

Keywords: Panel data, Panel Unit Root Test, Panel Co-integration Test, Vector Error Correction Model, Wald Test, Granger Causality Test

1 INTRODUCTION

1.1 Background of the study

Statistical modelling for the outbreak of COVID-19 has played an important role in understanding disease dynamics and taking necessary measures to control the spread of the virus, before antiviral drugs or vaccines exist. To evaluate the disease dynamics and to take necessary steps to arrange the number of hospitals, the level of protection, the rate of isolation of infected protection etc., it is paramount to develop statistical modelling approaches for such situations and to predict future COVID-19 cases and deaths.

1.2 Review of Literature

Sadiq et al.[1] used regression analysis to obtain confirmed positive cases of COVID-19 in India. They applied a 5% level of significance to obtain an ideal mathematical model to reach conclusions about positive cases of COVID-19 using R-Studio and Gretl to obtain a residual plot of the fitted model, in addition to actual confirmed, active, and recovered cases, and forecast of actual fitted deceased cases.

Livadiotis [2] performed a statistical analysis to understand the effect of environmental temperature on the exponential growth rate of COVID-19 cases in the US and Italian regions. Specifically, he analysed regional infected case datasets, derived the growth rates for regions characterized by a readable exponential growth phase in their evolution spread curve, plotted the results against the average environmental temperature within the same region, derived the relationship between temperature and growth rate, and evaluated the statistical confidence. The results showed a statistically significant negative correlation between the average environmental temperature and exponential growth rate of the number of infected cases.

* Corresponding author e-mail: rajarathinam@msuniv.ac.in

The critical temperature, which eliminates exponential growth and thus the spread of COVID-19 in the US region, was estimated to be TC = 86.1 $\pm$ 4.3 F0.

Chatterjee et al. [3] addressed the potential of data science to assess the risk factors correlated with COVID-19 after analysing existing datasets available at "ourworldindata.org (Oxford University database)" and newly simulated datasets, following the analysis of different univariate long short-term memory (LSTM) models for forecasting new cases and resulting deaths. The results showed that vanilla, stacked, and bidirectional LSTM models outperformed multilayer LSTM models. In addition, they discussed the findings related to the statistical analysis of simulated datasets. For correlation analysis, they included features, such as external temperature, rainfall, sunshine, population, infected cases, deaths, country, population, area, and population density for three months, i.e., January, February, and March, in 2020. For univariate time series forecasting using LSTM, they used datasets from 1 January 2020 to 22 April 2020.

Wang et al. [4] used a meta-analysis with a total of 269 patients from 47 studies. The mean age of operative patients with COVID-19 was 50.91 years, and 49% were female. A total of 28 patients died, with an overall mortality of 6%. All deceased patients had postoperative complications associated with surgery or COVID-19, including respiratory failure, acute respiratory distress syndrome (ARDS), shortness of breath, dyspnoea, fever, cough, fatigue or myalgia, cardiopulmonary system, shock/infection, acute kidney injury and severe lymphopenia. Patients who presented one or more of the symptoms of respiratory failure, ARDS, shortness of breath and dyspnoea after surgery were associated with significantly higher mortality (r = 0.891, $p < 0.001$), while fever, cough, fatigue or myalgia had marginally significant associations with postoperative mortality (r = 0.675, p = 0.023). Twenty studies reported information on medical staff infection: a total of 38 medical staff were infected, and medical staff who used biosafety level 3 (BSL-3) protective equipment did not become infected.

Wang et al. [5] stated that COVID-19 is spreading quickly all around the world, and publicly released data for 1212 COVID-19 patients in Henan, China, were analysed by employing various statistical and network analysis methods. They found that COVID-19 patients showed gender (55% vs. 45%) and age (81% aged between 21 and 60) differences, and possible causes were explored. The estimated average, mode and median incubation periods were 7.4, 4 and 7 days, respectively, and the incubation period of 92% of patients did not exceed 14 days. The epidemic in Henan experienced three stages that showed high correlation with the number of patients recently returned from Wuhan. Network analysis revealed that 208 cases were clustering infected, and various People's Hospitals were the main force in treating COVID-19.

