A network analysis and support vector regression approaches for visualising and predicting the COVID-19 outbreak in Malaysia

Siti Nurhidayah Sharin [a], Mohamad Khairil Radzali [b], Muhamad Shirwan Abdullah Sani [c,d,e,*]

[a] *Halal Products Research Institute, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[b] *Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[c] *International Institute for Halal Research and Training, International Islamic University Malaysia, Level 3, KICT Building, 53100 Kuala Lumpur, Malaysia*
[d] *Konsortium Institut Halal IPT Malaysia, Ministry of Higher Education, Block E8, Complex E, Federal Government Administrative Centre, 62604 Putrajaya, Malaysia*
[e] *The Catalytixs Solutions, No. 713, Jalan DPP 1/4, Desa Permai Pedas, 71400 Pedas, Negeri Sembilan, Malaysia*

ARTICLE INFO

ABSTRACT

This study aims to (1) correlate and visualise the Coronavirus disease 19 (COVID-19) pandemic spread via Spearman rank coefficients of network analysis (NA) and (2) predict the cumulative number of COVID-19 confirmed and death cases via support vector regression (SVR) based on COVID-19 dataset in Malaysia between July 2020 to June 2021. The NA indicated increasing connectivity between different states throughout the time frame, revealing the most complex network of COVID-19 transmission in the second quarter of 2021. The SVR model predicted future COVID-19 cases and deaths in Malaysia in the second half of 2021. The study demonstrated that the NA and SVR could provide relatively simple yet valuable artificial intelligence techniques for visualising the degree of connectivity and predicting pandemic risk based on confirmed COVID-19 cases and deaths. The Malaysian health authorities used the NA and SVR model results for preventive measures in highly populated states.

1. Introduction

World Health Organization (WHO) reported that 44 cases of unknown pneumonia aetiology were detected in Wuhan City, Hubei Province of China, between 31st December 2019 to 3rd January 2020 [1]. The Chinese authorities identified a new type of coronavirus known as coronavirus disease 2019 (COVID-19) on 12th January 2020 [2–4]. The WHO declared this communicable disease a global pandemic on 11th March 2020 due to its rapid worldwide [2–6].

The Ministry of Health Malaysia (MHM) confirmed the first COVID-19 case in local spreading on 25th January 2020 [7], where they successfully flattened the infection curve with less than 100 daily cases with a strict Movement Control Order (MCO) between July to September 2020 [8]. Nonetheless, a third outbreak hit Malaysia in late September 2020 [1], which caused a significant loss of human life and imposed an unprecedented threat to the economy, food systems, and public health. The supply chain manufacturing of essential products and services had slowed down [9] due to strict MCO, including the travel ban, stay-at-home and social distancing policies [10]. National businesses faced a deceleration of economic revenue growth and caused the retrenchment of millions of workers [11]. The MCO affected the food supply chains and led to food loss. Malaysian farmers had to

destroy food such as marine and vegetable products that reached expiry to ensure food safety for the consumers [12]. The COVID-19 inevitably impacted the health deterioration of more than 4.75 million communities in Malaysia and the death of more than 37000 patients to date. To prevent this sporadic and communicable disease from worsening, the MHM and enforcement authorities have established standard operating procedures (SOP) and guidelines [13]. Due to the rapid changes in the COVID-19 infection, academicians and statisticians have assisted the MHM by providing reliable insights on the infection via data analytics of the recorded COVID-19 cases. This study provided insights from the data analytics that may support the establishment of the SOP.

Data analytics on the recorded COVID-19 cases plays an essential role in response to the global COVID-19 pandemic. It offers a rich data source of COVID-19 for statisticians, physicists, and engineers from data-driven and computational perspectives. While some of these data were inaccessible before the epidemics, others were available but underutilised [14]. With the emerging programming languages using Python and R in data visualisation and prediction, public knowledge is expanding rapidly, rendering a greater understanding of the spreading of COVID-19. Data analysts employed various techniques of artificial intelligence (AI) to visualise and predict the dissemination

of COVID-19 worldwide. Among the AI techniques, network analysis (NA) and support vector regression (SVR) are among the visualisation and prediction techniques used in science, finance, economy, tourism and social sectors [7]. Despite that, negligible reports of these AI techniques are available for the health sector, especially for visualising and predicting COVID-19 cases. The NA is a simple yet powerful method to evaluate the pandemic risk by visualising the correlation among various regions based on real-time and historical data [15]. Data analysts in China employed the NA to investigate COVID-19 transmission among China's population in the major cities [16]. Another study has identified super spreaders among South Korean COVID-19 patients after removing top nodes with a higher degree of connectedness or key infection in NA, indicating that the NA is a potential AI to study COVID-19 transmission [17]. The relationship among states in Malaysia associated with COVID-19 cases is still unknown. Hence, this study extracted one-year recorded COVID-19 cases between July 2020 to June 2021 via https://github.com/MoH-Malaysia/covid19-public and investigated the connectivity of COVID-19 infection and transmission among the states in Malaysia using the NA followed by the SVR to predict the COVID-19 transmission.

The SVR is a well-known and influential supervised machine learning technique from a support vector machine (SVM) that finds the best-fitted line for linear and non-linear regressions in predicting COVID-19-related cases [18]. A study of COVID-19 in Bangladesh showed an excellent regression value of 0.8230 and 0.8322 for infected and death cases, respectively, using the radial basis function (RBF) kernel of SVR [19]. Also, a study of COVID-19 on the Turkish population showed that the SVR performed an excellent regression value at 94% with a low root mean square error (RMSE = 0.034) and mean absolute error (MAE = 0.036) compared to linear regression (LR), bagged tree (BT) and fine tree (FT) algorithms [20]. Besides transmission among the population, the SVR has successfully investigated COVID-19 transmission via food access in Guilford County, North Carolina [21]. While SVR is applied to predict COVID-19 transmission globally, a negligible report is available on its application in Malaysia. Hence, this study predicted COVID-19 cases using the SVR. Although this study provided a prediction of a future number of daily infection and mortality rates from July 2021 until December 2021, it indicated that the SVR was applicable for this purpose by providing insights on transmittable disease and assisting the government in preparing the mitigating measures against COVID-19 spread, especially in highly populated states.

Considering the increment of daily COVID-19 cases and fatalities in Malaysia, this study incorporated NA to understand the degree of connectedness among the Malaysian states and the SVR to forecast future COVID-19 cases. These statistical techniques are essential complementary tools for reliable visualisation and prediction of Malaysia's forthcoming confirmed cases and deaths. These techniques can also give an early insight and understanding into preventing and curbing this COVID-19 disease from rampantly spreading in Malaysia.

2. Methodology

2.1. Data source

In this study, three .csv files (cases_states, cases_malaysia and deaths_malaysia) were retrieved from an open-source website provided by the MHH via https://github.com/MoH-Malaysia/covid19-public. These three .csv files entailed COVID-19-related information from July 2020 to June 2021 (365 days).

2.2. Bar plot and Spearman's analysis

Before building the bar plot and Spearman's analysis using Python 3.3, cumulative confirmed cases (cases_malaysia) and deaths (deaths_malaysia) in Malaysia were combined into a .csv file dataset. In this analysis, a total number of 365 days was observed (3rd quarter

of the year 2020, 4th quarter of the year 2020, 1st quarter of the year 2021 and 2nd quarter of the year 2021; $n = 365$). A new column named 'days' was created in the same combined dataset to represent the date into particular days in 2020 and 2021 (365 rows × 4 columns). A bar plot was built to observe Malaysia's general COVID-19 cumulative confirmed cases and deaths from July 2020 to June 2021. Based on the skewness (confirmed cases = 1.14, deaths = 2.31), kurtosis (confirmed cases = 0.57, deaths = 4.52) and Shapiro–Wilk test at p-value < 0.05 for confirmed cases = 3.34×10^{-17} and deaths = 5.57×10^{-28}, the data was assumed as non-parametric (not normally distributed). Therefore, Spearman's analysis at p-value < 0.05 was conducted to determine the correlation strength and significant difference between cumulative confirmed cases and deaths in Malaysia.

