Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Quantifying collective intelligence and behaviours of SARS-CoV-2 via environmental resources from virus’ perspectives

Ray-Ming Chen

School of Mathematics and Statistics, Baise University, 21, Zhongshan No.2 Road, Guangxi Province, China

ARTICLE INFO

Keywords:
SARS-CoV-2
Collective intelligence
Minimal spanning trees
Quantification
Virus’s perspectives

ABSTRACT

Collective intelligence of viruses is witnessed in many research articles. Most of the researches focus on the qualitative properties or observations. In this research, we model the behaviours and collective intelligence of SARS-CoV-2 by minimal spanning trees (MSTs), which specify the underlying mechanisms of resource allocation in the viral colony. The vertices of the trees are 50 states, DC and NYC in the USA. The weights of the edges are assigned by the reciprocal of the sum of cases or deaths of COVID-19. The types of trees are decided by the chosen 18 factors. We sample 304 time-series data and compute their MST-based auto-correlations for stability analysis. Then we perform correlated analysis and comparative analysis on these stable factors. Our results show MST approach fits the collective intelligence modelling very well; the total cases and total deaths over areas are highly correlated in terms of MSTs; and these stable factors have little to do with the geographical distance. The results also indicate the colonisation of SARS-CoV-2 is pretty mature and organised. Based on the results, for environmental or health policies, we should also turn our attention to the transmission routes that are independent of or far away from human population or densities. The viruses’ colonies might already exist in the wild in a large scale, not only in the populated or polluted cities. We shall build or conduct a monitoring system of their colonisation and survival techniques, in order to terminate, contain or live with their communities.

1. Introduction

Though the origins and ways of diffusion and transmission of COVID-19 regarding environmental factor-to-human or human-to-human at country level is still in debate and research (Bontempi, 2020a, 2020b; Chen, 2021a), there are some observations and tracking of the development regarding different waves of infection (Bontempi, 2021; bib_Coccia_2021Coccia, 2021a,b). Like ZIKA viruses (Liu et al., 2021), SARS-CoV-2 has mutated in various forms across the globe (Wise, 2021; Wang, 2021). It is vital the evolutionary mechanisms and behaviours of SARS-CoV-2, and their interaction with human and environment could be fully explored and revealed in order to terminate, contain or live with these viruses. So far, there are many approaches to explore the properties of viruses via genomics, epidemiology, immunology, virology, biology, biochemistry, environmental science, mathematical models, etc. In essence, they focus on two main levels: (1) micro-level (or reductionism), for example genetic decoding, modelling dynamical viral diffusion (Kucharski et al., 2020; Coccia, 2020c); molecular biology (Deichmann, 2017) and dynamics (He et al., 2020); and (2) macro-level (or holism), in particular the interaction or observation from the environmental factors: air pollution, climate change, etc.; demographic factors or the mixed factors - like Index c (Coccia, 2020a), which takes various factors into consideration to yield an index for the prediction of infectious diseases by studying the correlation between these factors and deaths of COVID-19 (Coccia, 2020b). There are some practical methods to tackle these issues: the study of multi-scale model and confinement strategies (Bellomo et al., 2020); the study of collective transmission (Liu et al., 2021); and the mechanisms of collective infectious units (CIUs) (Asher et al., 2019).

The behaviours of viruses are studied in various research literature. For some viruses, it has already shown that they have some cooperative and altruistic behaviours at the population level to weaken and attack our immune systems (Dolgin, 2019; Bunimovich et al., 2020). They also employ some social tricks. For example, phages could actively discuss and make decisions about their replicating and breaking strategies. This shows they have some degree of collective intelligence (CI). For survival, viruses would adapt to environmental changes in order to spread the viral infections via adaptive variants (Domingo et al., 1997; Miyashita et al., 2015). These variants are also witnessed by SARS-CoV-2 in recent development (CDC, 2020). However, most of these researches are
focusing on the micro-scale modelling, in particular the structures of genetic parts (Chen, 2020a, 2021b, 2021b; Zhao, 2020) - even some are tackling the interactions between human body systems and viruses' invasion (Mohammed, 1993; Smith, 2012). These researches focus more on the qualitative and static aspects and observations of the collective intelligence or behaviours in the viral colonies. In this article, we delve into the mechanisms of CI of SARS-CoV-2 via identifying the minimum spanning trees (MSTs) of macro-scale networks with respect to (w.r.t.) 18 different chosen factors. MST (Prim, 1957; Dijkstra, 1959; Chen, 2020b) is a way (or MST principle) to effectively allocate resources in a colony.