Zuo et al. [6] provided a convenient method of data comparison that could be helpful for both governmental and private organizations. To date, facts and figures of the total confirmed cases, daily confirmed cases, total deaths, and daily deaths reported in Asian countries have been provided. Furthermore, a statistical model was recommended to provide the best description of the COVID-19 total death data in Asian countries.

1.3 Objectives of the present study

Based on the above information, the present study aims to assess the long-term and short-term relationships between the number of COVID-19 infection cases and the number of deaths due to COVID-19 in four countries, namely, the United States, Canada, India and the Ukraine, using daily infected cases and deaths for the period from 01-04-2020 to 26-12-2020.

1.4 Panel data model

Panel data contain observations of multiple phenomena collected over different time periods for the same group of individuals, units or entities. In short, econometric panel data are multidimensional data collected over a given period. A simple panel data regression model is specified as

$$Y_{it} = \alpha + \beta X_{it} + v_{it} \tag{1}$$

Here, Y is the dependent variable, X is the independent or explanatory variable,α and β are the intercept and slope, i stands for the i^{th} cross-sectional unit, t stands for the t^{th} time period, X is assumed to be nonstochastic, and the error term is assumed to follow the classical assumptions, namely, $E(V_{it}) = N(0, \sigma^2)$ In this study, i, that is, the number of cross-sections (districts), is 4 (i=1, 2, 3 and 4) and t=1, 2, 3, ..., 226. Detailed discussions of panel data modelling can be found in Baltagi [7], Gujarati et al. [8] , and Hsiao [9]. By combining time series of cross-sections of observations, panel data provide "more informative data, more variability, less collinearity among variables, more degrees of freedom and more efficiency" (Baltagi [7]).

2 MATERIALS AND METHODS

2.1 Materials

The daily COVID-19 infections and deaths from 01-04-2020 to 26-12-2020 were collected from the official website (https://www.wolfram.com/covid-19-resources) for the United States, Canada, India and the Ukraine. Several methodologies used to achieve the objectives of the present study are discussed in the methods section. EViews Ver.11 was used for the model estimation and it's parameters.

2.2 Methods

In this paper, panel regression tools, i.e., the panel unit root test (for testing stationarity, Maddala and Wu [10], Choi [11] and Hadri [12]) Fisher-Johansen co-integration test (test for the number of co-integrating relationships among the underlying variables), Vector Error Correction Model (VECM) (to determine long-term relationships among the variables), Wald test (to identify the short-term relationships among the variables) and Granger causality test (to detect the direction of causality), were employed.

2.3 Unit root tests

Unit roots in panel data can be tested using either the Levin-Lin-Chu (Levin, Lin and Chu [13]) test or the Hadri [12] Lagrange multiplier (LM) stationarity test. The null hypothesis is that the panels contain unit roots, and the alternative hypothesis is that the panels are stationary. In the results, if the p value is less than 0.05, then the null hypothesis can be rejected.

2.4 Panel co-integration test

The concept of co-integration was introduced to the econometric literature by Granger [14] and Engle and Granger [15]. If two integrated variables share a common stochastic trend such that a linear combination of these variables is stationary, they are called co-integrated.

Fisher [16] derives a combined test that uses the results of individual independent tests. Maddala and Wu [10] use Fisher's result to propose an alternative approach to test for co-integration in panel data by combining tests for individual cross-sections to obtain the test statistic for the full panel.

If π_i is the p-value from an individual co-integration test for cross-section i, then under the null hypothesis for the panel,

$$-2\sum_{i=1}^{N} log(\pi_i) \rightarrow \chi^2_{2N} \tag{2}$$

The χ^2 value is based on the MacKinnon-Haug-Michelis [17] p-values for Johansen's co-integration trace test and the maximum eigenvalue test.