2.3. Network analysis

A dataset of daily confirmed cases of states in Malaysia (cases states) was employed in this analysis consisting of 5840 observations (16 states × 3rd quarter of the year 2020 × 4th quarter of the year 2020 × 1st quarter of the year 2021 × 2nd quarter of the year 2021, $N = 5840$). A new column named 'quarter' is created in the dataset to classify the months into the respective quarter of 2020 and 2021, while the original columns remained (5840 rows × 4 columns). From a holistic approach, network analysis (NA) investigates the co-occurrence of features and reveals potential associations between features or functional pathways [22]. The cor.test() function in R or more robust tools for compositional data, such as the SparCC (sparse correlations for compositional data) package, can be used to calculate correlation coefficients (Spearman rank coefficient) and significant p-value [23]. Once calculated, the network can be analysed and visualised using Gephi [24], R library igraph [22], or Cytoscape [25].

The NA was conducted to study the connections between states in Malaysia in response to the COVID-19 spread. Basic network properties such as Spearman rank correlation and degree of interaction were computed using the iGraph R package (V 3.5.1) [22] before being fed into Cytoscape (V 3.8.2) for visualisation. The degree of interaction was determined based on the correlation (edge) between the states. By default, the states' correlation was considered a strong co-occurrence if the Spearman correlation coefficient is > 0.5 [23] and statistically significant if the computed p-value is < 0.05. Additionally, the NetworkAnalyzer Cytoscape plugin was applied to calculate the number of significant nodes and edges that denoted the networks' topological properties. In this study, nodes represented the states, whereas edges represented the correlations (nodes). Additionally, connected components are the maximum group of nodes connected by edges in a path, while network density is the ratio of observed edges over possible edges in a given network.

2.4. Support vector regression (SVR)

The support vector regression (SVR) relies on support vector machine (SVM) elements, where support vectors are generally closer points toward the generated hyperplane in an n-dimensional feature space that distinctly segregates the data points around the hyperplane. The SVR model performs the fitting. The generalised equation for hyperplane is represented as $y = wX + b$, where w is the weight and b is the intercept at X = 0. Meanwhile, the margin of tolerance is denoted by epsilon ϵ.

In this prediction modelling, COVID-19 cases in Malaysia in the 2nd quarter of 2021 showed a high correlation among all states based on the NA. Therefore, the same dataset in Section 2.2 was used to forecast cases for the Malaysian country. The steps of building the SVR model were adapted from Parbat & Chakraborty (2020) work [6]. This study developed the SVR model using a 70% training dataset via the Radial Basis Function (RBF) kernel with epsilon = 0.1. The developed SVR model was validated with a 30% testing dataset. Based on the SVR model, predictive values of future confirmed cases and deaths in Malaysia were computed by employing the Python 3.3 command:
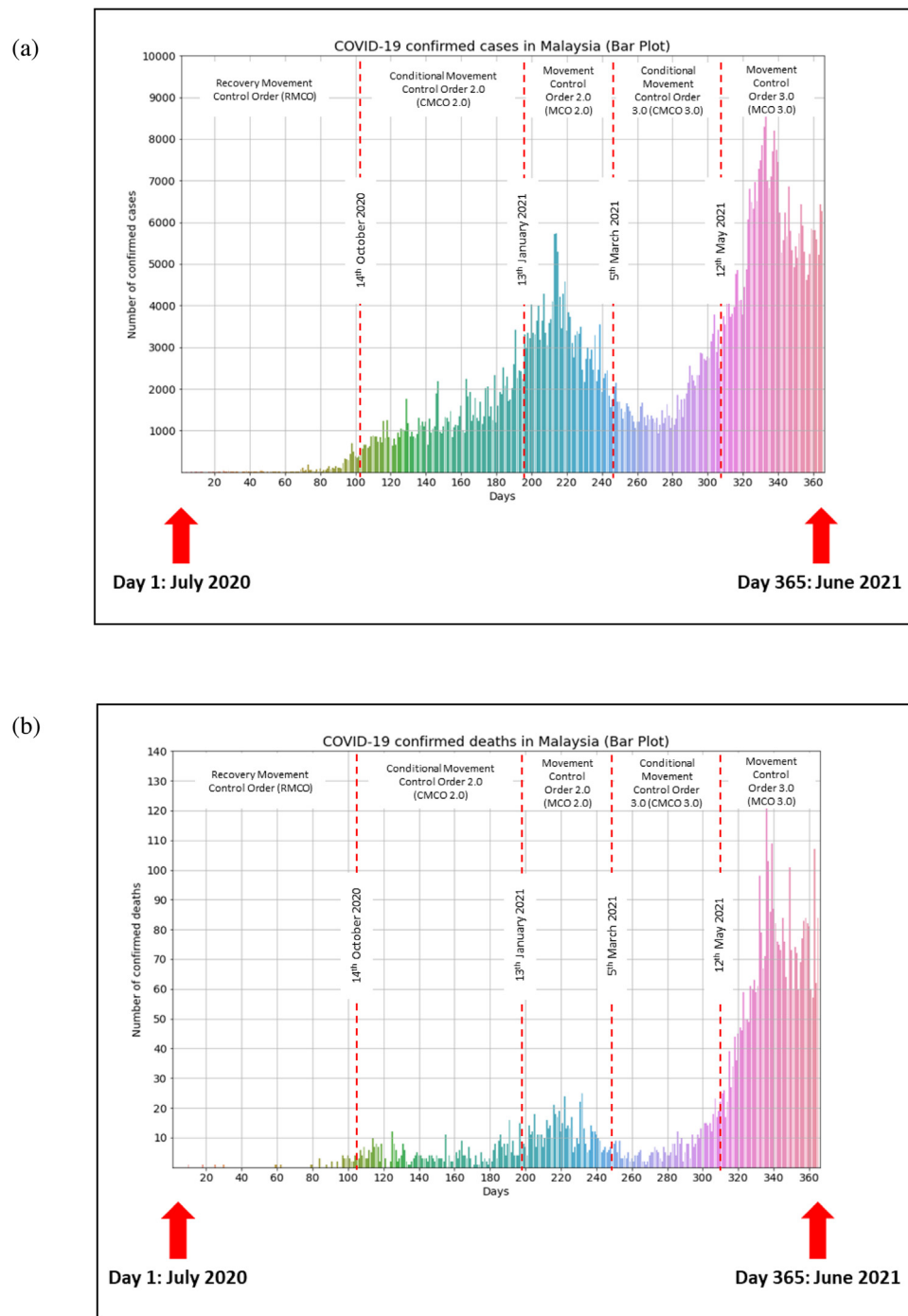
(a)

(b)

Fig. 1. The trend of COVID-19 (a) confirmed cases and (b) confirmed deaths from July 2020 until June 2021 (365 days).

y_pred = scy.inverse_transform(regressor.predict(scx. transform([[100]]))).

These values were manually inputted in Microsoft Excel to construct the SVR forecast. The performance of the SVR model was evaluated based on statistical goodness-of-fit criteria, e.g., regression score (R^2), mean squared error (MSE), and root mean square error (RMSE).