Collectively, SARS-CoV-2 is a colony with an aim to effectively build its kingdom. The underlying behaviours of their transmission and communication could be captured by MST. To begin with, the (52) vertices of the trees (or networks) are identified with the 50 USA states, the District of Columbia (DC) and New York City (NYC). The edges are represented by pairs of vertices. The weights assigned to the edges are decided by the reciprocal of the sum of magnitude of the values (COVID-19 cases and deaths). These weights basically witness the minimal paths or traces the viruses spread through. Up to this stage, we have constructed trees (or networks). Then we calculate and find the MSTs of these networks. These MSTs are then converted into incidence matrices.

Then we calculate the Hamming distances between these incidence matrices to measure the distance between MSTs, whose trees are formed by the 18 factors. Based on this, we perform stability analysis, which finds the auto-correlation and tells whether the factors create a series of stable MSTs respectively. The stable factors are then chosen for our further analytical processes: correlated analysis, which measures the cross-correlation among all the stable factors, and comparative analysis, which measures the cross-relation between the stable factors and the physical-distance factor. Our findings show that the stable factors exist (there are seven factors out of the 18 ones) - this indicates the suitability of MST principle for modelling the collective intelligence and behaviours of SARS-CoV-2. They also show that a very strong correlation between total cases of COVID-19 over areas of states and total deaths of COVID-19 over areas of states. The findings also show there is little correlation between physical distance and the chosen factors in terms of MST. It shows the viruses' density is the main factor that leads to the correlation between confirmed cases and deaths. A complete presentation of the results of this study is listed in section Results 3 and further explored for their environmental, health and social policy making in Section Conclusion 4.

2. Materials and methods

2.1. Research setting and sources

- Minimal Spanning Tree (MST)

MST is a concise method to trace the behaviours of SARS-CoV-2. The idea is these viruses exist for survival by their collective decisions in choosing the shortest paths to expand their kingdom. In order to find out their collective intelligence and decision processes behind COVID-19 pandemic, we locate (based on the data we could obtain) 18 factors (or MST factors) to pin down their surviving and expanding strategies. These factors would look into the viruses' strategies to survive and how they evolve. Here we choose 50 USA states, DC and NYC as our research targets (or vertices).

- SARS-CoV-2's perspectives of environmental factors

There are two macro-factors to look at: total cases and new cases. They indicate the statuses of viruses' colonisation. The total cases/deaths indicate the status of total population of viruses, while the new cases/deaths indicate the status of the speed of generation of their offspring. There are also macro-environmental factors (or resources), from the viruses' point of views, to be delved into: the size of the areas, human's population, human's population densities. The set of main factors we are exploring would be \( \{ \text{total cases, new cases} \} \times \{ \text{confirmed, deaths} \} \times \{ \text{area, population, population densities} \} \).
which of them are stable being MST, i.e., the channels have formed w.r.t.
the factor. This is done by separately analyse each factor and its time-
series results. We exploit adjacency matrices to represent their MSTs
and compute their Hamming distances. By observing these distances,
one locates the stable ones (w.r.t. time) and performs further analyses on
their relations between the stable ones with similar procedures, but the
MSTs are computed for pairing factors. This would reveal the underlying
similarities of the factors from the point of MST. Then we obtain the
geographical distances between states (a network) to yield their MST,
which is time-invariant, and calculate the distances between its MST and
the ones of all the stable factors via Hamming distancing on their ad-
jacency matrices.

