The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2043-9377.htm

Hybrid grey exponential smoothing approach for predicting transmission dynamics of the COVID-19 outbreak in Sri Lanka

Received 7 June 2021
Revised 14 August 2021
28 September 2021
Accepted 12 November 2021

D.M.K.N. Seneviratna
Department of Interdisciplinary Studies, Faculty of Engineering, University of Ruhuna, Galle, Sri Lanka, and
R.M. Kapila Tharanga Rathnayaka
Department of Physical Sciences and Technology, Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka

Abstract

Purpose – The Coronavirus (COVID-19) is one of the major pandemic diseases caused by a newly discovered virus that has been directly affecting the human respiratory system. Because of the gradually increasing magnitude of the COVID-19 pandemic across the world, it has been sparking emergencies and critical issues in the healthcare systems around the world. However, predicting the exact amount of daily reported new COVID cases is the most serious issue faced by governments around the world today. So, the purpose of this current study is to propose a novel hybrid grey exponential smoothing model (HGESM) to predicting transmission dynamics of the COVID-19 outbreak properly.

Design/methodology/approach – As a result of the complications relates to the traditional time series approaches, the proposed HGESM model is well defined to handle exponential data patterns in multidisciplinary systems. The proposed methodology consists of two parts as double exponential smoothing and grey exponential smoothing modeling approach respectively. The empirical analysis of this study was carried out on the basis of the 3rd outbreak of Covid-19 cases in Sri Lanka, from 1st March 2021 to 15th June 2021. Out of the total 90 daily observations, the first 85% of daily confirmed cases were used during the training, and the remaining 15% of the sample.

Findings – The new proposed HGESM is highly accurate (less than 10%) with the lowest root mean square error values in one head forecasting. Moreover, mean absolute deviation accuracy testing results confirmed that the new proposed model has given more significant results than other time-series predictions with the limited samples.

Originality/value – The findings suggested that the new proposed HGESM is more suitable and effective for forecasting time series with the exponential trend in a short-term manner.

Keywords Coronavirus, COVID-19, Exponential smoothing, GM (1, 1) model, Grey system theory

Paper type Research paper

1. Introduction

The Coronavirus (COVID-19) is one of the major pandemic diseases caused by a newly discovered virus that has been directly affecting the human respiratory system (WHO, 2020). The origin of the coronavirus is somewhat confusing and controversial (Shao, 2020). However, some other scientists have argued that viruses are the first living organisms which have been originated from the association of organic compounds in the ocean floor with the addition of inorganic compounds. Furthermore, someone believed that the modern coronavirus disease was reported initially because of the serious pandemics flu in 1918 (Normile, 2020).

Funding source: Sri Lanka Sabaragamuwa University of Research Grants (SUSL/RE/2017/04).

Data availability: The data used to support the findings of this study are included within the manuscript.

Conflicts of interest: The authors declare that there exists no conflict of interest.

2043-9377
DOI 10.1108/GS-06-2021-0085

The current ongoing pandemic of COVID-19 disease (SARS-CoV-2) was reported initially in the central Chinese city of Wuhan in late December 2019 and within a very short period of time spread around the world (Bogoch, 2020). The latest findings suggested that the SARS-CoV-2 is initially a spill as an animal coronavirus and later has moderated to human-to-human transmissions and has affecting seriously underlying poor health conditions such as cardiovascular disease, diabetes, chronic respiratory disease (Bogoch, 2020). Furthermore, the COVID 19 virus spread directly from person to person with close contacts (within about 6 feet) by respiratory droplets produced because of the cough or sneeze (Yang and Lu, 2020). So, at the present, special attention and efforts have been taken by governments around the world to reduce transmission in susceptible populations including children, health care providers and elderly people (Dong and Du, 2020; Hu *et al.*, 2017).

On January 11, China announced its first COVID-19 death in Wuhan. One week later, the number of deaths increased up to 17 (CDCP, 2020). On January 30, 2020, 7734 cases were confirmed in China and 90 other cases were reported from other Asian countries included Thailand, Vietnam, Malaysia, Singapore, Nepal, Japan, Republic of Korea, the United Arab Emirates, India, Cambodia, Philippines, Sri Lanka and outside the Asian continent include the United States, Australia, Canada, Finland, France and Germany (CSSE, 2020).

The first death of the COVID 19 outside mainland China occurred on 1st February 2020 in the Philippines (Rothan *et al.*, 2020), and outside Asia was happened in France on 14th February 2020 (NCIRD, 2020). At the time of preparing this manuscript on 25th September 2021, the World Health Organization (WHO) has confirmed approximately 156,681,524 COVID positive cases and a death toll of 3,269,340 among 222 countries and territories all over the world (Worldometers, 2021a, b). The WHO reported from 2020 June to up to date, the average number of death per day has been recorded continually around 5,000 amount worldwide. Indeed, about 80% of these deaths were in those over 60, and 75% had pre-existing health conditions including cardiovascular diseases and diabetes (Cao *et al.*, 2021; Wu *et al.*, 2020a, b). Based on the one-month confirmed daily cases from 21st January 2020 to 20th February 2020, Zhao *et al.* (2020) proposed a navel methodology to predict the number of positive COVID cases in China. The findings showed that, the prediction accuracy of the proposed model with the rolling sequence length of 7 is higher than that of it's derived models. The proposed rolling grey Verhulst models predicted the final number of confirmed patients and the date of reaching precisely. Indeed, the results confirmed that the proposed models have predicted the S-shaped change characteristics precisely (Luo *et al.*, 2020a, b).

With relation to the prediction of the daily number of confirmed cases of COVID-19 infectious diseases, Luo *et al.* (2020a, b) has made a significant contribution to the literature. Based on the principle of the grey prediction model and traditional Richards model, Luo *et al.* (2020a, b) proposed a new grey prediction model based on the optimal nonlinear terms. To validated and simulate the proposed model, small-sample and large-sample were selected separately from China, Italy, Britain and Russia. The comparative results show that the model weakens the dependence of the Richards model on single-peak and saturated S-shaped data, but more applicable, and optimize the nonlinear terms and the background values. Based on the grey model (GM[1,1]), nonlinear grey Bernoulli model (NGBM[1,1]) and fractional nonlinear grey Bernoulli model (FANGBM[1,1]), Sahin *et al.* (2020) conducted a study to forecast the number of confirmed cases of COVID-19 in Italy, the United Kingdom (UK) and the United States of America (USA).