3. Results

3.1. General insight of COVID-19 confirmed cases and deaths in Malaysia

Using a bar plot, this study visualised the trend of COVID-19 cumulative confirmed cases and deaths in Malaysia from July 2020 to June 2021 (Fig. 1). A trend was observed in cumulative confirmed cases in Malaysia starting from Day 70 (8th September 2020), with a triple-digit number of 100 cases (Fig. 1(a)). A similar pattern was observed on accumulated confirmed deaths on day 84 (22nd September 2020) with 3 cases (Fig. 1(b)). There were significant increments (p-value < 0.05) in daily confirmed cases and deaths in Malaysia from July 2020 to June 2021 (Table 1). The confirmed cases had a strong correlation with deaths ($r = 0.907$, p-value < 0.05), in which high cases influenced a high mortality rate (Table 1). Based on Fig. 1(a), the increment of confirmed cases for Day 280–Day 340 (6th April 2021–5th June 2021) was higher than the confirmed cases for Day 70–Day 215 (8th September 2020–31st January 2021).
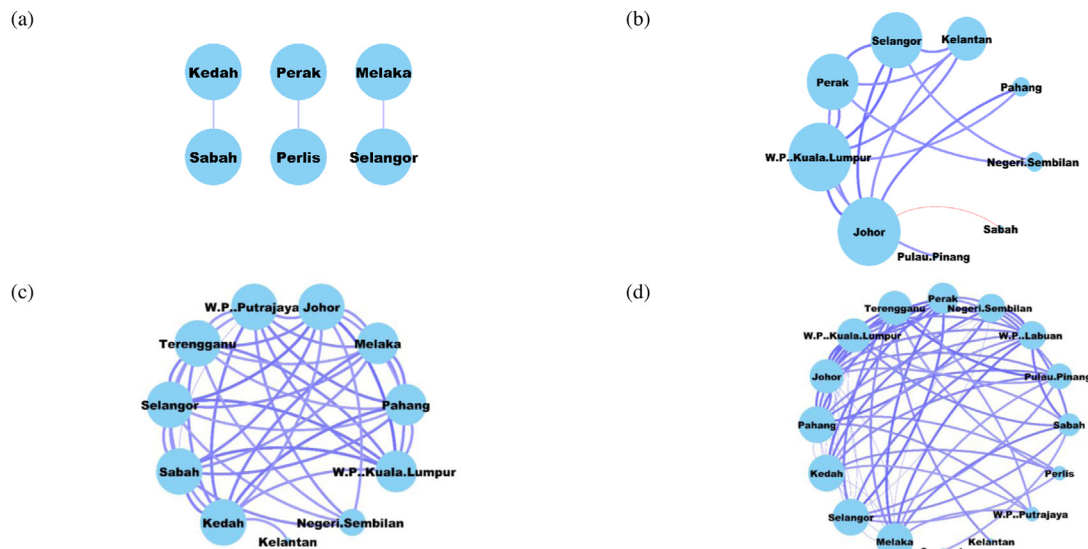
Fig. 2. Co-occurrence networks based on daily confirmed state cases for (a) Q3 of 2020, (b) Q4 of 2020, (c) Q1 of 2021, and (d) Q2 of 2021. The node size is proportional to the number of connections (i.e., degree), and the edge width (i.e., weight) is proportional to the r value. The edge is coloured blue and red for positive and negative correlation, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Spearman correlation coefficient of confirmed cases vs days confirmed deaths vs days and confirmed cases vs confirmed deaths in Malaysia from July 2020 until June 2021 (365 days).

Correlation type	Data	r-value	p-value
	Confirmed cases vs days	0.902	1.868e−134
Spearman's correlation	Confirmed deaths vs days	0.863	1.862e−109
	Confirmed cases vs confirmed deaths	0.907	1.260e−138

3.2. Correlation among states using network analysis

In the current study, the NA was constructed to determine the relationship of states in Malaysia associated with confirmed COVID-19 cases. States that were densely connected with others exhibited higher complexity of edges in the network graph, suggesting a critical centre of virus transmission throughout the networks [15]. This study used Spearman's rank coefficient to measure the polarity (−1 to 1) of correlation between states based on daily confirmed cases. A positive value of the Spearman rank correlation represents co-existence, whereas a negative value indicates opposition between two states. This study divided the COVID-19 dataset into four timeframes; Quarter 3 (Q3) of 2020 (July 2020–September 2020), Quarter 4 (Q4) of 2020 (October 2020–December 2020), Quarter 1 (Q1) of 2021 (January 2021–March 2021), and Quarter 2 (Q2) of 2021 (April 2021– June 2021) and investigated the correlation between states that led to the spiked number of cases.

Table 2 summarises the number of nodes, edges and analysis time of these quarters. The significantly different correlations (p-value < 0.05) were discussed in this section. In Q3 of 2020, Sabah and Kedah were highly correlated ($r = 0.329$) despite having a weak positive correlation compared to Perak with Perlis ($r = 0.322$) and Melaka with Selangor ($r = 0.326$) (Fig. 2(a)). The NA revealed a total of nine states, including FT Kuala Lumpur, Johor, Perak, Selangor, Kelantan, Pahang, Negeri Sembilan, Pulau Pinang and Sabah that were significantly correlated in which Johor and FT Kuala Lumpur had the highest degree of interaction among others based on the visualisation (Fig. 2(b)). Of the nine states, FT Kuala Lumpur and Selangor had a strong positive correlation ($r = 0.765$), followed by Johor and Selangor ($r = 0.756$). Other correlations among the states are summarised in Table 3.

In Q1 of 2021 (January 2021 to March 2021), Fig. 2(c) showed more complex networks in which 11 states formed 42 significantly strong positive correlations. Based on the visualisation, Johor, Kedah, Sabah, Selangor, Terengganu, and FT Putrajaya exhibited a similar and highest degree of interaction associated with COVID-19 transmission. Out of these states, Sabah and FT Putrajaya had the strongest positive correlation ($r = 0.834$), followed by Johor and FT Kuala Lumpur ($r = 0.754$). Additionally, Johor also exhibited positive correlations with Selangor ($r = 0.715$), Sabah ($r = 0.646$), Terengganu ($r = 0.637$), FT Putrajaya ($r = 0.574$), Melaka ($r = 0.566$), Kedah ($r = 0.552$), Pahang ($r = 0.536$), and Negeri Sembilan ($r = 0.535$). Table 4 summaries other correlations among the states.

The second quarter (Q2) of 2021 revealed the most complex network in the current finding (Fig. 2(d)). All 16 states significantly correlated in COVID-19 transmission nationwide and exponentially increased the number of confirmed and death cases. All four states, including Selangor, Pahang, Melaka, and Kedah, had the highest degree of interaction (12 edges), among others. Kedah and Selangor remained the states with the highest degree of interaction from 9 to 12 correlations (edges) from Q1 to Q2 of 2021, respectively. Selangor and FT Kuala Lumpur also had a strong positive correlation ($r = 0.886$). Subsequently, Melaka exhibited the highest positive correlations with Selangor ($r = 0.883$), Negeri Sembilan ($r = 0.860$), and Pahang ($r = 0.854$). Table 5 summarises other correlations among the states.

3.3. Prediction of confirmed cases and deaths in Malaysia using support vector regression model

This study employed SVR to observe the reliability of this model in predicting the future number of confirmed cases and deaths in Malaysia. The construction of the SVR model entailed a 70% training set of confirmed cases vs days, confirmed deaths vs days, and confirmed cases vs confirmed deaths. The model yielded a high R2 value (~1) of 0.846, 0.859 and 0.829, respectively (Table 6), and low MSE and RMSE (~0) indicated that the SVR model was considered a reliable predictive model [26]. Moreover, the low MSE and RMSE values also indicated the high accuracy of the SVR model. Meanwhile, the R2 values of 30% testing set for confirmed cases vs days, confirmed deaths vs days and confirmed deaths vs confirmed cases were 0.855, 0.909 and 0.836, as shown in Fig. 3(a), Fig. 3(b) and Fig. 3(c), respectively. Based on

Table 2

Summary statistics for each network analysed using the NetworkAnalyzer Cytoscape plugin.

Summary Statistics	Quarter 3 2020 (July–September 2020)	Quarter 4 2020 (October–December 2020)	Quarter 1 2021 (January–March 2021)	Quarter 2 2021 (April–June 2021)
Number of nodes (states)[a]	6	9	11	16
Number of edges (correlation)	3	16	42	67
Analysis time (s)	0.106	0.008	0.014	0.029

[a]The node represents states in Malaysia, whereas the edge represents a correlation between the states.

Table 3

Spearman correlation coefficient (r-value) between the states in Q4 of 2020 (October–December 2020).

State 1[a]	State 2[a]	r value (p-value < 0.05)
Johor (6)	Kelantan (4)	0.595
	Pahang (2)	0.674
	Perak (5)	0.607
	Sabah (1)	−0.530
	Selangor (5)	0.756
	FT Kuala Lumpur (6)	0.756
Kelantan (4)	Perak (5)	0.592
	Selangor (5)	0.608
	FT Kuala Lumpur (6)	0.640
Negeri Sembilan (2)	Perak (5)	0.539
	Selangor (5)	0.532
Pahang (2)	FT Kuala Lumpur (6)	0.542
Perak (5)	Selangor (5)	0.676
	FT Kuala Lumpur (6)	0.675
Pulau Pinang (1)	FT Kuala Lumpur (6)	0.523
Selangor (5)	FT Kuala Lumpur (6)	0.766

[a]Value in parenthesis indicates the degree of interaction of the state.