2.5 Vector error correction model(Engle and Granger [15])

VECM is a time series model that can be used to directly estimate the level to which a variable can be brought back to the equilibrium condition after a shock to other variables. VECM is a useful technique for assessing the short-term and long-term dynamics of the variables being studied.

The conventional Error Correction Model (ECM) for co-integrated series is

$$\Delta Y_t = \beta_0 + \sum_{i=1}^{n} \beta_i \Delta Y_{t-1} + \sum_{i=0}^{n} \delta_i \Delta X_{t-i} + \phi Z_{t-1} + \mu_t \tag{3}$$

where Z, the ECT, is the OLS residuals from the following long-term co-integration regression:

$$Y_t = \beta_0 + \beta_1 X_1 + \varepsilon_t \tag{4}$$

defined as

$$Z_{t-1} = ECT_{t-1} = Y_{t-1} - \beta_0 - \beta_1 X_{t-1} \tag{5}$$

The error correction relates to the fact that the previous period deviation from the long-run equilibrium (the error) influences the short-run dynamics of the dependent variable. Thus, the coefficient of ECT,ϕ, is the rate of adjustment that measures the speed at which Y returns to equilibrium after a change in X: this value must be negative and significant.

2.6 Wald test (Wald [18])

The short-run causality is also tested using the Wald test, which computes a test statistic based on the unrestricted regression. The Wald statistic measures how close the unrestricted estimates come to satisfying the restrictions under the null hypothesis. If the restrictions are in fact true, then the unrestricted estimates should come close to satisfying the restrictions.

2.7 Testing for causality (Granger [19])

The causal relationship between two stationary series X_t and Y_t can be assessed based on the following bivariate autoregression:

$$X_t = \phi_0 + \sum_{k=1}^{p} \phi_k Y_{t-k} + \sum_{k=1}^{p} \phi_k X_{t-k} + \nu_t \tag{6}$$

and

$$Y_t = \alpha_0 + \sum_{k=1}^{p} \alpha_k Y_{t-k} + \sum_{k=1}^{p} \beta_k X_{t-k} + u_t \tag{7}$$

where p is a suitably chosen positive integer; α_k and β_k , k=0,1,2,3,...,p, are constants; υ_t and u_t are the usual disturbance terms with zero mean and finite variance. The null hypothesis that X_t does not Granger-cause Y_t is rejected if β_k, $k > 0$ in the first equation are jointly significantly different from zero according to a standard joint test (e.g., an F test). Similarly, Y_t Granger-causes X_t if the coefficients of ϕ_k, $k > 0$ in the second equation are jointly different from zero. A bidirectional causality (or feedback) relation exists if both β_k and ϕ_k , $k > 0$, are jointly different from zero.

3 RESULTS AND DISCUSSION

3.1 Summary Statistics

The data in Table 1 and Fig. 1 reveal that the highest number of COVID-19 infected new cases was registered in the United States (18038065), followed by India (10097669), Canada (517058) and Ukraine (1006982). The data in Table 2 and Fig.

Table 1: Summary statistics for new cases of COVID-19

LOCATION	Obs.	Sum	Mean	Std. Dev.
Canada	266	517058	1943.82	2010.78
India	266	10097669	37961.16	28807.43
Ukraine	266	1006982	3785.65	4423.64
United States	266	18038065	67812.27	57892.81
All	1064	29659774	27875.73	42265.24

2 reveal that the highest number of deaths due to COVID-19 was registered in the United States (317487), followed by India (146409), Ukraine (17518) and Canada (14343).

Fig. 3 reveals the percentage of deaths due to COVID-19: although the number of deaths due to COVID-19 is high in the United States and India, the death percentage is very low. This may be due to care taken by the respective countries to avoid deaths by taking precautionary measures, i.e., medical facilities, awareness programmes, etc. The lowest death rate is observed in India.