The world has been suffering a crucial health crisis and hospital systems across the world are struggling with limited human and physical resources due to the unexpected spread of COVID-19 disease since the end of 2019 (Zhao *et al.*, 2021). Wollenstein *et al.* (2020) carried out a significant study to propose the best resource allocation methodology based on the health systems in Northern Italy, Ecuador and New York City. The current study was carried out based on 91,000 patients details ' in Mexico included demographics, prior medical conditions, SARS-CoV-2 test results, hospitalization, mortality and whether a patient has developed pneumonia or not.

Miscellaneous types of supervised and unsupervised classification methods such as robust versions of logistic regression and support vector machines, as well as random forests and gradient, boosted decision trees were used. The classification accuracies of this model suggested that prediction of hospitalization, mortality, need for ICU and need for a ventilator reached 61%, 76%, 83% and 84% respectively.

At the present, Sri Lanka has become one of the worst-hit countries by the third round of new modifying Indian beta coronavirus. The Specialist Physicians Society in Sri Lanka has estimated that the number of coronavirus deaths per day can exceed 200 inside the country in the coming two months. According to the Institute for Health Metrics and Evaluation (IHME), the University of Washington reports, in worse conditions, more than 20,000 COVID deaths could be estimated to report in Sri Lanka by the end of August. In extreme cases, that number is at risk of exceeding 30,000, with a maximum of 40,000 (Worldometers, 2021a, b).

The gradually increasing magnitude of the COVID-19 pandemic across the Sri Lanka has been sparking an urgent urgency around healthcare systems around the country. The growing debt overhang and weak public finances, the obvious constraints on public resources brought from the pandemic, the access to high-quality and affordable healthcare services for all as and when they needed have been creating a biggest challenge in the Sri Lanka today.

Furthermore, considering current facts and figures, it can be reasonably accepted that a large number of unidentified infected people are roaming freely in society. Under this crucial situation, if we can propose a proper mechanism to estimate the extent of the impact of this epidemic clearly, it would be great. So, the main objective of this research study is to propose a nonlinear stochastic model approach to predict the number of daily COVID- 19 actual death cases precisely.

Due to the volatility with exponential patterns, the ability of forecasting is notoriously embarrassing and represents a major challenge with traditional time series mechanisms; especially, most of the available approaches are weak to forecast unbalanced data frameworks. The purpose of this current study is to propose a HGESM to handle incomplete, noise and exponential types of data in multidisciplinary systems.

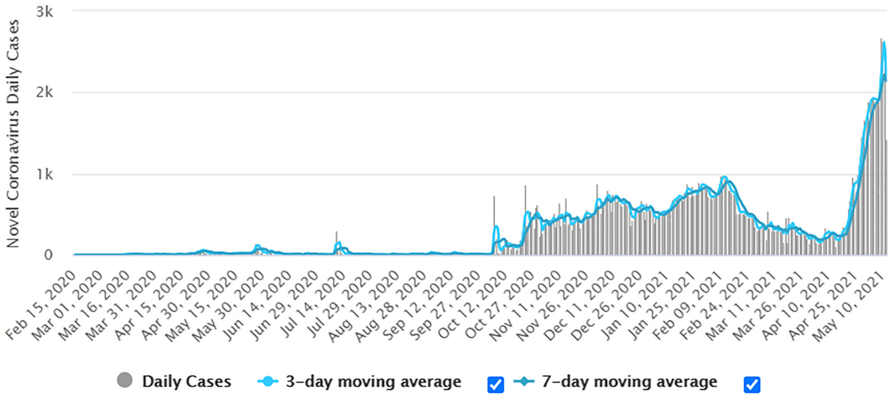
The remainder of this paper is set out as follows. The data preparation with the estimated grey exponential smoothing model is described in Section 2. Section 3 presents the preliminary data analysis and estimated results. Finally, ends up with concluding remarks with Section 4.

2. Methodology

In general, the feature of viral epidemics is the change in its appearance over time. For an example, even a change in a single protein in a virus can cause a significant change in its behavior and may accelerate the spread of the virus. In addition, it can also increase the severity of the disease caused by a viral infection (Cao *et al.*, 2021).

Sri Lanka has become one of the worst-hit countries today because of the third-round coronavirus attack (IHME, 2021). Figure 1 shows that, from February 2020 to November 2020, the Coronavirus daily cases were reported non-significantly. However, since April 2021, the COVID positive cases have been reported in an irregular manner because of the moderated strains such as B.1.1.7; especially, this has caused to spread of the disease faster than usual and has shown more severe symptoms (IHME, 2021). If the coronavirus outbreak reaches the peak continually during the coming weeks, there will be a possibility for Sri Lanka to enter on the red list soon.

The empirical analysis of this study was carried out on the basis of 3rd outbreak of Covid-19 cases in Sri Lanka, which were obtained from the Ministry of Health, Sri Lanka from 1st March 2021 to 15th June 2021. Out of the total 90 daily observations, the first 85% of daily



Source(s): https://www.worldometers.info/coronavirus/country/sri-lanka/

Figure 1.
Novel Coronavirus daily cases in Sri Lanka

confirmed cases were used during the training and the remaining 15% of the sample was considered as the out of sample to test the generalize capabilities of proposed models.

The proposed methodologies consist of two parts. They are double exponential smoothing (DES) and Grey exponential smoothing modeling approach respectively.

2.1 Grey forecasting methodology

The time series forecasting and analysis is a methodology which uses information regarding historical values and associated patterns to predict future demands under the trend analysis, cyclical variations and issues of seasonality. However, time series forecasting creates a very hard task due to the inherently uncertain nature of the real-world data patterns; especially, it seems to be complicated to estimate whether the time series is stochastic, deterministic chaotic or both.

Grey system theory and practices were introduced by Deng Julong in the 1980s with a strong ability to make short-term predictions (Li *et al.*, 2019; Rathnayaka *et al.*, 2020). At the present, this dynamical methodology has been successfully applied for non-negative limited raw data sequences which accords with exponential forms in many fields including transportation, management, science, medicine, engineering and etc (Ji *et al.*, 2007).