Table 4

Spearman correlation coefficient (r-value) between the states in Q1 2021 (January–March 2021).

State 1[a]	State 2[a]	r-value (p-value < 0.05)
Johor (9)	Kedah (9)	0.552
	Melaka (8)	0.566
	Negeri Sembilan (5)	0.535
	Pahang (8)	0.536
	Sabah (9)	0.646
	Selangor (9)	0.715
	Terengganu (9)	0.637
	FT Kuala Lumpur (8)	0.754
	FT Putrajaya (9)	0.574
Kedah (9)	Kelantan (1)	0.507
	Melaka (8)	0.666
	Pahang (8)	0.702
	Sabah (9)	0.667
	Selangor (9)	0.550
	Terengganu (9)	0.627
	FT Kuala Lumpur (8)	0.559
	FT Putrajaya (9)	0.682
Melaka (8)	Pahang (8)	0.645
	Sabah (9)	0.576
	Selangor (9)	0.609
	Terengganu (9)	0.514
	FT Kuala Lumpur (8)	0.615
	FT Putrajaya (9)	0.599
Negeri. Sembilan (5)	Sabah (9)	0.597
	Selangor (9)	0.514
	Terengganu (9)	0.516
	FT Putrajaya (9)	0.584
Pahang (8)	Sabah (9)	0.645
	Selangor (9)	0.508
	Terengganu (9)	0.504
	FT Kuala Lumpur (8)	0.557
	FT Putrajaya (9)	0.677
Sabah (9)	Selangor (9)	0.584
	Terengganu (9)	0.551
	FT Kuala Lumpur (8)	0.701
	FT Putrajaya (9)	0.834
Selangor (9)	Terengganu (9)	0.656
	FT Kuala Lumpur (8)	0.703
	FT Putrajaya (9)	0.518
Terengganu (9)	FT Kuala Lumpur (8)	0.643
	FT Putrajaya (9)	0.537
FT Kuala Lumpur (8)	FT Putrajaya (9)	0.618

[a]Value in parenthesis indicates the degree of interaction of the state.

Fig. 3(a) and Fig. 3(b), the predicted values of daily confirmed cases and deaths from Day 1 until Day 365 (July 2020 to June 2021) were lower but almost similar to the actual reported cases. Fig. 4 forecasts the future number of daily infection and mortality rates from July 2021 until December 2021. It was observed that the number of confirmed cases and deaths in Malaysia may spike between July and August 2021, and may show a downward trend starting in September 2021, provided that the Malaysian authorities continue the same interventions to curb the COVID-19 transmission.

4. Discussion

4.1. Spike of COVID-19 cases in Malaysia

Generally, the signs of COVID-19 usually appear after 1 – 14 days of the incubation period but commonly occur after five days [24]. Based on the estimation of serial interval and incubation period, it was estimated that 44% of transmission probably had occurred before symptoms appeared [25]. The previous study reported a significant relationship between viral load and incubation period, in which the initial viral load began to increase within 5 to 6 days before the first symptoms appeared [25]. The incubation period becomes shorter when the viral loads are high, corresponding to low cycle threshold (CT) values. Since viral loads evolve, high viral loads are probably the primary cause of transmission [27].

Previously, Malaysia had successfully curbed the first and second waves of the outbreak by lowering the confirmed cases in July until early September 2020 (Day 1 – Day 69), with less than 100 cases per day (Fig. 1(a)). However, the fourth week of September 2020 (Day 85–Day 92) showed an increment of confirmed cases (Fig. 1(a)) and deaths (Fig. 1(b)), commencing the third epidemic wave in Malaysia. The increase of confirmed cases occurred right after the state election in Sabah on 26th September 2020 (Day 88) [2]. Many cases are associated with high-risk areas in Sabah, which led to 29 clusters located in Sabah, where 26 clusters had an index case with travel history to Sabah mainly from the east of Sabah, including Lahad Datu, Semporna, Tawau, Kunak and Sandakan areas [28]. Despite the increment of confirmed cases in Sabah, the people's movements within Malaysia were not restricted. Also, COVID-19 tests were not mandatory before travelling among the states, resulting in the escalation of confirmed cases from single digit to thousands per day [29].

The condition in Sabah has become worse due to a lack of awareness about COVID-19 transmission and its symptoms, especially among people who live in rural areas, failure to compliance with the instructions
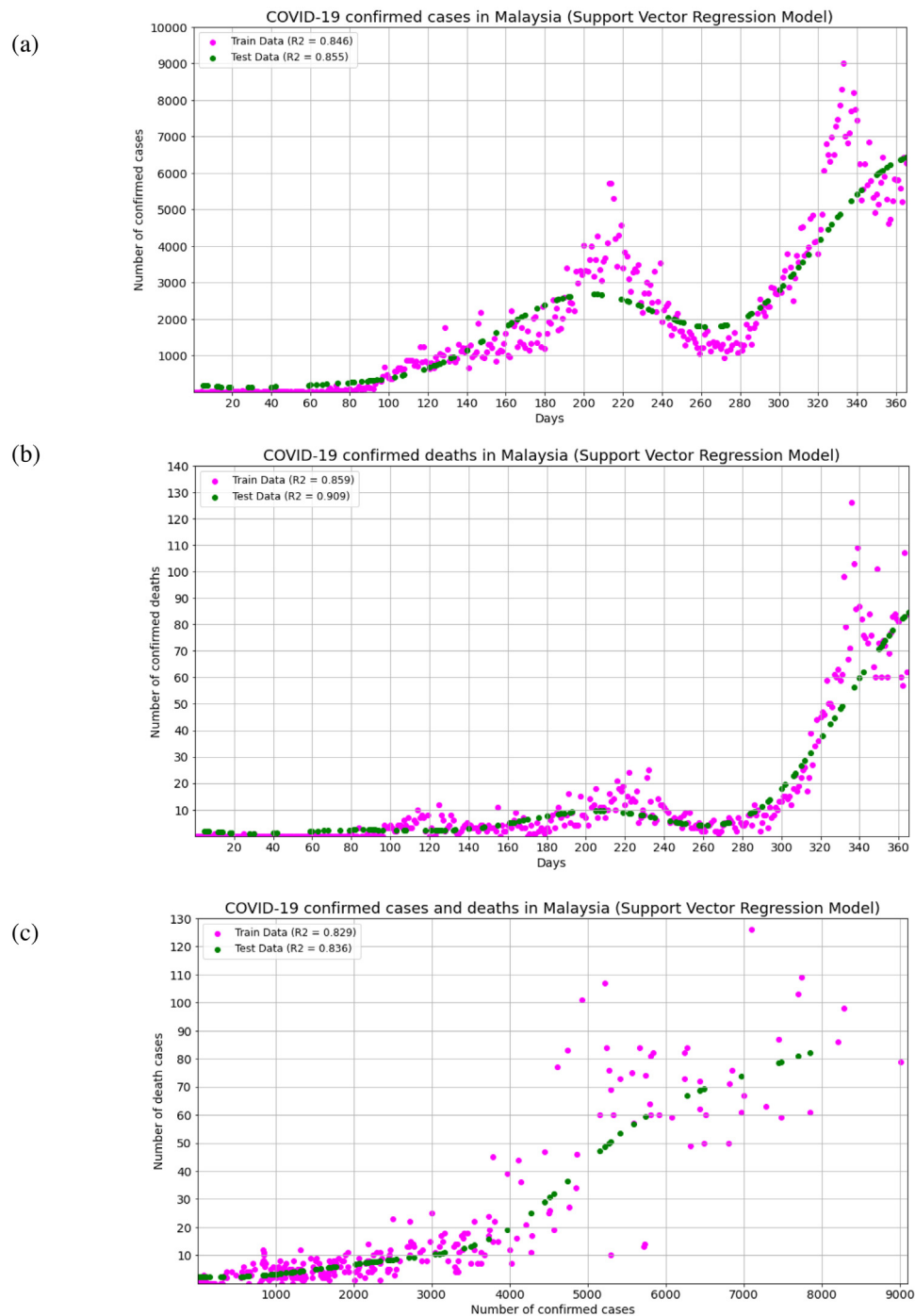
(a)