Table 2: Summary statistics for death due to COVID-19

LOCATION	Obs.	Sum	Mean	Std. Dev.
Canada	266	14343	53.92	54.10
India	266	146409	550.41	364.23
Ukraine	266	14518	65.86	75.36
United States	266	317487	1193.56	685.39
All	1064	495757	465.94	607.56

3.2 Panel unit root test

The panel unit root test is employed to assess whether the study variables are stationary; the results are presented in Tables 3, 4, 5 and 6. Both study variables have unit roots (non-stationary) at the level because the Levin, Lin & Chu t test statistics are non-significant but are stationary at the first difference because the Levin, Lin & Chu t test statistics values are highly significant.

Table 3: Characteristics of the Panel unit root test at the level for the COVID-19 infected cases

Method	Statistic	Prob.**	Cross - Sections	Obs.
Null: Unit root (assumes common unit root process)				
Levin, Lin Chu t*	0.83552	0.7983	4	1023
Null: Unit root (assumes individual unit root process)				
ADF - Fisher Chi-square	2.62642	0.9556	4	1023
PP - Fisher Chi-square	3.57328	0.8934	4	1060
** Probabilities for Fisher tests are computed using asymptotic Chi-Square distribution. All other tests are assuming asymptotic Normality.				

Table 4: Characteristics of the panel unit root test at first difference for deaths due to COVID-19

Method	Statistic	Prob.**	Cross - Sections	Obs.
Null: Unit root (assumes common unit root process)				
Levin, Lin Chu t*	-5.36261	0.0000	4	1018
Null: Unit root (assumes individual unit root process)				
ADF - Fisher Chi-square	42.6527	0.0000	4	1018
PP - Fisher Chi-square	454.954	0.0000	4	1056
** Probabilities for Fisher tests are computed using asymptotic Chi-Square distribution. All other tests are assuming asymptotic Normality.				

Table 5: Characteristics of the Panel unit root test at the level for the number of deaths due to COVID-19

Method	Statistic	Prob.**	Cross - Sections	Obs.
Null: Unit root (assumes common unit root process)				
Levin, Lin Chu t*	1.47842	0.9304	4	1027
Null: Unit root (assumes individual unit root process)				
ADF - Fisher Chi-square	2.84559	0.9437	4	1027
PP - Fisher Chi-square	24.5100	0.0019	4	1060
** Probabilities for Fisher tests are computed using asymptotic Chi-Square distribution. All other tests are assuming asymptotic Normality				

Here, all the study variables are integrated in the same order. Therefore, at the level, the variables are non-stationary, but after converting them to the first difference, they become stationary. When the variables are integrated of the same order, one can run the panel co-integration test.

Table 6: Characteristics of the panel unit root test at first difference for deaths due to COVID-19

Method	Statistic	Prob.**	Cross - Sections	Obs.
Null: Unit root (assumes common unit root process)				
Levin, Lin Chu t*	-18.9092	0.0000	4	1027
Null: Unit root (assumes individual unit root process)				
ADF - Fisher Chi-square	276.665	0.0000	4	1027
PP - Fisher Chi-square	399.569	0.0000	4	1056
** Probabilities for Fisher tests are computed using asymptotic Chi-Square distribution. All other tests are assuming asymptotic Normality.				

3.3 Johnson–Fisher co-integration test

Finally, the Johnsen-Fisher panel co-integration test was conducted, and the test results are presented in Table 7.

Table 7: Characteristics of the Johnsen-Fisher panel co-integration test

Hypothesized No. of CE(s)	Fisher Stat.* (from trace test)	Prob.	Fisher Stat.* (from max-eigen test)	Prob.
None	92.95	0.0000	98.56	0.0000
At most 1	10.89	0.2081	10.89	0.2081
* Probabilities are computed using asymptotic Chi-square distribution. *MacKinnon-Haug-Michelis (1999) p-values.				