In general, the Grey modeling (GM) approach processes throughout the five major steps as follows (Rathnayaka *et al.*, 2015).

Step 1: Identify the major factors which are influencing the objectives and develop an appropriate language model of thoughts and concepts

Step 2: Define terms and concepts to examine the causal relations and hypothesis among the factors identified in step (1)

Step 3: Formulate your research design and identify the significant inputs and outputs

Step 4: Testing the model accuracy and do validations under the different backgrounds

Step 5: Reporting your research findings and make conclusions.

2.1.1 The mathematical model of GM (1,1).

Definition 2.1.1. It is assumed that $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots\ldots\ldots, x^{(0)}(n)\}$ is sequence of non-negative discrete raw data series, which satisfied the condition that $x^{(0)}(k) \geq 0, k = 1, 2, \ldots, n$. Let construct the accumulated generating operator (AGO) or cumulative generated sequence of $X^{(0)}$ is $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\}$, where, $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, $k = 1, 2, \ldots, n$ is a monotonically increasing series (Rathnayaka *et al.*, 2016; Wang *et al.*, 2017).

The AGO is developed under the following theorems and definitions.

Lemma 2.1.1. The operator $D: X^{(0)} \to X^{(1)}$ is called the accumulating operator, where $X^{(1)}$ is a first-order generation by accumulation (Rathnayaka *et al.*, 2015).

Definition 2.1.2. The rth order generation by accumulated generating operation (AGO) of $X^{(0)}$ for $r > 1$ is given by

$$D^r\left(X^{(0)}\right) = X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), \ldots\ldots\ldots, x^{(r)}(n)\right) \tag{01}$$

Where, $x^{(r)}(1) = \sum_{i=1}^{k} \binom{k-i+r-1}{k-1} a x^{(0)}(i); k = 1, 2, \ldots, n$

Let the derivatives of the sequence $X^{(r)} = (x^{(r)}(1), x^{(r)}(2), \ldots\ldots\ldots, x^{(r)}(n))$ defined by

$$d(k) = \frac{dx^{(r)}(k)}{dt} = \lim_{t \to \min t} \frac{x^{(r)}(k) - x^{(r)}(k-1)}{t} \tag{02}$$

$$\frac{dX^{(r)}(k)}{dt} = \frac{\sum_{i-1}^{k} x^{(r-1)}(i) - \sum_{i-1}^{k-1} x^{(r-1)}(i)}{t} = x^{(r-1)}(k) \tag{03}$$

Definition 2.1.3. Let $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots\ldots\ldots, x^{(0)}(n)\}$ be a sequence of non-negative raw data where $n \geq 0$. Then $x^{(0)}(k) + az^{(1)}(k) = b$ which is called a differential equation of $GM(1,1)$ model, where $Z^{(1)}(k) = \frac{x^{(0)}(k) + x^{(0)}(k+1)}{2}; k = 2, 3, \ldots, n$ with development coefficient of a and b, respectively (Li *et al.*, 2019).

The average generating sequence of $X^{(1)}$ is defined by $Z^{(1)} = [z^{(1)}, z^{(2)}, z^{(3)}, \ldots, z^{(n)}]$ Where,

$$Z^{(1)} = \gamma x^{(1)}(k) + (1-\gamma) x^{(1)}(k-1), \quad k = 1, 2, \ldots, n \text{ and } \gamma = 0.5.$$

Theorem 2.1. Assume that $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots\ldots\ldots, x^{(0)}(n)\}$ is a non-negative raw data where $n \geq 0$. Let $X^{(r)} = (x^{(r)}(1), x^{(r)}(2), \ldots\ldots\ldots, x^{(r)}(n))$ is the $r-$AGO sequence. If $r \geq 1, x^{(r)}(k)$ is the increasing function of k, *where*; $k \geq n$.

Proof;

We produced by induction on r. When $r = 1$;

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \ldots\ldots\ldots, x^{(1)}(n)\right)$$

$$X^{(1)}(n) = \left(x^{(0)}(1), x^{(0)}(1) + x^{(0)}(2), \ldots\ldots\ldots, x^{(0)}(1) + x^{(0)}(2) + \cdots + x^{(0)}(n)\right) \tag{04}$$

If $r=2$

$$X^{(2)}=\left(x^{(2)}(1),x^{(2)}(2),\ldots\ldots\ldots,x^{(2)}(n)\right)$$

$$X^{(2)}(n)=\left(x^{(0)}(1),2x^{(0)}(1)+x^{(0)}(2),\ldots\ldots\ldots,nx^{(0)}(1)+(n-1)x^{(0)}(2)+\cdots+x^{(0)}(n)\right). \tag{05}$$

Actually, $\sum_{i=1}^{k}\binom{k-i+r-1}{k-i}x^{(0)}(i)-\sum_{i=1}^{k}\binom{k-2-i+r}{k-1-i}x^{(0)}(i):x^{(0)}(i)>0$

Thus, $X^{(r)}(k)$ is the increasing function of k. ■

According to the definition 2.1.1, 2.1.2 and 2.1.3, the GM(1,1) can be obtained by establishing the whitening equation for $X^{(0)}$ as equation (06);

$$x^{(0)}(k)+az^{(1)}(k)=b \tag{06}$$

$$\frac{d\widehat{x}^{(1)}}{dt}+a\widehat{x}^{(1)}=b \tag{07}$$

If $[\widehat{a}\quad\widehat{b}]^{T}$ is a non-negative sequence of parameters, then the least square estimate sequence of the grey differential equation satisfies;

$$\left[\widehat{a}\quad\widehat{b}\right]^{T}=\left(B^{T}B\right)^{-1}B^{T}Y \tag{08}$$

Where,

$$B=\begin{bmatrix}-z^{(1)}(2) & 1\\ -z^{(1)}(3) & 1\\ \vdots & \vdots\\ -z^{(1)}(n) & 1\end{bmatrix}\qquad Y=\begin{bmatrix}x^{(0)}(2)\\ x^{(0)}(3)\\ \vdots\\ x^{(0)}(n)\end{bmatrix}$$