(b)

(c)

Fig. 3. SVR models of (a) confirmed cases vs days, (b) confirmed deaths vs days, and (c) confirmed deaths vs confirmed cases from July 2020 until June 2021 (365 days).

by health officers, as well as the paucity of healthcare workers in Sabah's hospitals which had caused 10400 backlogged of COVID-19 test samples [30]. The Sabah was one of the top three states with the highest population composition at 11.7%, preceded by the Selangor at 20.1% and followed by Johor at 11.6% [31]. However, the population density per one square kilometre in Sabah ($52/km^2$) is not relatively high as in Federal Territory (FT) Kuala Lumpur ($7188/km^2$), FT Putrajaya ($2354/km^2$), Selangor ($674/km^2$), and Johor ($174/km^2$) [29]. Although the population density in Sabah is not as highly dense as in Peninsular Malaysia, the majority of 3.83 million people in Sabah are settling along the Sabah's coastline instead of the interior mountainous parts, which has caused the spike of COVID-19 cases in those areas after the state election of Sabah [32]. Besides, irregular and undocumented

migrants in Sabah refused to go for COVID-19 screening tests and contact tracing since they were at risk of detention or deportation, resulting in difficulty in getting robust and reliable data [33].

Other than that, the commencing of triple-digit COVID-19 cases was observed on Day 70 (8th September 2021) during the recovery movement control order (RMCO) and later exponentially increased during the conditional movement control order 2.0 (CMCO 2.0) between Day 106–Day 196 (14th October 2020– 12th January 2021) [34]. The exponential increment in COVID-19 cases during CMCO 2.0 was due to the emergence of new clusters post the Sabah state election held on 26th September 2020. The Malaysian government then decided to implement the MCO 2.0 again on 13th January 2021 (Day 197) after observing COVID-19 cases that reached thousands per day [35].
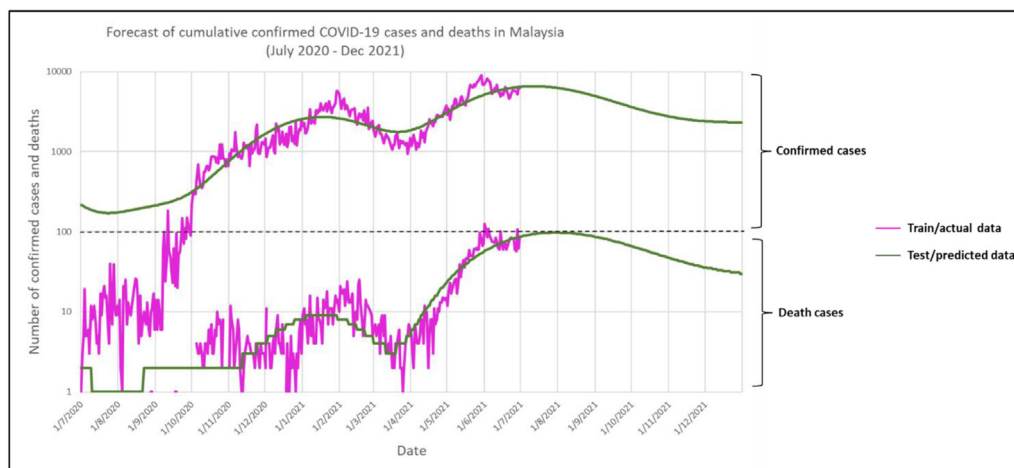
Fig. 4. Forecast of COVID-19 cumulative confirmed cases and deaths for the next 184 days (starting from July 2021 until December 2021).

Between Day 197– Day 247 (13th January 2021–4th March 2021), MCO 2.0 successfully showed a decreasing trend in COVID-19 cases per day. However, the implementation of MCO 2.0 did not last long. Once again, the government announced the third CMCO (CMCO 3.0) on 5th March 2021 for the safety of Malaysia's economy [36]. Although the MCO 2.0 execution was not stricter than MCO 1.0 and most businesses were allowed to operate, Malaysia still recorded a RM 600 million loss per day since most businesses struggled in the recovery phase, and investors remained pessimistic [37].

During the CMCO 3.0 (Day 280–Day 340), the spike in COVID-19 cases was observed to be higher than CMCO 2.0 and MCO 2.0 (Day 70–Day 215) due to the failure of mass COVID-19 testing and compliance with the standard operating procedures (SOPs) in Selangor and Penang, as well as the emergence of new coronavirus variants with higher infection rates comprising of United Kingdom variant (Alpha Variant B.1.1.7), South African variant (Beta Variant B.1.351), and Indian variant (Delta Variant B.1.617.2) [38]. Social gathering activities and the people's concentration in crowded spaces were the primary causes of spreading due to societies' difficulty in complying with the SOPs. In Selangor, the government state decided to fully utilise the antigen rapid test kit (RTK-Antigen) method during the mass COVID-19 testing since the testing results can be obtained on the same day compared to the reverse transcription-polymerase chain reaction (RT-PCR) method, which the testing results takes up to three days and caused backlog [39]. Mass COVID-19 testing using RTK-Antigen aims to detect and isolate the silent carriers and better understand the positivity rate and hotspots. Therefore, the increment of COVID-19 cases was anticipated. The increasing number of COVID-19 cases has also caused an overburden on the healthcare system, particularly in highly affected states such as Selangor, Sarawak, Penang, Kelantan and FT Kuala Lumpur, leading to the escalating of COVID-19 deaths [28].

4.2. Transmission complexity in second quarter (Q2) of 2021

Despite having a slight positive correlation, Sabah and Kedah were significantly connected ($r = 0.329$, $p < 0.05$) in Q3 2020 (Fig. 2 (a)). Sabah and Kedah had individual 1505 and 270 confirmed cases throughout the quarter, yet no reports linking the COVID-19 transmission between these two states. Sabah, which had reported the first cluster on 1st September 2020 at Lahad Datu District Police Headquarters lock-up, this state had accounted for 74.7% of the total new cases between 7th to 13th September 2020 [33]. Kedah had the earliest positive COVID-19 cases where the Sivagangga cluster emerged and spread to Perlis and Penang. Several factors were linked to COVID-19 transmission in Kedah, including the lack of physical distancing,

family gatherings that flouted standard operating procedures (SOP), and hospital visits [28]. Generally, the MHH expressed alarming concern about COVID-19 spreading in Q3 2020 since most respiratory viral tract infections were reported during rainy seasons in tropical regions [40], where Malaysia undergoes two monsoon seasons with rapid wind speed faced from late May to September and November to March in Southwest and Northeast Malaysia, respectively. These rainy seasons might contribute to the transmission of COVID-19 within this time frame [41].

Surprisingly, the confirmed COVID-19 cases increased from 2594 to 101786 from Q3 (Fig. 2(a)) to Q4 2020 ((Fig. 2(b)). Based on the NA, nine states were highly associated, including FT Kuala Lumpur, Johor, Perak, Selangor, Kelantan, Pahang, Negeri Sembilan, Pulau Pinang, and Sabah, with Johor and FT Kuala Lumpur had the highest degree of interaction among others (Fig. 2 (b)). The increasing number of COVID-19 cases might have been contributed by geographical factors such as high population density and population movement, especially in urban centres [42]. Manufacturing industries also contributed to the confirmed COVID-19 cases in FT Kuala Lumpur and Selangor [43]. Johor was also positively correlated with FT Kuala Lumpur ($r = 0.756$), Pahang ($r = 0.674$), Perak ($r = 0.607$), and Kelantan ($r = 0.595$). Other correlations (r values and degree of interaction) are summarised in Table 3.

However, Johor and Sabah showed a negative correlation ($r = -0.530$) in Table 3, suggesting strategic implementation in Sabah that might reduce the spread of COVID-19 in Johor (Fig. 2(b)). Several comprehensive implementations in Sabah included limited non-essential services, implementation of Targeted Enhanced Movement Control Order (TEMCO), increasing of healthcare equipment capacity (beds, ventilators, etc.), medical personnel mobilisation, point-of-entry testing, maximum daily RT-PCR testing capacity, mandatory 14-day quarantine at designated centres, quarantine centres for undocumented migrants and stringent border control. Apart from that, placing Johor under Conditional Movement Control Order (CMCO) and MCO, closing worship places, opening COVID-19 for quarantine and low-risk-treatment centres, and enforcing SOPs reduced the COVID-19 spreading from this state [44].

