IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 16, 2021, accepted August 28, 2021, date of publication September 22, 2021, date of current version October 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3114364

Application of Hidden Markov Models to Analyze, Group and Visualize Spatio-Temporal COVID-19 Data

SHANGLIN ZHOU[1], (Member, IEEE), PAOLO BRACA[2], (Senior Member, IEEE), STEFANO MARANO[3], (Senior Member, IEEE), PETER WILLETT[4], LEONARDO M. MILLEFIORI[2], (Member, IEEE), DOMENICO GAGLIONE[2], (Member, IEEE), AND KRISHNA R. PATTIPATI[4], (Life Fellow, IEEE)

[1]Department of Computer Science and Engineering, University of Connecticut (UCONN), Storrs, CT 06269, USA
[2]Research Department, NATO STO Centre for Maritime Research and Experimentation, 19126 La Spezia, Italy
[3]Department of Information and Electrical Engineering and Applied Mathematics (DIEM), University of Salerno, 84084 Fisciano, Salerno, Italy
[4]Department of Electrical and Computer Engineering, University of Connecticut (UCONN), Storrs, CT 06269, USA

Corresponding author: Shanglin Zhou (shanglin.zhou@uconn.edu)

The work of Peter Willett was supported by Air Force Office of Scientific Research (AFOSR) under Contract FA9500-18-1-0463. The work of Krishna R. Pattipati was supported in part by the U.S. Office of Naval Research; and in part by the U.S. Naval Research Laboratory under Grant N00014-18-1-1238, Grant N00014-21-1-2187, and Grant N00173-16-1-G905.

ABSTRACT The coronavirus epidemic (COVID-19) is a public health challenge due to its rapid global spread. Its unprecedented speed and pervasiveness have led many governments to implement a series of countermeasures, such as lock-downs, stopping/restricting travels, and mandating social distancing. To control and prevent the spread of COVID-19, it is essential to understand the latent dynamics of the disease's evolution and the effectiveness of the intervention policies. Hidden Markov models (HMMs) capture both randomnesses in spatio-temporal dynamics and uncertainty in observations. In this paper, we apply an overall HMM that, based on multiple nations' COVID-19 data including the USA, several European countries, and countries that have strict control policies, explore different types of observations, and we use it to infer the severity state on small geographical states or regions in the USA and Italy as test cases. Further, we aggregate the severity level of each region over a fixed time period to visualize the time evolution and propagation across regions. Such an analysis and visualization provide suggestions for interventions and responses in a calibrated manner. Results from HMM modeling are consistent with what is observed in Italy and the USA and these models can serve as visualization and proactive decision support tools to policymakers.

INDEX TERMS SARS-CoV-2, hidden Markov model, pandemic tracking, pandemic prediction, Viterbi decoding, aggregation, visualization.

I. INTRODUCTION

The coronavirus epidemic (COVID-19) is arguably one of the most life-threatening and economic disasters of the 21^{st} century, as, in addition to deep economic suffering, it has caused more than 198 million cases worldwide and over 4 million deaths, as of July 31, 2021. It spreads rapidly due (for example) to its long incubation period (median of 5.2 days) and asymptomatic spreading, and it impacts the respiratory tract, possibly leading to pneumonia and acute respiratory disease, long recovery time, and death. The disease's spread across several continents has affected a large swath of the world's population, and the World Health Organization (WHO) declared it a global public health emergency on March 11, 2020. Recently, new strains of COVID-19 that spread even more rapidly have been identified. Different countries and regions have adopted strict policies and restrictive measures to contain the virus, such as mandatory 14-day monitored home quarantine for travelers, implementation of social distancing measures, curtailed international air travel, and targeted lock-downs. In this vein, it is salient to extract the spatio-temporal evolution of the disease from the observed

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

data reports to understand the similarities and differences in the spreading patterns of COVID-19 across geographic states and regions.

The problem addressed here is to track, probabilistically, the severity of the epidemic in a geographic region, based on uncertain data (e.g., daily and cumulative infections, daily deaths) as a guide for policy decisions. We exploit the Hidden Markov Model (HMM) formalism because of its ability to capture the dynamics of the latent (unobservable) state variables of a system (in our case the local severity state of the disease) that are probabilistically related to noisy observations. Moreover, HMM is a parametric model characterized by state transition probabilities. It helps in understanding how likely it is that a geographic region transits from one severity state to another. Such an analysis also aids in discovering similarities in disease spreading patterns across geographic regions via severity state-based visualization, which may provide insights into virus propagation and may help to evaluate policies and plans to contain the disease.