The respond function or the solution function of the of $GM(1,1)$ is given by;

$$\widehat{x}^{(1)}(k+1)=\left(x^{(0)}(1)-\frac{\widehat{b}}{\widehat{a}}\right)e^{-\widehat{a}k}+\frac{\widehat{b}}{\widehat{a}}\qquad k=1,2,\ldots,n \tag{09}$$

Where, $\widehat{x}^{(1)}(k)$ and $\widehat{x}^{(0)}(k)$ are simulative values or the fitted values of $x^{(1)}(k)$ and $x^{(0)}(k)$ respectively. Furthermore, $\widehat{x}^{(0)}(1),\widehat{x}^{(0)}(2),\ldots,\widehat{x}^{(0)}(n)$ are called the GM(1,1) forecasted values. The restored values of $X^{(0)}$ are given by;

$$\widehat{x}^{(0)}(1)=x^{(0)}(1)$$

$$\widehat{x}^{(0)}(k+1)=\widehat{x}^{(1)}(k+1)-\widehat{x}^{(1)}(k)=\left(1-e^{-\widehat{a}}\right)\left(x^{(0)}(1)-\frac{\widehat{b}}{\widehat{a}}\right)e^{-\widehat{a}k},k=1,2,3,\ldots \tag{10}$$

2.2 Grey exponential smoothing model

Exponential smoothing is one of the window functions that commonly applied for smoothing data in signal processing to remove high-frequency noise.

Lemma 2.2.1. Let assumed that $X^{(0)}=\{x^{(0)}(1),x^{(0)}(2),\ldots\ldots\ldots,x^{(0)}(n)\}$ is original time series. The single exponential smoothing series is defined as;

$$\widehat{x}^{(0)}(k+1) = \propto x^{(0)}(k) + (1-\propto)\,\widehat{x}^{(0)}(k); \quad 0 \le \propto \le 1 \ \textit{where} \propto \text{is a smoothing factor.}$$

Definition 2.2.1. Let's assume that $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots\ldots\ldots, x^{(0)}(n)\}$ is original time series. The Grey double exponential smoothing can be defined as;

$$s^{/}(k) = \propto x^{(r)}(k) + (1-\propto)s^{/}(k-1) \tag{11}$$

$$s^{//}(k) = \propto s^{/}(k) + (1-\propto)s^{//}(k-1) \tag{12}$$

Where, $s^{/}(k)\ and\ s^{//}(k)$ are from single and DES for time k respectively. The forecasting methodology can be express as;

$$\widehat{X}^{(r)}(k) = a_k + k\ b_k \tag{13}$$

Where,

$$a_k = 2s^{/}(k) - s^{//}(k) \ \text{ and } \ b_k = \frac{\propto}{1-\propto}\left(s^{/}(k) - s^{//}(k)\right)$$

2.3 Hybrid grey exponential smoothing model

The combined methodologies under the grey exponential smoothing model and non-linear weighted average component have created high accuracy forecasting than single model approaches.

The proposed new hybrid methodology composed with two main phases based on their DES and grey exponential smoothing model on linear and non-linear domains as follows (Rathnayaka *et al.*, 2015).

$$Y_t = \widehat{X}^{(r)}(k) + N_t \tag{14}$$

Where; L_t and N_t denote the linear autocorrelation and non-linear component of the time series pattern Y_t, respectively. In the initial step, the exponential smoothing model is used to forecast under the stationary conditions.

Let the residual error series of the linear component can be defined as;

$$\varepsilon^{(r)}(k) = X^{(r)}(k) - \widehat{X}^{(r)}(k) \ \ \textit{for}\ k = 1, 2, \ldots, n \tag{15}$$

Where, $\varepsilon^{(r)}(k)$ denotes the residual error at k^{th} time. In the second phase, grey exponential smoothing model is used for $\widehat{\varepsilon}^{(r)}(k)$ forecasting. The approximation of exponential series can be defined as;

$$\varepsilon^{(r)}(k) = A_1 e^{s_1 k} + A_2 e^{s_2 k} + \cdots + A_n e^{s_n k} \tag{16}$$

Where, A_i and S_i are amplitude coefficient and exponent respectively. Let $z_i = e^{s_i}$, then residual error can be approximated as;

$$\widehat{\varepsilon}^{(r)}(k) = A_1 z_1^k + A_2 z_2^k + \cdots + A_n z_n^k \tag{17}$$

The approximated value of the HGESM can be estimated as;

$$\widehat{X}_e^{(r)}(k) = \widehat{X}^{(r)}(k) + \widehat{\varepsilon}^{(r)}(k) \tag{18}$$

Where, $\widehat{X}^{(r)}(k)$ and $\widehat{\varepsilon}^{(r)}(k)$ are approximated values of exponential smoothing series and grey exponential smoothing modeling respectively.

2.4 Model accuracy testing evaluation

Model accuracy testing is a significant way for identifying the suitable model for forecasting accuracy results. Numerous methods have been carried out in the literature to accomplish this goal. In this study, Mean absolute deviation (MAD), mean square error (MSE) and root mean square error (RMSE) were mainly proposed. The formulas of MAD, MSE and RMSE are defined as follows:

$$\varepsilon_{MAD} = \frac{1}{M}\sum_{j=1}^{M}|X_{Aj} - X_{Pj}| \tag{19}$$

$$\varepsilon_{MSE} = \frac{1}{M}\sum_{j=1}^{M}(X_{Aj} - X_{Pj})^2 \tag{20}$$

$$\varepsilon_{RMSE} = \sqrt{\frac{1}{M}\sum_{j=1}^{M}(X_A - X_P)^2} \tag{21}$$

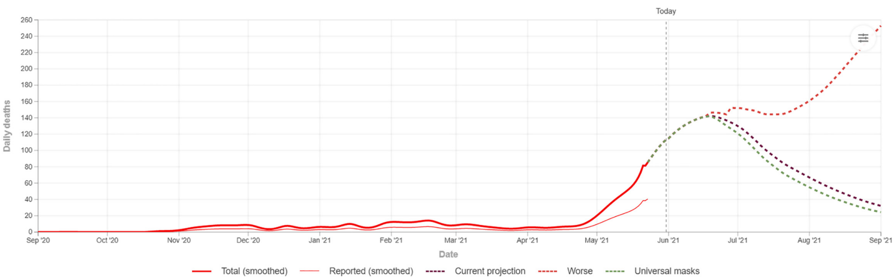
Where, X_A and X_P represent the actual value and predicted value of price at t time respectively.