2.2. Measures of variables

Based on the methodology mentioned in Section 2, we devise a set
of procedures to implement the methodology. According to our analytical
purposes, the implementations are divided into three parts: stability (or
auto-correlation), correlation (or cross-correlation), and comparative
implementations. The results of the implementations would be pre-
sented in Section 2.3.

- stability implementation

This implementation is to locate the factors that form the MST is a
stable and consistent way. There are several steps in the
implementation.

1. Download the data related to daily COVID-19 cases and deaths, the
data related to the area and populations for the states, DC and NYC -
these will serve as the vertices of the tree (or network) and the
fundamental bases for weights of edges. Let us name the downloaded
data DATA.

2. Filter out the non-essential data, add and extract new added factors
shown in Section 2; the sampled data lie between March 16, 2020
and January 13, 2021 (in total, 304 days). These data would serve
our source data. Let us name the data DATAcmb.

3. Choose the factor(s) one is interested in and extract them to a set
named DAY.

4. Assign the weights to the edges between the vertices (in our case 50
states, DC and NYC) and save them as EDGE_DAY.

5. Find the MST of EDGE_DAY and save them as mst_fac.

6. Relate each MST to an incidence matrix and save it as ADJ_FAC.
Collectively they are saved as ADJ_DAY.

7. Calculate the average Hamming distances between the MSTs for the
consecutive days and save them as HAM_FAC.

- correlation implementation

Now we have located the factors whose formed networks (dynamical
behaviours) largely or strictly follow the MST principle. Next, let us find
the correlations between all the following stable factors:

Factor 1: total cases;
Factor 3: total cases over areas;
Factor 9: total deaths;
Factor 11: total deaths over areas;
Factor 13: total deaths over populations;
Factor 15: total deaths over total cases;
Factor 17: total deaths over population densities.

This is done by comparing the pairwise MSTs for the pairs of the
factors. To implement it, the MSTs are converted into adjacency
matrices and the Hamming distances are applied to measure their
relations.

- comparative implementation

Finally, we want to find the relation between the stable factors and
the MST of geographical distances to see whether they are correlated.
Again, to implement it, we convert the MSTs into adjacency matrices
and apply the Hamming distance on them.

2.3. Model and data analysis

This section implements the processes in Section 2.2. The results are
produced from the databases or websites (Area, 2021a; Area, 2021b;
Population, 2021; NYC, 2021), including Google Map Api (Google,
2021) - for the physical distances between states, DC and NYC - fetched
and run by R program 4.0.3. The nominal numbers (or vertices) are the
states (including DC and NYC) presented in Table 1, where the acronyms
for the states could refer to Appendix A 7.1.

In the following analysis, we simply call the 50 states, DC and NYC
52 vertexes. The raw data are presented in blocks chronologically, where
each block consists of 52 rows - each of which is a vector, with length 18,
of values for the 18 factors. An extract of the raw data is presented at
Table 5 (or DATAcmb) in Appendix 7.2. The data are sampled from
March 16th, 2020 to January 13th, 2021.

Based on Table 5 and the method of assigning weights to edges in
Section 2, we construct the lengths for the edges. The results are pre-
sented at Table 6 in Appendix 7.3. The table is separated by 18 factors
(from left to right) and 304 days (from top to bottom), i.e., 18 times 304
blocks - each block contains a 52-by-52 matrix whose elements are the
weights for the edges between the 52 vertexes. For example, for factor 1
day 1 (i.e., March 16th, 2020), the weight assigned to the edge be
 tween state AK (or 1 by nominal ordering) and AL (or 2 by nominal
ordering) is \( \frac{1}{52} = 0.032 \); and for factor 18 at day 304 (i.e., January
13th, 2021), the weight between Alabama (or 2 by nominal ordering)
and West Virginia (or 51 by nominal ordering) is \( \frac{1}{18} = 0.056 \).