In this paper, we integrate different observation modalities (the aforementioned daily infections and daily deaths) into a vector of uncertain observation sequences to learn the time evolution of the spread of COVID-19. Specifically, we learn the HMM model across multiple nations, which includes the USA, some European countries, and several countries with stricter control policies (Singapore, Taiwan, New Zealand, Australia). This complexity of the data introduces diversity of information into the model. We then use the model to infer the hidden state sequences of small regions or states in the USA and Italy to quantify their COVID-19 severity levels. Moreover, we aggregate the HMM state sequence of each geographical state or region over a fixed time period (say, a month) and perform grouping based on the mean value of the state sequences over that period to glean the COVID-19 disease severity and spreading patterns.

The remainder of this paper is organized as follows. Section II reviews the application of HMM in biological studies and existing analysis on COVID-19. Section III describes the basic concepts of HMM modeling with other concepts leveraged in the paper, with details relegated to Appendix A. Section IV describes the datasets and the derived features for analysis and visualization. Section V illustrates the proposed framework, including details on learning the HMM models, the severity state grouping procedure, and the computational results. Discussion and concluding remarks are provided in Sections VI and VII. Some material is deferred to appendices not to fragment the exposition.

II. RELATED WORK

A. APPLICATION OF HMM IN BIOLOGICAL STUDIES

Churchill [1] was the first to introduce Hidden Markov Models to computational biology. Since then, HMM has become a promising tool for various biological problems [2]. Gene transmission can be viewed as a 2D hidden Markov process, which has led to such advances as the Elston-Stewart algorithm [3] and the Lander-Green algorithm [4], for genetic reconstruction and for extracting the inheritance information. In biological sequence analysis, authors in [5] characterized a biological sequence, such as protein and DNA, by an HMM, using the sequence of monomers as observations and the match and gap occurring in each site as a hidden state. HMMs are extensively leveraged for single sequence pattern recognition [6]–[8]. Similarly, by considering the genomic DNA as the observed symbols, and the gene structures and the characteristic subsequences as hidden states, HMM can also be applied to gene finding and feature discovery domains, such as sequence pattern extraction, motif search, and non-coding RNA [9]. Epidemiological analysis, prediction, and surveillance are also popular application areas of hidden Markov models [10], [11]. Progression of a chronic disease [12], as well as epidemic dynamics [13], can be well-modeled using HMMs because the dispersion of communicable pathogen in the population satisfies a temporal-spatial Markov dependence [14], so that the future status of a disease only depends on the current state, but not the past states.

B. EPIDEMIOLOGICAL MODELS OF COVID-19 PANDEMIC

The Susceptible-Infected-Recovered (SIR) model is the first and the most popular model applied to COVID-19 for analyzing the space-time dependence of the disease. In [15], the authors predicted an exponential growth of cases based on the SIR model, while authors in [16] pointed out that when the epidemic peaks, death rates would have an exponential growth following the power-law behavior based on the SIR model. The authors in [17] made modifications to the standard SIR/Susceptible-Exposed-Infected-Recovered (SEIR) epidemiological models, and included social distance into the analysis for predicting the disease trends. The author in [18] combined the SIR model with a statistical learning model to analyze the importance of lock-downs on COVID-19 progression. In [19], the authors proposed a Bayesian sequential estimation and forecasting algorithm tailored to the stochastic SIR model to estimate the state of the epidemic. A deterministic compartmental model was proposed in [20]. It takes the clinical progression, epidemiological status, and intervention measures into account. The authors of [21] proposed an extended SEIR model by considering transmission across cities, but without taking into account the adopted control measures. In [22], a difference equation (DE) model to predict the trends in COVID-19 epidemic progression is developed. The authors of [23] proposed a quickest detection model — the mean-agnostic sequential test (MAST) — to study the onset of COVID-19 pandemic waves. In [24] a discrete-time stochastic model to estimate the effect of travel restrictions on COVID-19 has been introduced. The authors of [25] proposed a stochastic compartmental model to capture the effects of intervention measures. Because of the different nature of the epidemic phases (meaning, for example that the SIR parameters can change from recovery to explosive growth in infections), the authors in [26] and [27] focused on an easily-implementable version of Page's CUSUM quickest-detection test for composite hypothesis scenarios.