3. Results and discussion

The Institute for Health Metrics and Evaluation (IHME), the University of Washington reported that because of the poor usage of masks, neglect of social distance and irresponsible political actions, the daily patient growth rate in Sri Lanka has been moving to a worst-case scenario (Hellewell *et al.*, 2020) (see Figure 2).

At the time of writing on 26th September 2021, 513,609 confirmed cases have been reported with 12,680 deaths. Furthermore, up to now, Sri Lankan authorities have tracked down over 300,000 people who had contacted directly with the identified patients and had ordered them to do self-quarantine inside their houses (World meters, 2021a, b).

According to the current practices, the only way to avoid imminent danger is to identify the exact number of confirmed and unconfirmed COVID-19 cases precisely. It is very important for the medical officers to make immediate arrangements in the medical sector; especially, with the limited amount of human and physical resources, the exact estimations of COVID patients are important and essential to allocate ICU beds and ventilators properly to COVID patients. The current study mainly proposed a novel forecasting approach to predict the exact number of confirmed COVID-19 cases for the coming week.



Source(s): https://covid19.healthdata.org/sri-lanka?view=daily-deaths&tab=trend

Figure 2.
Number of estimated death rates in Sri Lanka

3.1 Case study

The main objective of this research study is to propose a nonlinear stochastic model approach to predict the number of daily COVID- 19 death cases precisely. So, in the current study, the data for the period 01st of April 2021 to 23rd May 2021 (50 days) are were used during the training (in-sample or training sample), and the remaining data for the period 24th May 2021 to 10th June 2021 (about 15% of the sample) were considered as the testing sample to test the generalization capabilities of selected models.

3.2 Double exponential smoothing for estimating upcoming confirmed COVID 19 cases

The DES is a forecasting method for univariate time series data, which provides short-term forecasting for handling trends without considering a seasonal component. The DES model in Figure 3 displays the smoothing constants (weights) for the level and trend components for one-step-ahead forecasts.

The results from Figures 3 and 4 warned that, the number of confirmed COVID cases and deaths have been rising exponentially in the coming days. The accuracy measures, mean absolute percentage error (MAPE) and MAD, were respectively 12.7, 79.2 and 59.8, 20 for number of confirmed COVID and number of death for two DES fit's respectively.

According to the forecasting results, the next few weeks will be very decisive; especially death toll per day has been moving beyond 40 is not something to be taken lightly. So, at this juncture, the government, the health as well as the people must work together to adhere to the health guidelines. If any of the three parties take off their responsibilities, it could lead to a serious situation.

3.3 New proposed hybrid grey exponential smoothing model

As a result of high volatility and unstable patterns, the traditional grey forecasting approaches have not achieved success in both linear and non-linear domains with limited data samples. So, combined methodologies under the linear autocorrelation structure and

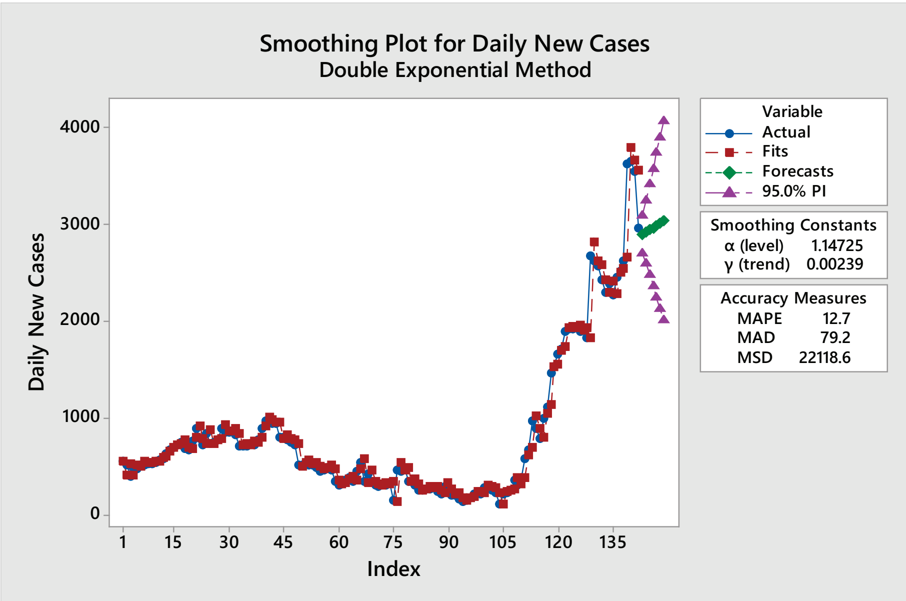


Figure 3.
Smoothing plot for confirmed COVID 19 cases in Sri Lanka from 1st January 2021 to 22nd May 2021

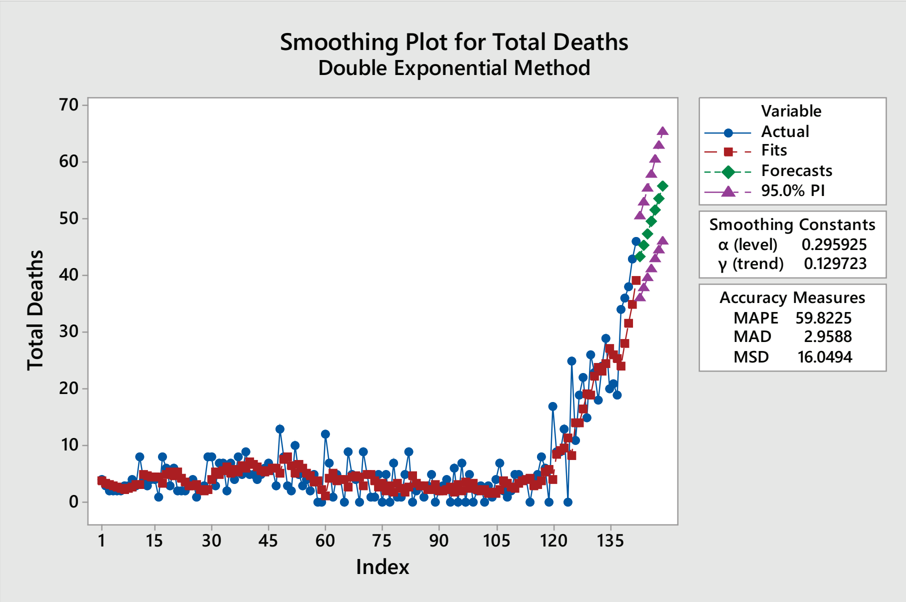


Figure 4. Smoothing plot for total death because of the COVID 19 in Sri Lanka from 1st January 2021 to 22nd May 2021

non-linear weighted average component have created high accuracy forecasting than single model approaches.