- stability analysis

Based on the constructed daily network with respect to factors in
Table 6, we run R programs to find their MSTs, which are represented by
incidence matrices (due to space limitation, the visualised MSTs and
incidence matrices are not presented explicitly here). Then we calculate
their Hamming distances between these incidence matrices. The results
are listed in Figs. 1, 2 and 3. In total, there are 18 plots spread over the
three figures - each plot is a time series of Hamming distances between
MST structures of the 52 vertexes - there are 303 pairs of consecutive
MST structures. There plots show the auto-correlation of the given fac-
tor. One would expect that when a factor is stable, the time series of
auto-correlation of the MST structures should substantially converge,
which is at least witnessed by a negative trend of the time series. Sta-
bility analysis is intended to find the stable variables in terms of MST
principle. It serves two purposes: to pin down those variables that are
stable and their dynamics are self-explanatory by MST principle; and to
filter out those inexplicable by MST principle. Stability analysis also
reveals the coordinated behaviours or collective intelligence (or strate-
gies) among the viruses. The results are further analysed in Section 3.

- correlated analysis

From Figs. 1, 2 and 3, we could then locate the factors that (almost)
strictly follow MST principle (to allocate their resources of the SARS-
CoV-2 colony). The stable and consistent factors of the 18 ones are
located and collected in the set: factor 1, factor 3, factor 9, factor 11,
factor 13, factor 15, factor 17. Then we compute their pairwise Ham-
ing distances of the converted MSTs (or incidence matrices). The re-
results are presented in Figs. 4, 5, and 6. There are \( \frac{7}{2} = 21 \) plots spread
over these figures. Each plot is a time series of Hamming distance be-
tween factor-based structures. It shows the cross-correlation between
two factors. This indicator shows the similarities between allocating strategies across different factors. Furthermore, the average Hamming distances between the factors are calculated and presented in Table 2.

Each value in the cell is the sum of all the 304 Hamming distances for each factor. The minimal spanning tree is computed and listed in Table 3. This table could average out all the 304 Hamming distance for each factor. The result is presented at the last row in Table 2. These values show the stabilising factors and the physical-distance factor. These values show the similarities between allocating strategies across different factors. Furthermore, the average Hamming distances between the factors are calculated and presented in Table 2.

One could observe that there is little correlation between the stable factors and the physical-distance factor. This indicator shows the similarities between allocating strategies across different factors. Furthermore, the average Hamming distances between the factors are calculated and presented in Table 2.

3. Results

This research focuses on understanding the behaviours of SARS-CoV-2 via delving into their collective intelligence. By identifying their CI via MSTs, we obtain a series of main results as follows:

1. From Figs. 1–3, the auto-correlated stability analysis for the time series of the structures shows that the factor total cases or deaths are overwhelmingly stabler than the ones of new cases/deaths. This indicates the viruses’ colonisation is already in a mature stage - the totality of their colonisation or kingdom is almost complete - despite the fact that their attacking structures or methods are still dynamically evolving and expanding. This might also suggest that they probably would coexist with humans begins and compete with us for quite a long time.

2. From Figs. 1–3, one also observes that the dynamical structures for new cases over areas and over human population are already in the polar stages, this might indicate their attacking structures or methods are still dynamically evolving and expanding. This might also suggest that they probably would coexist with human begins and compete with us for quite a long time.

3. Results

This research focuses on understanding the behaviours of SARS-CoV-2 via delving into their collective intelligence. By identifying their CI via MSTs, we obtain a series of main results as follows:

1. From Figs. 1–3, the auto-correlated stability analysis for the time series of the structures shows that the factor total cases or deaths are overwhelmingly stabler than the ones of new cases/deaths. This indicates the viruses’ colonisation is already in a mature stage - the totality of their colonisation or kingdom is almost complete - despite the fact that their attacking structures or methods are still dynamically evolving and expanding. This might also suggest that they probably would coexist with human begins and compete with us for quite a long time.

2. From Figs. 1–3, one also observes that the dynamical structures for new cases over areas and over human population are already in the polar stages, this might indicate their attacking strategies on based on some learning processes; moreover, one inspects the new cases over human population densities and finds out there is a high volatility of the dynamical structures - this might indicate the colonisation is still interacting and competing with with the environmental factor: human population densities.

3. From Fig. 2, the structures for total deaths over total cases is stabler than the ones for new deaths over new cases - this indicates the long-term strategy for viruses’ colonisation is well regulated and managed, but the short-term strategy is to keep on exploring various possibilities for transmission.