The results presented in the above table reveal that the null hypothesis of at most one co-integration equation is not rejected, as Fisher's trace and the max-eigen test values are non-significant at the 5% level. That is, there is at most one co-integrating equation among the variables. The results show that a long-term equilibrium relationship exists between the variables. However, in the short term, there may be deviations from this equilibrium, and it is necessary to verify whether such disequilibrium converges to the long-run equilibrium. Thus, VECM is used to generate the short-term dynamics.

3.4 Vector error correction model

As the study variables are co-integrated, the VECM can be applied. Here, the number of deaths due to COVID-19 is considered the dependent variable(Y), and the number of new COVID-19 infections is the independent variable(X). The characteristics of the target model are presented in Tables 8 and 9.

$$\Delta Y_t = -0.0991^{**} ect_{t-1} + 0.1122^{**} \Delta Y_{t-2} - 0.0004 \Delta X_{t-1} - 0.0066^{**} \Delta X_{t-2} + 3.8991 \quad (8)$$

** indicates highly significant $(p < 0.000)$ The co-integration equation (long-run model) is defined as follows:

$$ect_{t-1} = 1.000000 Y_{t-1} - 0.011159 X_{t-1} - 153.1994 \quad (9)$$

Here, C(1), the coefficient of error correction, also called the speed of adjustment towards long-run equilibrium, is negative and significant, indicating long-run causality from the number of deaths due to COVID-19 to the number of COVID-19 infections. The speed of the adjustment is 9.9%. The intuition behind this finding is that if the number of deaths is greater than its long-run equilibrium during the COVID-19 pandemic, the negative coefficient pulls the number back down toward equilibrium. The model is also highly significant, and all the coefficients are highly significant, except C(4) and C(6). To investigate the short-run causality running between the number of COVID-19 infected new cases and the number of deaths due to COVID-19, Wald statistics were calculated for the null hypothesis H_0=C(2)=C(3)=0 (H_0: No influence in the short run), and the results are reported in Table 10. The Chi-square statistics are significant, indicating that C(2) and C(3) are not jointly zero and that there is short-run causality running from the independent variable to the dependent variable. Thus, the independent variable has long-run and short-run influences on the dependent variable.

Table 8: Characteristics of the fitted VECM model

Co-integrating Eq:	CointEq1	
Y(-1)	1.000000	
X(-1)	-0.011159	
	(0.00146)	
	[-7.63804]	
C	-153.1994	
Error Correction:	D(Y)	D(X)
CointEq1	-0.099113	-1.008671
	(0.01700)	(0.53033)
	[-5.83094]	[-1.90198]
D(Y(-1))	0.112220	10.28192
	(0.03494)	(1.09024)
	[3.21143]	[9.43086]
D(Y(-2))	-0.195255	5.497175
	(0.03479)	(1.08534)
	[-5.61290]	[5.06493]
D(X(-1))	-0.000376	-0.487527
	(0.00113)	(0.03532)
	[-0.33179]	[-13.8031]
D(X(-2))	-0.006639	-0.236023
	(0.00113)	(0.03524)
	[-5.87780]	[-6.69784]
C	3.899141	314.5788
	(6.18400)	(192.939)
	[0.63052]	[1.63045]
Standard errors in () & t-statistics in [ ]		