The new proposed hybrid methodology can be described as follows using our proposed case study. So, in the current study, the data for the period 2nd May 2021 to 23rd May 2021 (21 days) are were used during the training (in-sample or training sample), and the remaining data for the period 24th May 2021 to 10th June 2021 (about 15% of the sample) were considered as the testing sample to test the generalization capabilities of selected models.

Step 01. The original data series as follows.

$$X^{(0)}(k) = [9, 9, 13, 1, 25, 11, 19, 22, 15, 26, 23, 18, 24, 29, 20, 21, 19, 34, 36, 38, 43] \tag{22}$$

The accumulated generating sequence obtained as follows.

$$X^{(1)}(k) = [9, 18, 31, 32, 57, 68, 87, 109, 124, 150, 173, 191, 215, 244, 264, 285, 304, 338, 374, 412, 455] \tag{23}$$

$$Z^{(1)}(k) = [13.5, 24.5, 31.5, 44.5, 62.5, 77.5, 9, 116.5, 137, 161.5, 182, 203, 229.5, 254, 274, 294.5, 321, 356, 393, 433.5]^T \tag{24}$$

$[a\ \ b]^T$ is a non-negative sequence of parameters. Thus we can obtain the least square estimate sequence of the grey differential equation as;

$$\left[\hat{a} \quad \hat{b}\right]^T = \left(B^T B\right)^{-1} B^T Y = [-0.06495 \quad 10.25]^T \tag{25}$$

Where,

$$B^T B = \begin{bmatrix} 1503380.0 & -4703.0 \\ -4703.00 & 22.00 \end{bmatrix}$$

So, we can update the approximate function of the GM(1,1) model as;

$$\widehat{x}_1^{(0)}(k+1) = \left(x^{(0)}(1) + 10.199\right)e^{-0.0654k} - 0.0654 \quad k = 1, 2, 3, \ldots, n \tag{26}$$

Step 02. The grey DES as;

$$s^{/}(k) = 0.85x^{(r)}(k) + (1 - 0.85)s^{/}(k-1)$$

$$s^{//}(k) = 0.85s^{/}(k) + (1 - 0.85)s^{//}(k-1)$$

Step 03. The approximated value of the *HGESM* estimated as;

$$\widehat{X}_e^{(0)}(k) = \widehat{X}^{(0)}(k) + \widehat{\varepsilon}^{(0)}(k) = \{9, 11, 12, 13, 15, 16, \ldots.., 36, 39, 41, 44\} \tag{27}$$

Where; $\widehat{X}^{(r)}(k)$ *and* $\widehat{\varepsilon}^{(r)}(k)$ are approximated values of $GM(1,1)$ and *Exponential Smoothing series,* respectively.

Table 1 results conformed suggested, with minimum MAD, MSE and MAPE (%), the proposed *HGESM* have always given the best performances compared with other single models.

As a comparison mode the selected methods are used to assess the out-of-sample forecasting performance for the horizon of one week ahead (testing samples) from 24th May 2021 to 30th May 2021 (Sample 01) and 31st May 2021 to 5th June 2021 (Sample 02).

According to the error analysis results in Table 1 and Figure 5, the new proposed HGESM is highly accurate (less than 10%) with lowest RMSE values in a one head as well as weakly forecasting's than traditional single exponential smoothing and DES models; especially, MAD accuracy testing results confirmed that new proposed HGESM model has given more accurate results for forecasting exponential growth time series predictions.

4. Conclusion and discussion

The number of COVID-19 virus has infected approximately 168 million people worldwide from the end of December 2019 till today. The virus has infected the entire world, except for a handful of countries and the death toll reached nearly 3.5 million by the end of May 2021 (NCIRD, 2020). Different types of forecasting methodologies can be seen to forecast daily Covid-19 deaths in the literature. However, most of the empirical studies suggested that

Table 1. The model accuracy for coming week

	Forecasting accuracy (%)							
Model accuracy	Single exponential smoothing (SES)		Double exponential smoothing (DES)		GM(1,1)		*HGESM*	
	Sam.01	Sam.02	Sam.01	Sam.02	Sam.01	Sam.02	Sam.01	Sam.02
MAD (%)	15.42	12.47	10.85	11.54	6.86	4.01	4.14	2.71
MSE	264.29	250.26	135.43	112.25	64.57	19.42	25.57	10.14
RMSE	16.26	15.82	11.64	10.59	8.03	4.40	5.05	3.18