4. From Figs. 4–6 or Table 2, one finds the best correlated pair of factors w.r.t. MST is “total cases COVID-19 over areas of states” and “total deaths of COVID-19 over areas of states”. This indicates the viruses’ density is the top priority for their collective decision making in

### Table 1
Numbering the 50 acronymous states, District of Columbia and New York City.

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|----|----|
| AK | AL | AR | AZ | CA | CO | CT | DC | DE | FL | GA |
| HI | IA | ID | IL | IN | KS | KY | LA | MA | MD | ME |
| MI | MN | MO | MS | MT | NC | ND | NE | NJ | NM | NY |
| NV | NY | OH | OK | OR | PA | RI | SC | SD | TN | TX |

| factor 1 | factor 3 | factor 9 | factor 11 | factor 13 | factor 15 | factor 17 |
|----------|----------|----------|-----------|-----------|-----------|-----------|
| factor 1 | 0        | 67.77    | 67.86     | 67.78     | 69.67     | 75.31     | 87.82     |
| factor 3 | 67.77    | 0        | 15.68     | 0.07      | 5.37      | 9.17      | 101.22    |
| factor 9 | 67.86    | 15.68    | 0         | 15.69     | 13.44     | 21.47     | 90.19     |
| factor 11| 67.78    | 0.07     | 15.69     | 0         | 5.38      | 9.17      | 101.22    |
| factor 13| 69.67    | 5.37     | 13.44     | 0         | 0         | 10.38     | 98.61     |
| factor 15| 75.31    | 9.17     | 21.47     | 9.17      | 0         | 100.20    | 90.19     |
| factor 17| 87.82    | 101.22   | 90.19     | 101.20    | 98.61     | 100.20    | 0         |
| GeoDis   | 99.53    | 101.22   | 90.19     | 101.20    | 101.90    | 100.59    | 100.02    |
constructing their viruses’ kingdom. It also shows the underlying colony of the viruses is already very mature in forming their attacking and propagating techniques; 5. From the last row of Table 2, it has shown the transmitted paths does not follow the MST of physical distance. It shall indicate they spreading is not conducted in a rippling way, but with a widespread transmission - an indication that other environmental factors like wind or climate change might contribute to the widespread (Coccia, 2020d, 2020f, 2021b). But our results probably hint a step further that the widespread might also happen in some less polluted cities. This shall declare that a sole physical lockdown policy or social distancing might produce little effect on the transmission of SARS-CoV-2.

4. Conclusion

By observing viruses’ collective intelligence through MST with various environmental factors from the viruses’ point of views, we obtain a series of results regarding their correlation and propagating behaviours. These findings should help us adopt some environmental policies, health policies and social policies in decision making. For environmental policies: The finding also suggests the viruses’ transmission has little to do with geographical distance, i.e., the spread of viruses might resort to other media, not mainly through human-to-human or distance-related pollution. This might indicate there might exist other environments with sparse human population that incubate the viruses colonies. This is also indirectly witnessed by the fact that the confirmed cases and deaths are highly correlated via the environmental factor area”. This indicates the densities of viruses matter, not the density of human population. This might imply they already form their communities and colonies that are largely independent of interaction with human environment. This might also hint we should terminate the viruses largely in the environment, i.e., their existence might be accompanied by other media or carriers. For health policies: From result One, there is an evidence that the viruses’ colonisation is pretty stable. This shall indicate that the viruses might coexist with us for some years to come and might evolve or mutate by its own approaches. In our future research, one should keep track of their colonisation, their identities, and their colonisations are revealed or explored in latter research. Another limitation is the follow-up work. Since this research is performed when the global vaccination is still underway, a follow-up research should be also performed to compare the shift of their MST structures after the majority of people are vaccinated.

Author contribution

The author declares that he is the sole author of this manuscript and has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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

This work is supported by the Humanities and Social Science Research Planning Fund Project under the Ministry of Education of China (Grant No. 20XJAGAT001).