III. BASIC THEORIES AND CONCEPTS

A. BASIC THEORY OF HIDDEN MARKOV MODELS

A Markov process is a stochastic process in which the conditional probability distribution of a future state depends only on the current state while, given the current state, it is conditionally independent of the past. A *hidden* Markov model (HMM) is an extension in which the state sequence is latent and is only revealed indirectly via a probabilistic mechanism [28]. In other words, an HMM is a doubly embedded stochastic process with an underlying stochastic dynamics (e.g., the severity state of COVID-19 pandemic in a geographic region or state) that, although unobservable (hidden), can be inferred through the observation of another set of related stochastic processes (e.g., infection rates, deaths). HMMs provide the required theoretical machinery to learn a probabilistic model from data; and by letting an HMM analyze COVID-19 observations in a region, it is possible to estimate which severity state currently characterizes it, and also to predict the most likely evolution of the severity state over time [29].

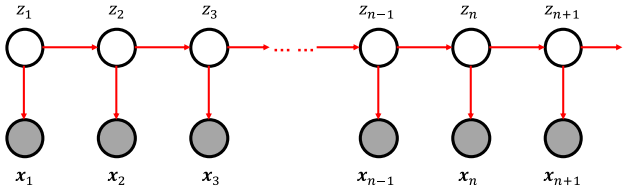


FIGURE 1. Hidden Markov Model.

As shown in Figure 1, $\{z_t\}_{t=1}^T$ is the (hidden, latent) state sequence and $\{\mathbf{x}_t\}_{t=1}^T$ is the observation sequence generated by the hidden state sequence, where T is the total number of observations. The hidden states, denoting the severity level of the pandemic, are discrete, but the observation sequences, representing a normalized number of infections per 100,000 population, normalized number of daily deaths per 5,000,000 population, etc., are ratios of integers, and are assumed to be continuous (the rationale for the choice of normalization for deaths is explored in detail in Section VI-A). The state sequence $\{z_t\}_{t=1}^T$ satisfies the requirements of a Markov process and the observation sequence $\{\mathbf{x}_t\}_{t=1}^T$ is a function of the state sequence $\{z_t\}_{t=1}^T$ as detailed below.

Each HMM is characterized by $\boldsymbol{\lambda} = (N, \boldsymbol{\pi}, A \text{ and } \{b_j(\mathbf{x})\})$, defined as follows [30]:

1) The number of states in the model, N. Then, the set of severity states can be written as $S = \{S_1, S_2, \ldots, S_N\}$. If we denote the state of model at time t as z_t, then $z_t \in \{S_1, S_2, \ldots, S_N\}$
2) The initial probability vector $\boldsymbol{\pi}$, which indicates how likely it is for a new input sequence to start in a given hidden state. Each element π_i represents the unconditional probability of being in state S_i at time $t = 1$. The sum of π_i's must, naturally, be unity.

$$\boldsymbol{\pi} = (\pi_i)_N \text{ where } \pi_i = P(z_1 = S_i),\ 1 \le i \le N. \quad (1)$$

3) A state transition probability matrix A, whose elements indicate the conditional probability of transitioning from one hidden state to another. An element a_{ij} in this matrix represents the probability of transitioning from state S_i at time t to state S_j at time $t + 1$. All row sums of A are unity.

$$A = [a_{ij}]_{N\times N} \quad \text{where } a_{ij} = P\left(z_{t+1} = S_j | z_t = S_i\right),\ 1 \le i, j \le N. \quad (2)$$

4) The emission probability density $b_j(\mathbf{x})$ in state S_j indicates the observation likelihood given the hidden state. It can be any parametric density with parameter $\boldsymbol{\theta}$ conditioned on the current (hidden) state. Here, we assume $b_j(\mathbf{x})$ to be Gaussian:

$$b_j(\mathbf{x}) = \mathcal{N}(\mathbf{x}, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j), \quad \text{where } 1 \le j \le N, \quad (3)$$

and where $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$ are the mean vector and covariance matrix associated with hidden state S_j. The Gaussian assumption is convenient for estimation, and its unboundedness causes no appreciable concern since generally $\boldsymbol{\mu}_j \gg \sqrt{\boldsymbol{Diag}(\boldsymbol{\Sigma}_j)}\boldsymbol{e}$, where $\boldsymbol{e}$ is a column vector of ones, $\boldsymbol{Diag}(\boldsymbol{\Sigma}_j)$ means diagonal elements of Σ_j, and "$\gg$" operates entry-by-entry.