Note(s): *Denotes the model with the minimum error values

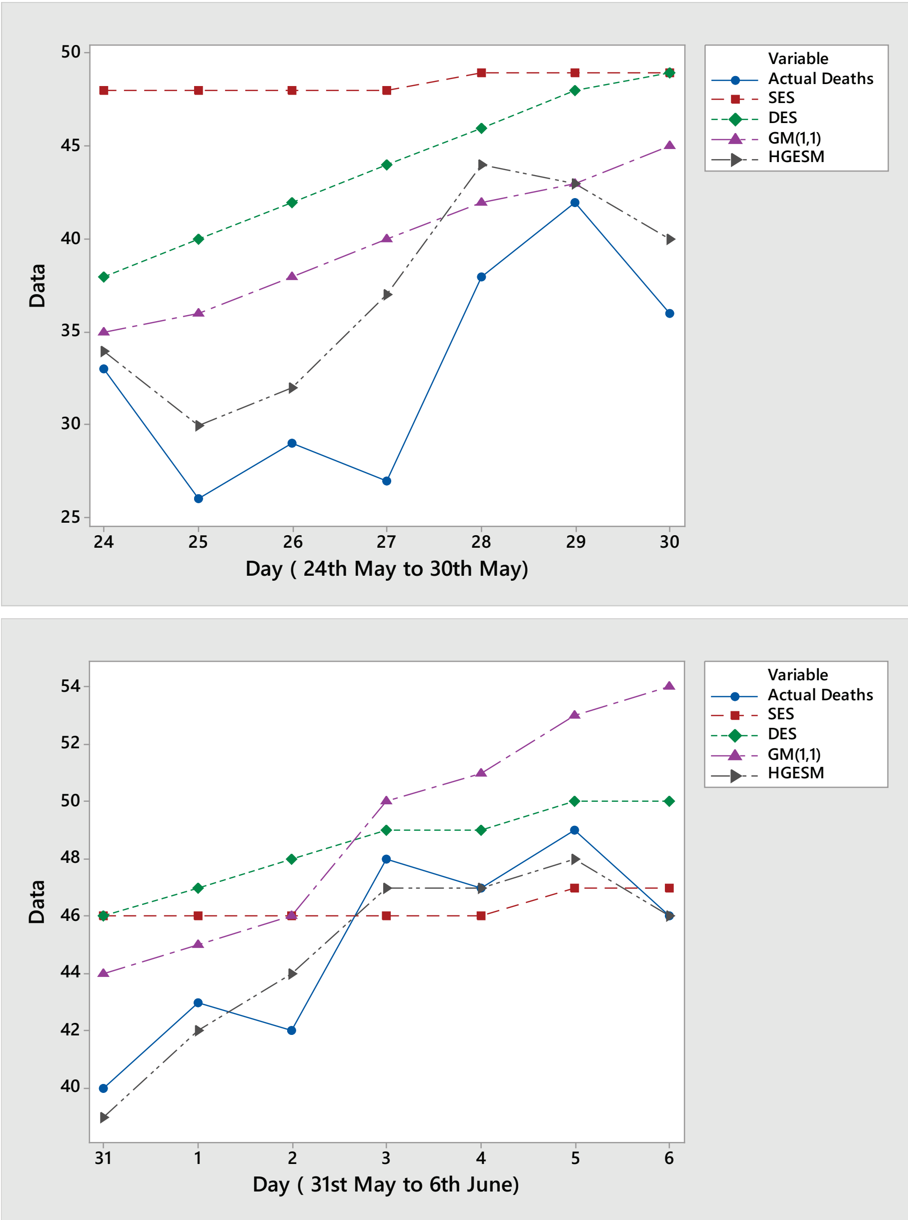


Figure 5.
The simulative, predictive values confirmed COVID-19 Cases

nonlinear separate methods are not sufficient and not enough to forecast modern data indices under the conditions of uncertainty; especially data studies without normality, linearity and stationary assumptions. For example, some forecasting models are great at short-term predictions, but cannot capture seasonality or variability with very limited numbers of sample observations.

This current study presents exponential smoothing, DES, GM(1,1) and grey exponential smoothing modeling approaches to predict the number of COVID-19 death cases in Sri Lanka based on the period from 1st of April 2021 to 23rd May 2021 (50 days). The prediction performances of the models are tested by MAD), MSE and RMSE. The main conclusion of this study can be given as:

(1) The empirical findings suggested that the new proposed HGESM is selected as the best prediction model with having the lowest MAD, MSE and RMSE values.

(2) The findings confirm that the new proposed HGESM is more suitable and effective for forecasting time series with the exponential trends in a short-term manner.

Moreover, following suggestions can be given for further studies on this topic.

(1) Time series analysis is an essential methodology for analyzing the characteristics for making future adjudications in decision making. As a result of the complications, with regard to the traditional time series approaches with uncertainty analysis, neural network computing models with the new hybrid methodology are more suitable and appropriate for handling incomplete, noisy and uncertain data in multidisciplinary systems.

The most important thing that can be done in practice is to take strong and effective measures to control the spread of the disease immediately. At the time of preparing this manuscript, the government of Sri Lanka introduced various sequential measures to improve social distancing such as the closure of schools and education institutes, introducing their work from home concept for the government and private sector, reduce the public gathering, introducing travel bans for international arrivals. Furthermore, the Sri Lankan government drastically imposed an island-wide lockdown for 14 days to minimize the burden of the disease to the health system and the entire community.

Furthermore, the prevailing situation suggests, without ending up this COVID 19 epidemic totally from the society, we will no longer be able to adopt that comfortable lifestyle back. Whether the world is totally free from this virus, someday it will come back to the world again and again. That date could be another month, two, three or six. So, if we do not prepare for the present challenge properly, we will be defeated as a country. It is the responsibility of all of us, not only of the government, to look into these issues in the future as well as to address the current situation.

References

Bogoch, A.W.B. (2020), "Pneumonia of unknown etiology in Wuhan, China: potential for international spread", *Med*, Vol. 27 No. 2, doi: 10.1093/jtm/taaa008.

Cao, J.L., Tu, W.J., Hu, X.R. and Liu, Q. (2021), "Clinical features and short-term outcomes of 102 patients with Coronavirus disease 2019 in Wuhan, China", *Clinical Infectious Diseases*, Vol. 12 No. 11, pp. 25-36.

Centers for Disease Control and Prevention (CDC) (n.d.), U.S. Department of Health and Human Services, available at: https://covid.cdc.gov/covid-data-tracker/#county-view?list_select_state=all_states&list_select_county=all_counties&data-type=.

CSSE, A.J. (2020), *Coronavirus COVID-19 Global Cases*, Center for Systems Science and Engineering (CSSE) . Johns Hopkins University (JHU).

Dong, E. and Du, H. (2020), "An interactive web-based dashboard to track COVID-19 in real time", *Lancet Infectious Disease*, Vol. 1 No. 3, doi: 10.1016/S1473-3099(20)30120-1.

Hellewell, J., Abbott, S., Gimma, A., *et al.* (2020), "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts", *The Lancet Global Health*, Vol. 8 No. 4, pp. e488-e496.

Hu, B., Zeng, L.P. and Yang, X.L. (2017), "Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus", *PLoS Pathogens*, Vol. 13 No. 11, pp. 57-64.