7. Appendix

7.1. Appendix A: Full names for acronyms

Table 4

| AK   | AL   | AR    | AZ     | CA    | CO     | CT     |
|------|------|-------|--------|-------|--------|--------|
| Alaska | Alabama | Arkansas | Arizona | California | Colorado | Connecticut |
| DC    | DE    | FL     | GA     | HI    | IA     | ID     |
| District of Columbia | Delaware | Florida | Georgia | Hawaii | Iowa | Idaho |
| IL    | IN    | KS     | KY     | LA    | MA     | MD     |
| Illinois | Indiana | Kansas | Kentucky | Louisiana | Massachusetts | Maryland |
| ME    | MI    | MN     | MO     | MS    | MT     | NC     |
| Maine | Michigan | Minnesota | Missouri | Mississippi | Montana | North Carolina |
| ND    | NJ    | NH     | NJ     | NM    | NV     | NY     |
| North Dakota | Nebraska | New Hampshire | New Jersey | New Mexico | Nevada | New York |
| NYC   | OH    | OK     | OR     | PA    | RI     | SC     |
| New York City | Ohio | Oklahoma | Oregon | Pennsylvania | Rhode Island | South Carolina |
| SD    | TN    | TX     | UT     | VA    | VT     | WA     |

(continued on next page)
Table 4 (continued)

|    | AK  | AL  | AR  | AZ  | CA  | CO  | CT  |
|----|-----|-----|-----|-----|-----|-----|-----|
| South Dakota | Tennessee | Texas | Utah | Virginia | Vermont | Washington |
| WI | WV | West Virginia | WY | Wyoming |

7.2. Appendix B: raw data with respect to 304 dates and 18 factors

Table 5

304 batches of daily raw data of 52 vertexes (states) against 18 factors

| no. | date       | state | factor 1 | factor 2 | ... | factor 17 | factor 18 |
|-----|------------|-------|----------|----------|-----|-----------|-----------|
| 1   | 03/16/2020 | AK    | 1        | 0        | ... | 0.00      | 0.00      |
| 2   | 03/16/2020 | AL    | 29       | 7        | ... | 0.00      | 0.00      |
| 3   | 03/16/2020 | AR    | 17       | 1        | ... | 0.00      | 0.00      |
| ... | ...        | ...   | ...      | ...      | ... | ...       | ...       |
| 50  | 03/16/2020 | WI    | 53       | 26       | ... | 0.00      | 0.00      |
| 51  | 03/16/2020 | WV    | 0        | 0        | ... | 0.00      | 0.00      |
| 52  | 03/16/2020 | WY    | 3        | 0        | ... | 0.00      | 0.00      |
| 1   | 03/17/2020 | AK    | 3        | 2        | ... | 0.00      | 0.00      |
| 2   | 03/17/2020 | AL    | 39       | 10       | ... | 0.00      | 0.00      |
| 3   | 03/17/2020 | AR    | 24       | 7        | ... | 0.00      | 0.00      |
| ... | ...        | ...   | ...      | ...      | ... | ...       | ...       |
| 50  | 03/17/2020 | WI    | 88       | 35       | ... | 0.00      | 0.00      |
| 51  | 03/17/2020 | WV    | 1        | 1        | ... | 0.00      | 0.00      |
| 52  | 03/17/2020 | WY    | 11       | 8        | ... | 0.00      | 0.00      |
| 1   | 01/13/2021 | AK    | 49,203   | 406      | ... | 545.95    | 4.85      |
| 2   | 01/13/2021 | AL    | 410,995  | 3147     | ... | 163.61    | 5.31      |
| 3   | 01/13/2021 | AR    | 262,020  | 2467     | ... | 197.72    | 3.07      |
| ... | ...        | ...   | ...      | ...      | ... | ...       | ...       |
| 50  | 01/13/2021 | WI    | 558,020  | 2467     | ... | 197.72    | 3.07      |
| 51  | 01/13/2021 | WV    | 104,392  | 2467     | ... | 197.72    | 3.07      |
| 52  | 01/13/2021 | WY    | 48,289   | 1189     | ... | 56.59     | 1.25      |