Once an HMM is specified as $\boldsymbol{\lambda} = (N, \boldsymbol{\pi}, A, \{b_j(\mathbf{x})\})$, it can be used to (1) generate an observation sequence $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$; (2) learn the parameters from observed data; (3) compute the likelihood of observing a given sequence, given model parameters; and (4) determine the most likely evolution of the state sequence over time. Details are provided in Appendix A.

B. HAMPEL FILTER FOR OUTLIER REMOVAL

It is common to have outliers in the real-world datasets [31]. In the case of COVID-19 data, this is especially true during weekends (when cases are accumulated and reported en masse afterwards) and in the case of late reporting of deaths. Thus, outlier detection and interpolation are necessary. In our experiments, we used the Hampel filter, which is more robust than the standard "three-sigma" rule [32]. Hampel filtering consists of a sliding window of configurable width that slides across the time series, within which the median and a robustified version of the standard deviation are calculated [33], [34]. If the point of interest lies more than a prescribed multiple of the standard deviation from the window's median, then it is identified as an outlier and is replaced by the median [35].

Given a time series $\{x_1, x_2, \ldots, x_n\}$ and a sliding window with length L, the median m_i, and the standard deviation σ_i used in Hampel filtering are defined as:

$$\begin{aligned} m_i &= \text{median}(x_{i-L}, x_{i-L+1}, \ldots, x_i, \ldots, x_{i+L-1}, x_{i+L}), \\ \sigma_i &= \kappa \cdot \text{median}(|x_{i-L} - m_i|, \ldots, |x_{i+L} - m_i|), \\ &\text{where } \kappa = \frac{1}{\sqrt{2}\,\text{erfc}^{-1}(1/2)} \approx 1.4826, \\ &\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \end{aligned} \quad (4)$$

The quantity σ_i/κ is referred to as the median absolute deviation (MAD). Any sample x_i for which $|x_i - m_i| > c\,\sigma_i$, for some predefined threshold $c > 0$, is replaced by m_i.

IV. DESCRIPTION OF COVID-19 DATASETS AND PREPROCESSING

A. SOURCES OF DATA

In this section, we describe three datasets analyzed in this work. The first dataset pertains to the spread of COVID-19 in the United States, the second is related to the spread of COVID-19 in Italy, and the last is concerned with the spread of COVID-19 in other places of the world. The data at the national level for all countries are extracted from the last dataset for building the HMM models. Regions or geographical states data for the USA and Italy are extracted from the first two datasets for prediction purposes.

1) CORONAVIRUS (COVID-19) DATA IN THE UNITED STATES

For HMM analysis at the state level, we use the datasets from "The COVID Tracking Project[1]", a volunteer organization launched by *The Atlantic*. It reports the available data from all 50 states, 5 territories, and the District of Columbia in the United States until March 7, 2021. The starting date of observation sequences in each state and territory can be different because each state began to track the COVID-19 evolution only after cases were detected in it. The dataset is comprised of both cumulative and daily positive cases, hospitalizations, deaths, the daily and cumulative intensive care unit (ICU) occupancy, negative polymerase chain reaction (PCR) tests and negative antibody tests, daily and cumulative ventilator cases, and both daily and cumulative COVID-19 tests administered. Since we compare across countries, our analysis focuses on daily positive cases and deaths because these are present in all datasets.

2) CORONAVIRUS (COVID-19) DATA IN ITALY

The regional level COVID-19 data for Italy are downloaded from the repository of the Italian Department of Civil Protection.[2] The starting date is February 24, 2020, for all the Italian regions. The repository is updated daily at 6:00 PM local time zone. The datasets report cumulative and daily number of cases, deaths, and hospitalized cases of each of the 20 regions (the region Trentino Alto Adige reports data for each of the autonomous provinces of Bolzano and Trento, resulting in a total of 21 timeseries). Other useful information from this repository includes the current home confinement cases, the current positive cases (hospitalized patients plus home confinement), the recovered cases, number of people tested, and daily admissions to the intensive care.