Table 9: Summary of Significant Statistics of VECM

	Coefficient	Std. Error	t-Statistic	Prob.
C(1) (ect_{t-1})	-0.099113	0.016998	-5.830937	0.0000
C(2) (ΔY_{t-1})	0.112220	0.034944	3.211428	0.0014
C(3) (ΔY_{t-2})	-0.195255	0.034787	-5.612896	0.0000
C(4) (ΔX_{t-1})	-0.000376	0.001132	-0.331792	0.7401
C(5) (ΔX_{t-2})	-0.006639	0.001129	-5.877795	0.0000
C(6) (β_0)	3.899141	6.183996	0.630521	0.5285
Root MSE	199.7094	R-squared		0.1687
Mean dependent var	2.4943	Adjusted R-squared		0.1648
S.D. dependent var	219.1476	S.E. of regression		200.2814
Akaike info criterion	13.4430	Sum squared resid		41957813
Schwarz criterion	13.4713	Log likelihood		-7065.0240
Hannan-Quinn criter.	13.4537	F-statistic		42.4662
Durbin-Watson stat	2.18449	Prob(F-statistic)		0.0000

Table 10: Characteristics of the Wald test

Test Statistic	Value	df	Probability
F-statistic	24.87356	(2, 1046)	0.0000
Chi-square	49.74711	2	0.0000
Restrictions are linear in coefficients			

3.5 Test of causality

To assess whether causal relationships exist among the variables and to determine the direction of the causality, the Granger test of causality was employed, and the results are presented in Table 11. The null hypothesis of no causality between the independent and dependent variables running in either direction is rejected; hence, bidirectional causality exists between the study variables.

Table 11: Characteristics of the pairwise Granger causality test between independent and dependent variables

Null Hypothesis	Obs.	F-Statistic	Prob.
X does not Granger-cause	1056	19.4581	5E-09
Y does not Granger-cause		29.9723	2E-13