IHME (2021), "Data release information sheet, institute for health Metrics and evaluation", available at: http://www.healthdata.org/about/terms-and-conditions.

Ji, P., Luo, X. and Zou, H. (2007), "A study on properties of GM(1,1) model and direct GM(1,1) model", *Proceedings of 2007 IEEE International conference on DGrey Systems and Intelligent Services*, Nanjing, Vol. 12 No 2, pp. 399-403.

Li, B., Yang, W. and Li, X. (2019), "Application of combined model with DGM(1,1) and linear regression in grain yield prediction", *Grey Systems: Theory and Application*, Vol. 1, pp. 25-34, doi: 10.1108/GS-07-2017-0020.

Luo, X., Duan, H. and Xu, K. (2020a), "A novel grey model based on traditional Richards model and its application in COVID-19", *Chaos, Solitons and Fractals*, Vol. 10 No. 1, pp. 56-63, doi: 10.1016/j.chaos.2020.110480.

Luo, X. Duan, H. and Xu, K. (2020b), "A Novel Grey model based on traditional Richards model and its application in COVID-19", *Chaos, Solitons and Fractals*, Vol. 13 No. 1, doi: 10.1016/j.chaos.2020.110480.

NCIRD (2020), *National Center for Immunization and Respiratory Diseases COVID-19 _Travel Health Notices*.

Normile, D. (2020), "Coronavirus cases have dropped sharply in South Korea. What's the secret to its success?", *Science*, Vol. 3 No. 6, doi: 10.1126/science.abb7566.

Rathnayaka, R.M.K.T. and Seneviratna, D.M.K.N. (2020), "Taylor series approximation based unbiased GM(1,1) hybrid statistical approach for forecasting stock market", *The Journal of Grey System*, Vol. 32 No. 3, pp. 125-135.

Rathnayaka, R.M.K.T., Seneviratna, D.M.K.N. and Jianguo, W. (2015), "Grey system based novel approach for stock market forecasting", *Grey Systems: Theory and Application*, Vol. 5 No. 2, pp. 178-193, doi: 10.1108/GS-04-2015-0014.

Rathnayaka, R.M.K.T., Seneviratna, D.M.K.N., Jianguo, W. and Arumawadu, H.I. (2016), "An unbiased GM(1,1)-based new hybrid approach for time series forecasting", *Grey Systems: Theory and Application*, Vol. 6 No. 3, pp. 322-340, doi: 10.1108/GS-04-2016-0009.

Rothan, H. and Nagadenahalli, S. (2020), "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak", *Journal of Autoimmunity*, Vol. 15 No. 10, doi: 10.1016/j.jaut.2020.102433.

Sahin, U. and Sahin, T. (2020), "Forecasting the cumulative number of confirmed cases of COVID-19 in Italy, UK and USA using fractional nonlinear grey Bernoulli model", *Chaos, Solitons and Fractals Nonlinear Science, and Nonequilibrium and Complex Phenomena*, Vol. 138 No. 1, doi: 10.1016/j.chaos.2020.109948.

Shao, P. (2020), "Impact of city and residential unit lockdowns on prevention and control of COVID-19", *MedRxiv*, Vol. 10 No. 11, doi: 10.1101/2020.03.13.20035253.

Wang, J., Wang, Z. and Li, Q. (2017), "Export injury early warning of the new energy industries in China: a combined application of GM(1,1) and PCA methods", *Grey Systems: Theory and Application*, Vol. 7 No. 2, pp. 272-285, doi: 10.1108/GS-02-2017-0003.

WHO (2020), "WHO statement regarding cluster of pneumonia cases in Wuhan", available at: https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-ofpneumonia-cases-in-wuhan-china.

Wollenstein-Betech, S., Cassandras, C.G. and Paschalidis, I.C. (2020), "Personalized predictive models for symptomatic COVID-19 patients using basic preconditions: hospitalizations, mortality, and the need for an ICU or ventilator", *medRxiv*, Vol. 10 No. 1, doi: 10.1101/2020.05.03.20089813.

Worldometers (2021a), "COVID-19 coronavirus pandemic", available at: https://www.worldometers.info/coronavirus/#countries.

Worldometers (2021b), "COVID-19 coronavirus pandemic", available at: https://www.worldometers.info/coronavirus/country/sri-lanka/.

Wu, J.T., Leung, K. and Bushman, M. (2020a), "Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China", *Nature Medicine*, Vol. 26 No. 4, pp. 506-510.

Wu, J.T., Leung, K. and Leung, G.M. (2020b), "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study", available at: http://www.thelancet.

Yang, Y. and Lu, Q. (2020), "Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China", medRxivAljazeera, Vol. 10 No. 13.

Zhao, Y.F., Shou, M.H. and Wang, Z.X. (2020), "Prediction of the number of patients infected with COVID-19 based on rolling grey verhulst models", *International Journal of Environment Research Public Health*, Vol. 17, p. 4582, doi: 10.3390/ijerph17124582.

Zhao, T.F., Shou, M.H. and Wang, Z.X. (2021), "Prediction of the number of patients infected with COVID-19 based on rolling grey verhulst models", *International Journal of Environmental Research Public Health*, Vol. 17, p. 4582, doi: 10.3390/ijerph17124582.

Further reading

Heath Promotion Bureau (2021), "COVID-19: live situational analysis dashboard of Sri Lanka", available at: https://hpb.health.gov.lk/covid19-dashboard/.

Kapila Tharanga Rathnayaka, R.M., Seneviratna, D. M. K. N., Jianguo, W. and Arumawadu, H.I. (2015), "A hybrid statistical approach for stock market forecasting based on Artificial Neural Network and ARIMA time series models", *2015 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC)*, pp. 54-60, doi: 10.1109/BESC.2015.7365958.

Rathnayaka, R.M.K.T. and Seneviratna, D.M.K.N. (2019), "Taylor series approximation and unbiased GM(1,1) based hybrid statistical approach for forecasting daily gold price demands", *Grey Systems: Theory and Application*, Vol. 9 No. 1, pp. 5-18, doi: 10.1108/GS-08-2018-0032.

Corresponding author
D.M.K.N Seneviratna can be contacted at: seneviratna@is.ruh.ac.lk