3) CORONAVIRUS (COVID-19) DATA FOR OTHER COUNTRIES OF THE WORLD USED TO LEARN HMM

Data for Singapore, Taiwan, New Zealand, Australia, and selected countries in Europe are downloaded from the "JHU CSSE COVID-19 Data[3]". This repository is operated and supported by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE), the ESRI Living Atlas Team, and the Johns Hopkins University's Applied Physics Lab (JHU APL), as the 2019 Novel Coronavirus Visual Dashboard [36]. Cumulative positive cases, deaths, and recovered cases of COVID-19 for most of the countries around the world are reported since January 21, 2020.

B. OBSERVATION SEQUENCES

In our experiments, we used different data sequences and derived features from them to form vector observations for learning the parameters of the HMMs. The data sequences used and the derived features are described below.

1) POSITIVE CASES

In the experiments, we leveraged normalized daily new positive cases to infer the spread of the COVID-19 pandemic, denoted by $\widetilde{P}(t)$. It is calculated by dividing the sequence of the number of daily positive cases in one region (say, Abruzzo in Italy) by the inhabitants in that region and then multiplying it by 100,000. This normalized sequence corresponds to positive cases per 100,000 population, a metric that is frequently used to indicate the spread of the disease relative to the total population size. It gives standardized information about community transmission and spread of COVID-19 [37]. Normalized sequences can be written in the format shown in (5), wherein $P(t)$ denotes the number of positive cases, and $\widetilde{P}(t)$ its normalized version:

$$\widetilde{P}(t) = \frac{100,000\,P(t)}{\text{population}}. \tag{5}$$

2) DEATHS

As with the number of positive cases, the normalized version $\widetilde{D}(t)$ of the daily deaths $D(t)$ is adopted to study how fast the disease is spreading. In order to keep the daily positive cases and daily deaths in the same dynamic range, normalization of daily deaths is calculated by dividing the sequence by the inhabitants in that region and then multiplying by 5, 000, 000, i.e.,

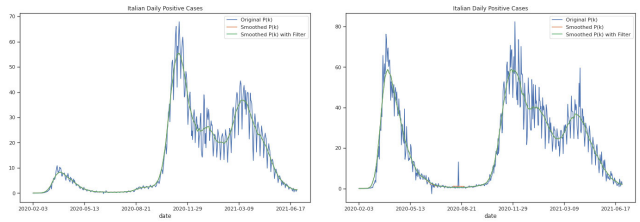
$$\widetilde{D}(t) = \frac{5,000,000\,D(t)}{\text{population}}. \tag{6}$$

C. OUTLIER REMOVAL

As Figure 2 shows, there is a visible outlier that appears in daily deaths in Italy on August 15, 2020. This jump in the curve might be due to a delay or an artifact due to bunching in reporting the data. To address this problem, we used the Hampel Filter for outlier removal and interpolation. During the experiments, window size (length of sliding window) is set to 21 to align with the length of the moving average filter we adopted on the observation sequences (Section IV-D).

[1] https://covidtracking.com/data/api
[2] https://github.com/pcm-dpc/COVID-19/
[3] https://github.com/CSSEGISandData/COVID-19/

The threshold factor c for outlier detection (in standard deviations) is set to 3, according to Pearson's rule [38].



(a) (b)

FIGURE 2. (a) Daily new positive individuals in Italy since February 3, 2020, its moving average obtained with a window of 21 days (orange line), and its moving average after outlier removal and interpolation (green line). (b) Daily new deaths in Italy since February 3, 2020, its moving average obtained with a window of 21 days (orange line), and its moving average after outlier removal and interpolation (green line).

D. SMOOTHING

To address the presence of potential problems in the data sets, such as missing values, delays in reporting the data, or errors in recording, we smooth the sequences by a moving average filter with uniform weights. After several experiments, we selected its length to be 21 days. Henceforth, such filter is denoted by MA(21). Figure 2 shows the daily new positive individuals and daily new deaths in Italy, with orange lines being the corresponding smoothed sequences without outlier removal, and green lines being the smoothed sequences after outlier removal and interpolation. In the experiments, after the outlier detection, both of the observation sequences (daily positives and daily deaths) are smoothed by MA(21) before performing normalization.