A more complex network consisting of 11 states and 42 significant positive correlations was identified in Q1 2021 (January to March 2021) in Fig. 2 (c). This finding was in accordance with the shifting of the National Transmission State Assessment stage from Stage 2 (Localised Community Transmission) to Stage 3 (Large-scale Community Transmission – Low Confidence), signifying that the COVID-19 spread was high among the communities. During this Q1 2021, the increase in COVID-19 spread was potentially due to inter-state travel during

Table 5
Spearman correlation coefficient (r-value) between the states in Q2 2021 (April–June 2021).

State 1[a]	State 2[a]	r-value (p-value < 0.05)
	Kedah (12)	0.783
	Melaka (12)	0.776
	Negeri Sembilan (9)	0.727
	Pahang (12)	0.712
	Perak (10)	0.750
Johor (11)	Pulau Pinang (8)	0.607
	Sabah (7)	0.545
	Selangor (12)	0.799
	Terengganu (11)	0.764
	FT Kuala Lumpur (11)	0.761
	FT Labuan (9)	0.615
	Melaka (12)	0.799
	Negeri Sembilan (9)	0.702
	Pahang (12)	0.808
	Perak (10)	0.736
	Perlis (4)	0.585
Kedah (12)	Pulau Pinang (8)	0.635
	Selangor (12)	0.832
	Terengganu (11)	0.770
	FT Kuala Lumpur (11)	0.771
	FT Labuan (9)	0.539
	FT Putrajaya (4)	0.506
Kelantan (1)	Terengganu (11)	0.605
	Negeri Sembilan (9)	0.860
	Pahang (12)	0.854
	Perak (10)	0.760
	Pulau Pinang (8)	0.524
Melaka (12)	Sabah (7)	0.674
	Selangor (12)	0.883
	Terengganu (11)	0.619
	FT Kuala Lumpur (11)	0.824
	FT Labuan (9)	0.745
	FT Putrajaya (4)	0.536
	Pahang (12)	0.725
	Perak (10)	0.654
Negeri Sembilan (9)	Sabah (7)	0.670
	Selangor (12)	0.843
	FT Kuala Lumpur (11)	0.824
	FT Labuan (9)	0.757
	Perak (10)	0.722
	Perlis (4)	0.549
	Pulau Pinang (8)	0.642
	Selangor (12)	0.821
Pahang (12)	Terengganu (11)	0.683
	FT Kuala Lumpur (11)	0.739
	FT Labuan (9)	0.579
	FT Putrajaya (4)	0.574
	Pulau Pinang (8)	0.552
	Selangor (12)	0.752
Perak (10)	Terengganu (11)	0.652
	FT Kuala Lumpur (11)	0.672
	FT Labuan (9)	0.585
Perlis (4)	Selangor (12)	0.511
	Terengganu (11)	0.530
	Selangor (12)	0.576
Pulau Pinang (8)	Terengganu (11)	0.762
	FT Kuala Lumpur (11)	0.528
	Sarawak (1)	0.510
	Selangor (12)	0.609
Sabah (7)	FT Kuala Lumpur (11)	0.622
	FT Labuan (9)	0.726
	Terengganu (11)	0.680
Selangor (12)	FT Kuala Lumpur (11)	0.886
	FT Labuan (9)	0.684
	FT Kuala Lumpur (11)	0.607
Terengganu (11)	FT Putrajaya (4)	0.544
FT Kuala Lumpur (11)	FT Labuan (9)	0.622

[a]Value in parenthesis indicates the degree of interaction of the state

holiday celebrations, mainly in FT Kuala Lumpur, Selangor, Johor, Penang, Sabah, Kedah, Perak, Negeri Sembilan, and Melaka [28]. A few festive seasons (Q1 2021) that applied to these states, including New Year's Day (1st January 2021), Thaipusam (28th January 2021), and Chinese New Year (12th–13th February 2021), hence might lead to an increase in population movement within the time frame. Data from Google Mobility Report also indicated a surge of cumulative population movement (workplace, retail and recreations, parks, grocery and pharmacy, and transit stations) in Johor, Kedah, Sabah, Selangor, Terengganu, and FT Putrajaya within the quarter (Fig. 5) [45], suggesting that festive season may be one of the potential factors of COVID-19 spread [46]. Hence, the MHM shall take the necessary actions prior to the event. The snow-ball effect on the infection from these festive seasons could be seen in the following second quarter (Q2) of 2021.

The second quarter (Q2) of 2021 was the most complicated network in the present study, where the highest degree of interaction in Selangor, Pahang, Melaka, and Kedah occurred (Fig. 2(d)). The Malaysian authorities changed the stage of the National Transmission Stage Assessment from Stage 3 (Large-Scale Community Transmission – Low Confidence) to Stage 3 (Large-Scale Community Transmission – Moderate Confidence) effective on 26th April 2021 and further shifted to Stage 3 (Large-Scale Community Transmission – High Confidence) on 10th May 2021. From Q1 to Q2 of 2021, Kedah and Selangor remained the states with the highest degree of interaction, ranging from 9 to 12 correlations (edges). During the time frame, the surge cases in Kedah and Selangor were linked to densely populated areas and those who contracted the virus at factories [47].

Selangor and FT Kuala Lumpur had a strong positive correlation ($r = 0.886$) and consistently reported a high proportion of confirmed cases compared to Sarawak, Penang, Johor, and Kelantan due to the burden of the healthcare system. Moreover, multiple hospitals across FT Kuala Lumpur and Selangor struggled with the surged admission of critically ill COVID-19 patients requiring oxygen support [48], leaving more contracted patients to stay at home and leading to more infection. Additionally, 132673 and 11873 confirmed COVID-19 cases in Selangor and Melaka had been reported in the same quarter. However, no reports between Melaka and Selangor were found despite having a strong positive correlation ($r = 0.883$). Thus, this study inferred that the transmission might be due to inter-state travel and the rapid spread of COVID-19 within the local community, educational institutions, and places of worship. Considering the rise of population movement (residential, grocery and pharmacy) (Fig. 6), the asymptomatic carriers and the emergence of new COVID-19 variants in Q1 of 2021 could potentially cause the virus to be more transmissible among all states, including the Melaka and Selangor [49]. In addition, several national festive seasons in Q2 of 2021 (April–July 2021), including Labour Day (1st May 2021), Eid Fitr (13th–14th May 2021) and Wesak Day (26th May 2021), might link to the increase of population movements and contribute to the increment of COVID-19 infection. Likewise, the Malaysian authorities shall take precautions by inculcating the communities to comply with SOPs and imposing a targeted movement restriction in the states with high populations.

4.3. Potential of support vector regression (SVR) model in predicting confirmed and death cases associated with COVID-19 in Malaysia

Although SVR is one of the standard predictive tools in predicting COVID-19-related cases, our findings revealed higher R^2 and lowered MSE and RMSE values than in a previous study [50]. It was observed that the predicted values of daily confirmed cases and deaths from Day 1 until Day 365 (July 2020 to June 2021) were lower but almost similar to the actual reported cases in Fig. 3(a) and Fig. 3(b). This finding indicated that the SVR was a reliable and robust method to briefly predict the impending number of daily infections and mortality rates. Nevertheless, in this study, the prediction of the SVR model was solely based on historical data and did not consider the reproduction number

Table 6
Model performance of support vector regression for confirmed cases, confirmed deaths and days.