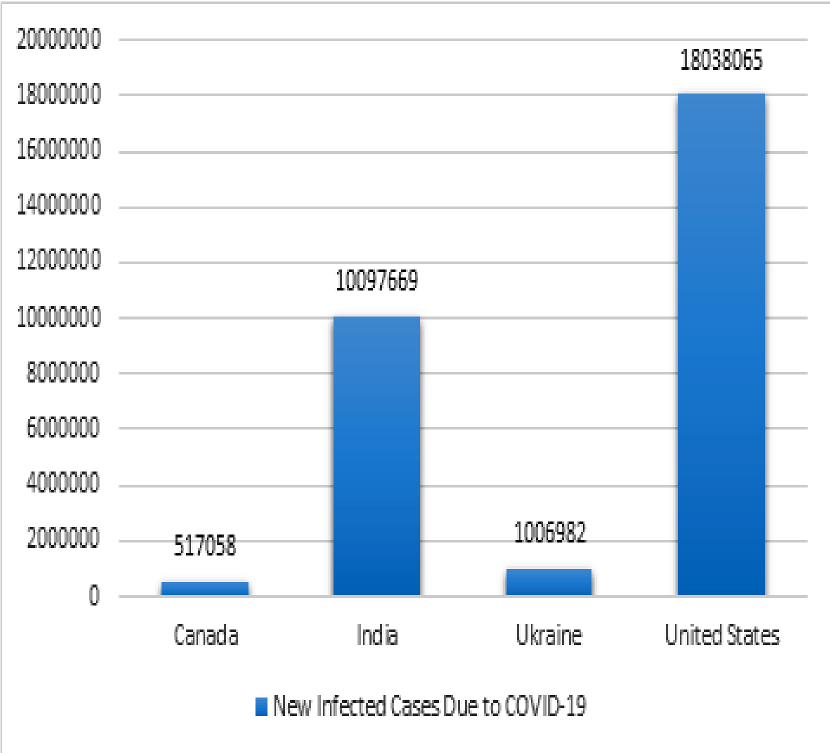

Fig. 1: Number of COVID-19 infections in all four countries

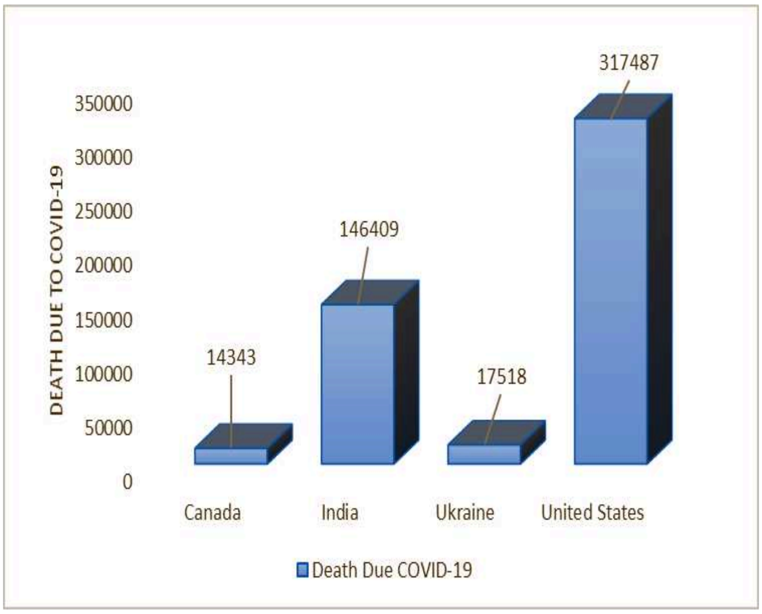

Fig. 2: Number of deaths due to COVID-19 in all four countries

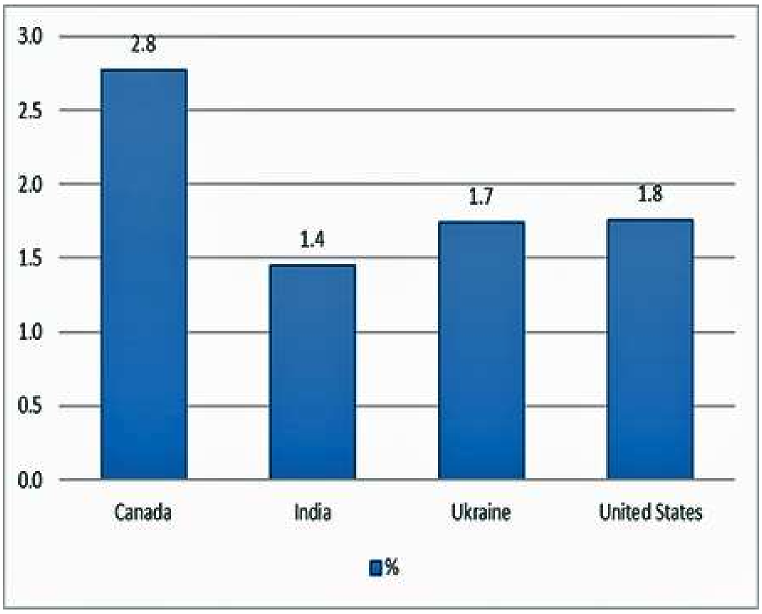

Fig. 3: Death Percentage due to COVID-19 in all four countries

4 CONCLUSIONS

The results of this paper reveal that the highest numbers of COVID-19 infection were registered in the United States, followed by India, Canada and Ukraine; the highest numbers of deaths due to COVID-19 were registered in the United States, followed by India, Ukraine and Canada. The death percentage is exceedingly high in Canada, followed by the United States, Ukraine and India. The Johnsen-Fisher co-integration test results reveal the existence of one co-integration equation. The VECM and Ganger causality test reveal that long-run and short-term causality exists between COVID-19 infected new cases and deaths due to COVID-19.The speed of adjustment is found to be 9.9%.