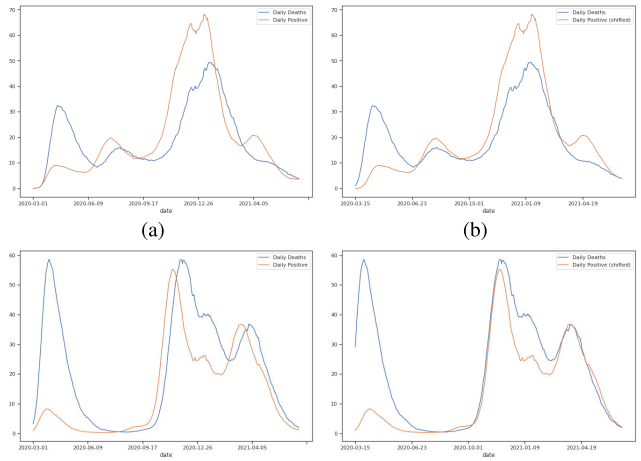
V. COMPUTATIONAL DETAIL, RESULTS AND INSIGHTS

In this section, we first describe the details of our computational experiments, including preprocessing and learning of HMM model parameters. The model is learned using data from a number of nations, including the USA, a number of European countries, as well as nations with strict COVID-19 policies. Data from these countries constituted our training data. We used the trained model in our analysis on the spread of COVID-19 in the geographic states in the USA and various regions in Italy. Data from the individual states in the USA and Italian regions formed our test set.

A. INVESTIGATION OVERVIEW

The observation sequences of daily new positive cases and daily deaths are extracted from the national and state-level datasets. Data preprocessing involved the use of a Hampel filter with a window size of 21 and a threshold factor c of 3 for outlier detection and interpolation, see Eq. (4), and then applying a MA(21) filter (defined in Section IV-D) to smooth the resulting sequences.

The left two plots in Figure 3 give examples of the raw observation sequences for the two features, viz., new positives and deaths, in the USA and Italy, respectively. We note that the two curves exhibit the same shape, but are time-shifted from each other. For both of the plots, as the contagion grows, the first rising curve is that of the new positives, then followed by daily deaths. The same order is observed in phases of decreasing contagion. This suggests that the two features are good indicators of the contagion phases. Daily positives are useful as early alerts of subsequent deaths. On the other hand, the number of deaths seems to represent a more robust index with respect to the new positives, because it is less affected by the number of tests performed and is less susceptible to inaccuracies in data reporting.



(a) (b) (c) (d)

FIGURE 3. (a) Scaled daily deaths and daily new positive cases in the USA. (b) Scaled daily deaths and 14 days right-shifted daily new positive cases in the USA. (c) Scaled daily deaths and daily new positive cases in Italy. (d) Scaled daily deaths and 14 right-shifted daily new positive cases in Italy. Note that the normalized deaths are higher during the first wave than the normalized positive because of the limited number of swab tests.

Because we consider vector observation sequences to learn HMM models, time shift is appropriate when concatenating sequences of daily positives and daily deaths together. The length of the time shift is chosen by calculating the cross-correlation [39] between the two sequences $\widetilde{P}(t)$ and $\widetilde{D}(t)$, and then selecting the index at which the cross-correlation attains its maximum. In our experiments, a time shift of 14 days between $\widetilde{D}(t)$ and $\widetilde{P}(t)$ provided the best alignment for all the nations considered. As can be seen in the two plots on the right of Figure 3, with a 14 days shift on daily new positive sequence, daily deaths and daily new positive cases are nearly perfectly aligned.

Our process for training and testing the HMM models is as follows. First, HMM models are learned from data corresponding to the USA, several European countries and four countries with stricter COVID-19 control policies (Singapore, Taiwan, New Zealand, Australia). Because we want the hidden states to correspond to the severity levels of COVID-19, we reorder the hidden states of the model in ascending order of the mean of daily deaths. Different numbers of hidden states were tried, with the constraint that