7.3. Appendix C: weights of edges from factor 1 to factor 18 for 304 days

Table 6

weights of edges for 52 vertexes (states) per day per factor

|      | 1    | 2    | ... | 51   | 52   |            | 1    | 2    | ... | 51   | 52   |            | 1    | 2    |
|------|------|------|-----|------|------|-----------|------|------|-----|------|------|-----------|------|------|
| factor 1 | 0.333 | 0.032 | ... | 0.500 | 0.200 | ...       | 1.000 | 1.000 | ... | 1.000 | 1.000 | ...       | 1.000 | 1.000 |
| factor 2 | 0.032 | 0.017 | ... | 0.033 | 0.030 | ...       | 1.000 | 1.000 | ... | 1.000 | 1.000 | ...       | 1.000 | 1.000 |
| ...   | ...   | ...   | ... | ...   | ...   | ...       | ...   | ...   | ... | ...   | ...   | ...       | ...   | ...   |
| factor 18 | ...   | ...   | ... | ...   | ...   | ...       | ...   | ...   | ... | ...   | ...   | ...       | ...   | ...   |
7.4. Appendix D: Autocorrelated Hamming distance between incidence matrices

Fig. 1. Stability analysis for factor 1 to factor 8 (read from left to right, row-wise).
Fig. 2. Stability analysis for factor 9 to factor 16 (read from left to right, row-wise).
Fig. 3. Stability analysis for factor 17 to factor 18 (read from left to right, row-wise).
7.5. Appendix E: Cross-correlated Hamming distance between incidence matrices

Fig. 4. cross-relations for factor 1 and factor 3, 9, 11, 13, 15, 17 and for factor 3 and factor 9,11 (read from left to right, row-wise).
Fig. 5. Cross-relations for factor 3 and factors 13, 15, 17; for factor 9 and factors 11, 13, 15, 17; for factor 11 and factor 13 (read from left to right, row-wise).
Fig. 6. cross-relations for factor 11 and factor 15,17; for factor 13 and factor 15,17; for factor 15 and factor 17 (read from left to right, row-wise).

Credit author statement

Ray-Ming Chen is the sole author.

References

Annual Estimates of the Resident Population for the United States, Regions, States, and the District of Columbia: April 1, 2010 to July 1, 2020 (NST-EST2020).

Asher, Leeks, Sanjuan, Rafael, West, Stuart A., 2019. The evolution of collective infectious units in viruses. Virus Res. 265, 94–101. https://doi.org/10.1016/j.virusres.2019.03.013. ISSN 0168-1702.

Bellomo, Nicola, Bingham, Richard, Mark, A., Chaplain, J., Dosi, Giovanni, Guido, Forni, Knopoff, Damian A., Lowengrub, John, Twarock, Reidun, Enrica Virgillito, Maria, 2020. A multiscale model of virus pandemic: heterogeneous interactive entities in a globally connected world. Math. Model Methods Appl. Sci. 30 (8), 1591–1651. https://doi.org/10.1142/S0218202520500323.

Bontempi, E., 2021. The europe second wave of COVID-19 infection and the Italy "strange" situation. Environ. Res. 193, 110476. https://doi.org/10.1016/j.envres.2020.110476. ISSN 0013-9351.

Bontempi, Elza, Vergalli, Sergio, Squazioni, Flaminio, 2020. Understanding COVID-19 diffusion requires an interdisciplinary, multi-dimensional approach. Environ. Res. 188, 109814. https://doi.org/10.1016/j.envres.2020.109814. ISSN 0013-9351.

Buninovich, L., Shu, L., 2020. Local immunodeficiency: role of neutral viruses. Bull. Math. Biol. 82 (11), 140. https://doi.org/10.1007/s11538-020-00813-z.

Centers for Disease Control and Prevention, 2020. Interim: Implications of the Emerging SARS-CoV-2 Variant VOC 202012/01.