Prediction model	Data	70% training dataset[b]			30% testing dataset[b]		
		MSE	RMSE	R^2	MSE	RMSE	R^2
Support vector regression[a]	Confirmed cases vs days	0.154	0.393	0.846	0.151	0.388	0.855
	Confirmed deaths vs days	0.141	0.376	0.859	0.108	0.329	0.909
	Confirmed cases vs confirmed deaths	0.171	0.413	0.829	0.196	0.443	0.836

[a]Execution of support vector regression was based on radial basis function (RBF) kernel with epsilon value 0.1.
[b]MSE = Mean squared error; RMSE = Root mean square error; and R^2 = Regression score.
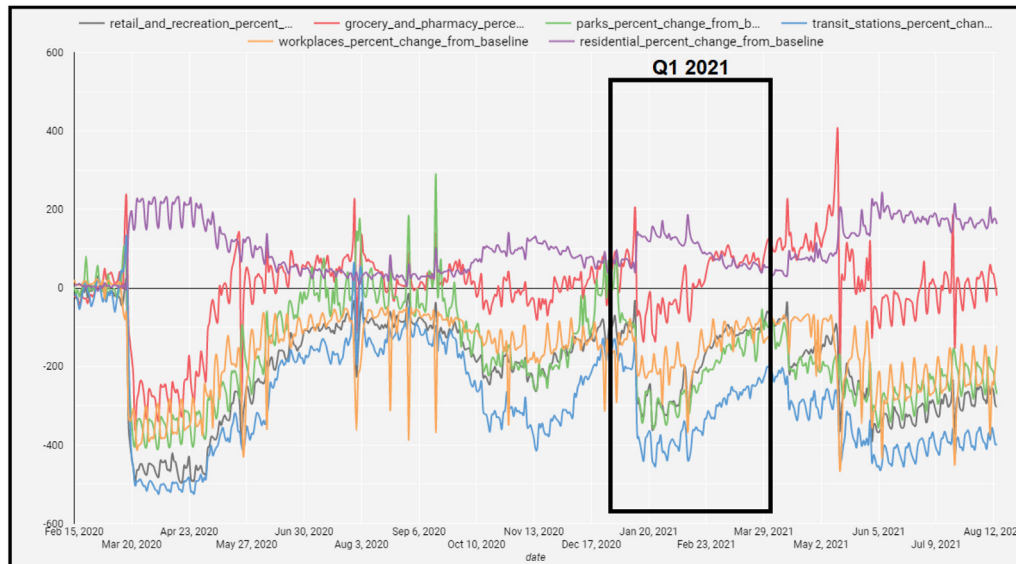
Fig. 5. Cumulative mobility data (%) over time for Johor, Kedah, Sabah, Selangor, Terengganu, and F.T. Putrajaya in Q1 of 2021 (Google LLC, 2021).

(R_0). The R_0 is the estimated number of cases an infected individual causes in spreading the disease to other individuals who are not yet infected. The R_0 was utilised to determine the potential for a disease to spread in that population [7]. Recently, the determination of the R_0 value has been vital since this value is able to indicate the severity rate of the outbreak to spread among individuals [50]. Since our current aim only focuses on observing the SVR model's reliability in predicting the forthcoming COVID-19 cases, the R_0 value may be proposed together with the SVR model for future study. The forecast for the future number of daily infection and mortality rates predicted from July 2021 until December 2021 (Fig. 4) indicated that the number of confirmed cases and deaths in Malaysia spiked between July and August 2021, and a downward trend to start in September 2021, provided that the MHH and Malaysia government remain the similar interventions to curb the COVID-19 transmission. However, this study emphasised that this forecast was merely based on daily confirmed cases and deaths variables, and the prediction via SVR needs more variables to investigate the influence of other variables on the COVID-19 trend in Malaysia.

4.4. Linkage between the results of network analysis and support vector regression

Based on the NA of Q2 2022 in Fig. 2(d) and the decrease of predicted confirmed death cases in August 2021 (Fig. 4), the Malaysian government should implement stringent preventive measures mainly in states that have a high degree of interactions, such as Melaka, Selangor, Kedah and Pahang. Following New Zealand's COVID-19 elimination strategies, preventive measures, including extensive border management, public communication on physical distancing, hand washing, mask-wearing, cough etiquette, and hospital capacity, may reduce

COVID-19 transmission. Moreover, the public should be advocating for the awareness of common COVID-19 symptoms such as fever, dry cough, sore throat and the urge to quarantine themselves based on self-test results. It is also advisable for the public to take vaccine shots to reduce the risk of getting and spreading the COVID-19 virus. All measures mentioned earlier shall be stringently implemented starting from the highly interacted states to contain the virus transmission.

5. Conclusion

Despite relying solely on the confirmed COVID-19 dataset, our approach visualised COVID-19 pandemic risk through NA interconnections between the Malaysian states. During the study period, state connectivity grew, identifying a higher degree of interaction among a few states, which may become the probable transmission keys. The SVR model could predict the number of COVID-19 cases and deaths in the future. With low MSE and RMSE values and reliable regressor scores, the SVR model's reliability was comparable to other predictive models; consequently, it can be advocated in future prediction studies. Nonetheless, our findings were confined to verified daily cases and fatalities only. To investigate the effect of other factors on the COVID-19 trend in Malaysia, this study via SVR needs more variables. Also, the SVR's prediction ability shall be compared with other prediction models, such as logistic regression, autoregressive integrated moving average (ARIMA), and long short-term memory (LSTM) in predicting COVID-19 cases. From this study, the NA and SVR model aided in understanding the spread of the virus among communities and provided early warning to the Malaysian authorities in preparing preventive measures based on the daily confirmed cases and fatalities.
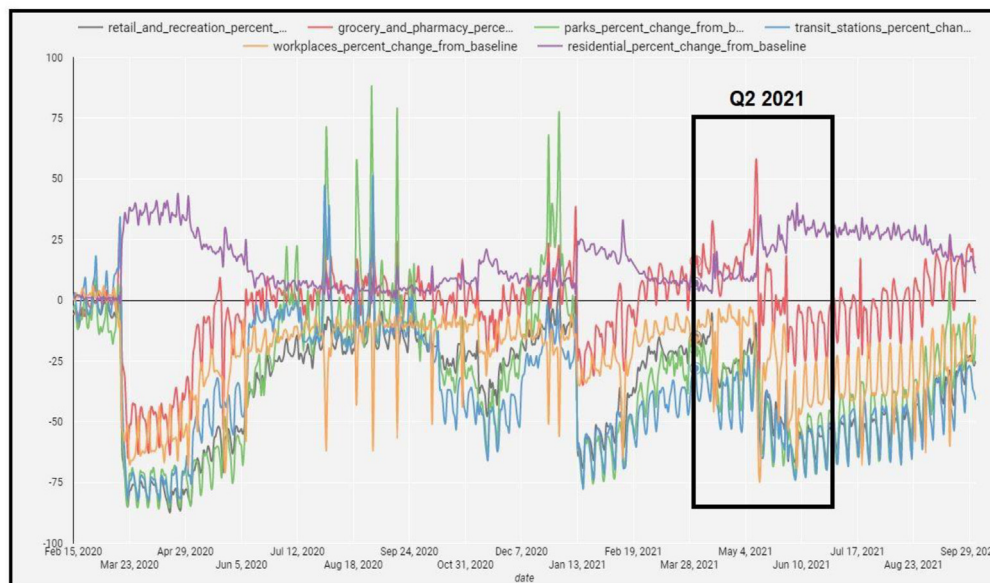
(a)

(b)

Fig. 6. Cumulative mobility data (%) over time for (a) Selangor and (b) Melaka in Q2 of 2021 (Google LLC, 2021).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors self-financially supported the present work.

References

[1] Z. Gao, S. Wang, J. Gu, et al., A community-level study on COVID-19 transmission and policy interventions in Wuhan, China, Cities 127 (2022) 103745.

[2] J.H. Hashim, M.A. Adman, Z. Hashim, et al., COVID-19 epidemic in Malaysia: Epidemic progression, challenges, and response, Front Public Heal. 9 (2021) 1–19.

[3] World Health Organization. COVID-19 - China, World Health Organization (WHO), 2020, pp. 1–3.

[4] A. Elengoe, COVID-19 outbreak in Malaysia, Osong Public Heal. Res. Perspec. 11 (2020) 93–100.

[5] V. Gallego, H. Nishiura, R. Sah, et al., The COVID-19 outbreak and implications for the Tokyo 2020 Summer Olympic Games, Travel Med. Infect. Dis. 34 (2020) 1–4.

[6] D. Parbat, M. Chakraborty, A python based support vector regression model for prediction of COVID19 cases in India, Chaos Solitons Fractals 138 (2020) 109942.