Acknowledgement

Both authors express their sincere thanks and gratitude to Dr. K. Senthamari Kannan, Senior Professor and Head, Department of Statistics, Manonmaniam Sundaranar University,Tirunelveli-627 012,Tamil Nadu State, India for providing necessary facilities and his constant encouragement during the preparation of this paper.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] I.A. Sadiq, S.K. Sharma and R. Ado, Statistical analysis of future prospective in COVID-19 cases in India, *International Journal of Engineering Research & Technology* **9(11)**,627-631 (2020).
[2] G. Livadiotis, Statistical analysis of the impact of environmental temperature on the exponential growth rate of cases COVID-19, *PLoS ONE*, https://doi.org/10.1371/journal.pone.0233875, **15(5)**: e023387 (2020).
[3] A. Chatterjee, M.K. Gerdes and S.G. Martinez, Statistical explorations and univariate time-series analysis on COVID-19 datasets to understatnd the trend of Disease spreading and death, *Sensors*, doi:10.3390/s20113089, **20**, 3089 (2020).
[4] K. Wang, C. Wu, J. Xu, B. Zhang, X. Zang, Z. Gao and Z. Xia, Factors affecting the mortality of patients with COVID-19 undergoing surgery and the safety of medical staff : A systematic review and meta-analysis. *E-Clinical Medicine*, http://doi.org/10.1016/j.eclinm.2020.100612, 29-30 (2020).
[5] P. Wang, J. Lu, Y. Jin, M. Zhu, L. Wang and S. Chen, Statistical and network analysis of 1212 COVID-19 patients in Henan, China, *International Journal of Infectious Diseases*, https://doi.org/10.1016/j.ijid.2020.04.051, **195**, 391 – 398,(2020).
[6] M. Zuo, S.K. Khosa, Z. Ahmad and Z. Almaspoor, Comparison of COVID-19 pandemic dynamics in Asian countries with statistical modeling, *Computational and mathematical methods in medicine*, Article ID 4296806, https://doi.org/10.1155/2020/4296806 (2020).
[7] B.H. Baltagi. *Econometric Analysis of Panel Data*, 3^{rd} Edition,West Sussex, England:John Wiley & Sons (2005).
[8] D.N. Gujarati, D.C. Porter and S. Gunasekar *Basic Econometrics*, 5^{th} Edition, New Delhi, NCT: Tata McGraw Hill Education, (2012).
[9] C. Hsiao. *Analysis of Panel Data*, 2^{nd} Edition, Cambridge University Press, Cambridge,MA (2003).
[10] G.S. Maddala and Shaowen Wu, Comparative Study of Unit Root Tests with Panel Data and A New Simple Test, *Oxford Bulletin of Economics and Statistics*, **61**,631–52 (1999).
[11] In Choi, Unit root tests for panel data, *Journal of international money and Finance*, **20(2)**,249-272 (2001).
[12] K. Hadri, Testing for stationarity in heterogeneous panel data, *The Econometrics Journal*, **3**, 148-161 (1999).
[13] A. Levin, C.F. Lin and C.S.J. Chu, Unit Root Tests in Panel Data:Asymptotic and Finite-Sample Properties, *Journal of Econometrics*, **108**, 1-24 (2002).
[14] C.W.J. Granger, Some Properties of Time Series Data and Their Use in Econometric Model Specification,*Journal of Econometrics*, 121-130 (1981).
[15] R.F. Engle and C.W.J. Granger, Co-Integration and Error Correction: Representation,Estimation and Testing, *Econometrica*, **55**,251-276 (1987).
[16] R.A. Fisher. *Statistical Methods for Research Workers*, 4^{th} Edition, Edinburgh: Oliver & Boyd (1932).
[17] J.G. MacKinnon, A.H. Alfred and M. Leo, Numerical Distribution Functions of Likelihood Ratio Tests for Co-integration, *Journal of Applied Econometrics*, **14**, 563-577 (1999).
[18] A. Wald, Tests of statistical hypotheses concerning several parameters when the number of observations is large, *Transactions of the American Mathematical Society*, **54**, 426-482 (1943).
[19] C.W.J. Granger, Investigating Causal Relations by Econometric Models and Cross Spectral Methods, *Econometrics*, **35**, 224-238 (1969).