Chen, Ray-Ming, 2020a. Randomness for nucleotide sequences of SARS-CoV-2 and its related subfamilies. Computational and Mathematical Methods in Medicine vol.. https://doi.org/10.1155/2020/8819942.
Chen, Ray-Ming, 2020b. A unified proof for finding a minimal spanning tree. In: ICoMS 2020: Proceedings of the 2020 3rd International Conference on Mathematics and Statistics, June 2020, pp. 34–41. https://doi.org/10.1145/3409915.3409925.

Chen, Ray-Ming, 2021a. On COVID-19 country containment metrics: a new approach. J. Decis. Syst. https://doi.org/10.1080/12460125.2021.1866625.

Chen, Ray-Ming, 2021b. Journal of Bioinformatics and Computational Biology. https://doi.org/10.1142/S0219716521500116.

Coccia, M., 2020. An index to quantify environmental risk of exposure to future epidemics of the COVID-19 and similar viral agents: theory and practice. Environ. Res. 191 (December), 110155. https://doi.org/10.1016/j.envres.2020.110155.

Coccia, M., 2020a. How (Un)sustainable environments are related to the diffusion of COVID-19: the relation between coronavirus disease 2019, air pollution, wind resource and energy. Sustainability 12, 9709. https://doi.org/10.3390/su12229709.

Coccia, M., 2020b. How do low wind speeds and high levels of air pollution support the spread of COVID-19? Atmospheric Pollution Research 12 (1), 437–445. https://doi.org/10.1016/j.apr.2020.10.002.

Coccia, M., 2020c. How (Un)sustainable environments are related to the diffusion of COVID-19: the relation between coronavirus disease 2019, air pollution, wind resource and energy. Sustainability 12, 9709. https://doi.org/10.3390/su12229709.

Coccia, M., 2021a. Effects of Air Pollution on COVID-19 and Public Health. Research Article-Environmental Economics-Environmental Policy, Research Square. https://doi.org/10.21203/rs.3.rs-41354/v1; https://www.researchsquare.com/article/rs41354/v1.

Coccia, M., 2021b. Effects of the spread of COVID-19 on public health of polluted cities: results of the first wave for explaining the deja vu in the second wave of COVID-19 pandemic and epidemics of future viral agents. Environ. Sci. Pollut. Res. 28, 19147–19154. https://doi.org/10.1007/s11356-020-11662-7.

Coccia, M., 2021c. The effects of atmospheric stability with low wind speed and of air pollution on the accelerated transmission dynamics of COVID-19. Journal: Environmental International Journal of Environmental Studies GENV 78 (1), 1–27. https://doi.org/10.1080/00207233.2020.1802937.

Coccia, M., 2021d. How do low wind speeds and high levels of air pollution support the spread of COVID-19? Atmospheric Pollution Research 12 (1), 437–445. https://doi.org/10.1016/j.apr.2020.10.002.

Deichmann, U., 2017. Hierarchy, determinism, and specificity in theories of development and evolution. Hist. Philos. Life Sci. 39 (4), 33. https://doi.org/10.1007/s40656-017-0160-3.

Dijkstra, E.W., 1959. A note on two problems in connexion with graph. Numer. Math. 1, 269–271. https://doi.org/10.1007/BF01386390.

Dolgin, Elie, 2019. The secret social lives of viruses. Nature 570, 290–292. https://doi.org/10.1038/s41586-019-01880-6.

Domingo, E., Holland, 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. https://doi.org/10.1146/annurev.micro.51.1.151.

Dolgin, Elie, 2019. The secret social lives of viruses. Nature 570, 290–292. https://doi.org/10.1038/s41586-019-01880-6.

Dolgin, Elie, 2019. The secret social lives of viruses. Nature 570, 290–292. https://doi.org/10.1038/s41586-019-01880-6.

Dolgin, Elie, 2019. The secret social lives of viruses. Nature 570, 290–292. https://doi.org/10.1038/s41586-019-01880-6.

Dolgin, Elie, 2019. The secret social lives of viruses. Nature 570, 290–292. https://doi.org/10.1038/s41586-019-01880-6.

Dolgin, Elie, 2019. The secret social lives of viruses. Nature 570, 290–292. https://doi.org/10.1038/s41586-019-01880-6.