[7] A. Alsayed, H. Sadir, R. Kamil, et al., Prediction of epidemic peak and infected cases for COVID-19 disease in Malaysia,2020, Int. J. Environ Res Public Health 17 (2020) 1–15.

[8] N.A. Aziz, J. Othman, H. Lugova, et al., Malaysia's approach in handling COVID-19 onslaught: Report on the Movement Control Order (MCO) and targeted

screening to reduce community infection rate and impact on public health and economy, J. Infect. Public Health 13 (2020) 1823–1829.

[9] S.M. Zahraee, N. Shiwakoti, P. Stasinopoulos, Agricultural biomass supply chain resilience: COVID-19 outbreak vs. sustainability compliance, technological change, uncertainties, and policies, Clean Logist. Supply Chain 4 (2022) 100049.

[10] K. Omar, D. Baha Raja, N.A. Abdul Taib, et al., Risk stratification and assessment framework for international travel and border measures Amidst the COVID-19 pandemic – A Malaysian perspective, Travel Med. Infect. Dis. 47 (2022) 102318.

[11] Y.Y. Sun, M. Li, M. Lenzen, et al., Tourism, job vulnerability and income inequality during the COVID-19 pandemic: A global perspective, Ann Tour Res. Empir Insights 3 (2022) 100046.

[12] A. Barman, R. Das, P.K. De, Impact of COVID-19 in food supply chain: Disruptions and recovery strategy, Curr. Res. Behav. Sci. 2 (2021) http://dx.doi.org/10.1016/j.crbeha.2021.100017, Epub ahead of print.

[13] T.C. Voo, A. Ballantyne, C.J. Ng, et al., Public acceptability of COVID-19 control measures in Singapore, Hong Kong, and Malaysia: A cross-sectional survey, Int. J. Infect. Dis. 120 (2022) 51–58.

[14] Q. Zhang, J. Gao, J.T. Wu, et al., Data science approaches to confronting the COVID-19 pandemic: A narrative review, Phil. Trans. A (2021) 1–20.

[15] M.K.P. So, A. Tiwari, A.M.Y. Chu, et al., Visualizing COVID-19 pandemic risk through network connectedness, Int. J. Infect. Dis. 96 (2020) 558–561.

[16] J. Liu, J. Hao, Y. Sun, et al., Network analysis of population flow among major cities and its influence on COVID-19 transmission in China, Cities, 112.

[17] W. Jo, D. Chang, M. You, et al., A social network analysis of the spread of COVID-19 in South Korea and policy implications, Sci. Rep. 11 (2021) 1–10.

[18] M.R. Afrash, H. Kazemi-Arpanahi, M. Shanbehzadeh, et al., Predicting hospital readmission risk in patients with COVID-19: A machine learning approach, Inform. Med. Unlocked 30 (2022) 100908.

[19] A.A. Nayan, B. Kijsirikul, Y. Iwahori, Coronavirus disease situation analysis and prediction using machine learning: a study on Bangladeshi population, Int. J. Electr. Comput. Eng. 12 (2022) 4217–4227.

[20] I. Atik, Performance comparison of regression learning methods: COVID-19 case prediction for Turkey, Int. J. Mech. Eng 7 (2022) 6297–6306.

[21] A. Almalki, B. Gokaraju, Y. Acquaah, et al., Regression analysis for COVID-19 infections and deaths based on food access and health issues, Healthc 10 (2022) http://dx.doi.org/10.3390/healthcare10020324, Epub ahead of print.

[22] G. Csardi, T. Nepusz, The igraph software package for complex network research, InterJournal Complex Syst. (2006) 1–10.

[23] H. Akoglu, User's guide to correlation coefficients, Turkish J. Emerg. Med. 18 (2018) 91–93.

[24] B. Hu, H. Guo, P. Zhou, et al., Characteristics of SARS-CoV-2 and COVID-19, Nat. Rev. Microbiol. 19 (2021) 141–154.

[25] X. He, E.H.Y. Lau, P. Wu, et al., Temporal dynamics in viral shedding and transmissibility of COVID-19, Nat. Med. 26 (2020) 672–675.

[26] D. Chicco, M.J. Warrens, G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, PeerJ Comput. Sci. 7 (2021) 1–24.

[27] L. Cornelissen, E. André, Understanding the drivers of transmission of SARS-CoV-2, Lancet Infect. Dis. 21 (2021) 580–581.

[28] World Health Organization, COVID-19 in Malaysia Situation Report 14, vol. 2020, World Health Organization (WHO, 2020, pp. 1–4.

[29] N.T.P. Pang, A. Kamu, M.A. Mohd Kassim, et al., Monitoring the impact of movement control order (MCO) in flattening the cummulative daily cases curve of Covid-19 in Malaysia: A generalized logistic growth modeling approach, Infect. Dis. Model 6 (2021) 898–908.

[30] World Health Organization., COVID-19 in Malaysia Situation Report 22, vol. 2020, World Health Organization (WHO), 2020, pp. 1–8.

[31] Sabah @ a Glance, vol. 1, Department of Statistics Malaysia, 2016.

[32] J.T. Lim, K. Maung, S.T. Tan, et al., Estimating direct and spill-over impacts of political elections on COVID-19 transmission using synthetic control methods, PLoS Comput. Biol. 17 (2021) 1–15.

[33] World Health Organization, COVID-19 in Malaysia Situation Report 17, vol. 2020, World Health Organization (WHO, 2020, pp. 1–6.

[34] New Straits Times, CMCO in All But 3 Peninsular States, New Straits Time, 2020, pp. 1–2.

[35] N.H. Rodzi, Malaysia to impose mco for 2 weeks from 13th january in several states to curb covid-19 cases: muhyiddin, The Straits Times (2021) 1–8.

[36] Malaysiakini. Selangor, KL, johor, penang under CMCO strating march 5, Malaysiakini (2021) 6–11.

[37] Afrina Arfa, How Has MCO Affected the Malaysian Economy?, Taylor's University, 2021, pp. 1–6.

[38] Astro Awani, COVID-19: 4 Varian Membimbangkan Termasuk Dari UK, India Dikesan Di Malaysia, Astro Awani, 2021, pp. 1–8.

[39] M. Hassandarvish, Mass screening: Here's what selangor's mass Covid-19 testing tells us about the pandemic, Malay Mail (2021) 1–17.

[40] R.H.M. Price, C. Graham, S. Ramalingam, Association between viral seasonality and meteorological factors, Sci. Rep. 9 (2019) 929.

[41] W.B. Wan Nik, M.F. Ahmad, M.Z. Ibrahim, et al., Wind energy potential at east coast of peninsular Malaysia, Int. J. Appl. Eng. Res. 4 (2019) 9–16.

[42] S.B. Aw, B.T. Teh, G.H.T. Ling, et al., The covid-19 pandemic situation in malaysia: Lessons learned from the perspective of population density, Int. J. En-viron. Res. Public Health 18 (2021) http://dx.doi.org/10.3390/ijerph18126566, Epub ahead of print.

[43] Supriya Surendran, Covid-19 cuts top glove both ways, Edge Malays. (2020) 1–11.

[44] World Health Organization, COVID-19 in Malaysia Situation Report 45, vol. 2021, World Health Organization (WHO, 2021, pp. 1–9.

[45] Google LLC., Google Mobility Report, Google LLC, 2021, https://www.google.com/covid19/mobility/ (accessed 13th 2021).

[46] N. Zhao, W. Li, S.F. Zhang, et al., Network analysis of depressive symptoms among residents of wuhan in the later stage of the COVID-19 pandemic, Front Psychiatry 12 (2021) 1–11.

[47] R. An, N.H. Rodzi, Kl to place six selangor districts under first-tier MCO, avoiding widespread Covid-19 lockdowns, The Straits Times (2021) 1–11.

[48] A. Zainuddin, Klang valley hospitals on the brink of collapse, CodeBlue 2021 (2021) 1–12.

[49] M.A. Johansson, T.M. Quandelacy, S. Kada, et al., SARS-CoV-2 transmission from people without COVID-19 symptoms, JAMA Netw. Open 4 (2021) 1–8.

[50] J. Chu, A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain, Epub, 2021, http://dx.doi.org/10.1371/journal.pone.0249037, ahead of print.