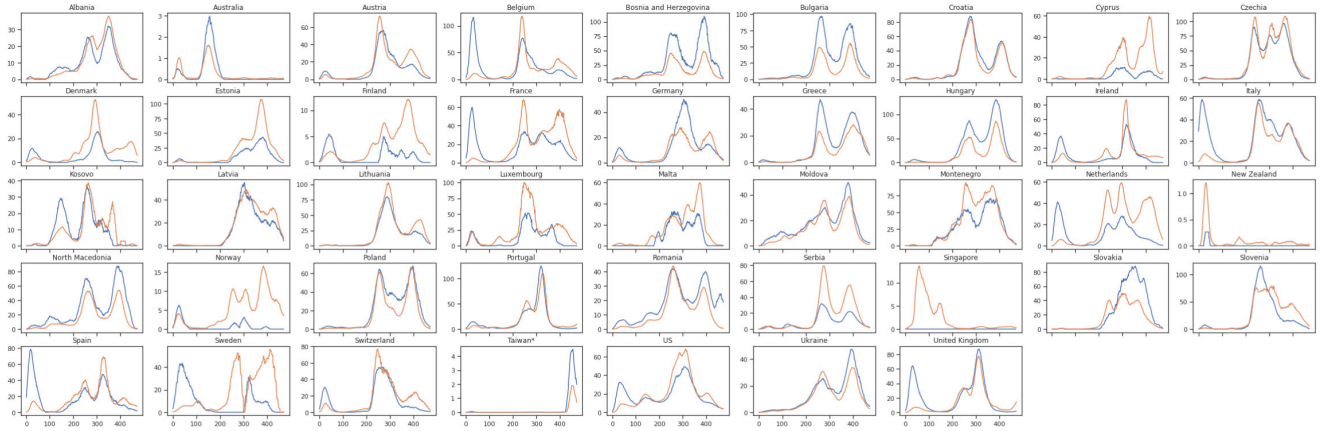


FIGURE 4. Scaled daily deaths (blue line) and right-shifted daily new positive cases (orange line) of countries used in training the model. Note the different ranges of values on the vertical axis.

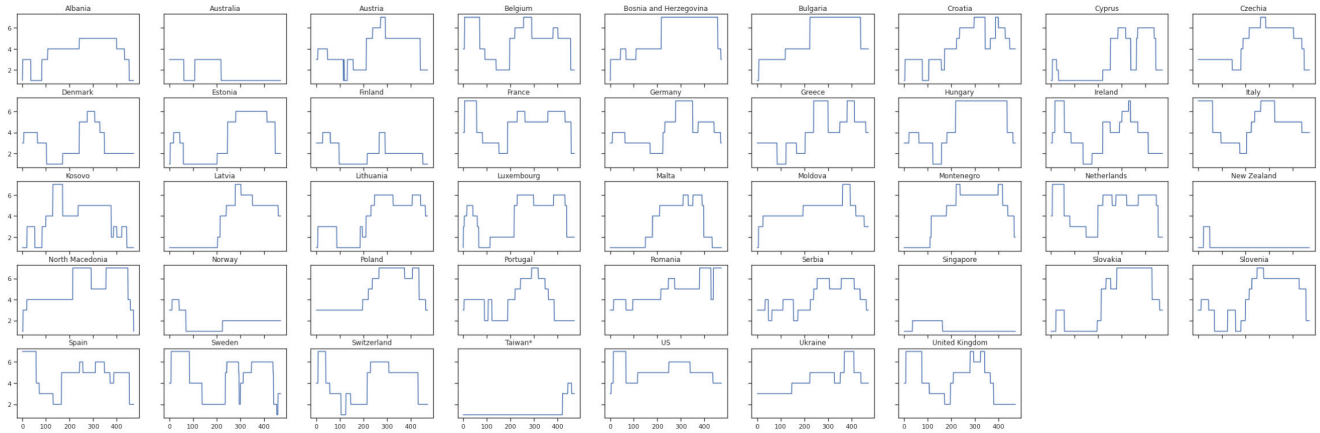


FIGURE 5. HMM Viterbi sequences of the countries used in training data.

the hidden Markov chain should be aperiodic and irreducible. This constraint implies that all the HMM states are strongly connected and that every hidden state is reachable from every other hidden state [40]. Details of the model will be presented in Section V-B.

The trained HMM model is used to infer the region-level most-likely HMM state sequences via the Viterbi algorithm. For each region or state in Italy and the USA, we estimate its HMM state sequence based on the model with the vector observation sequence from this region. This HMM state sequence is a time series, with each data point representing the severity level of this region on the corresponding day.

We are also interested in how a region's severity level changes on an aggregated time scale. In order to investigate this, the inferred HMM state sequences of each state or region are split into monthly sequences (30 days). For each such sequence, we calculate the mean, which represents the average severity level of the region over a 30-day time period. Then, aggregated state sequences can be visualized to see how the disease is propagating in the region over a longer time horizon.

B. ANALYSIS OF THE MODEL

Figure 4 shows the aligned raw observation sequences of nations that were used in building the model. The blue line in each plot shows the (scaled) daily death sequence while the orange line shows 14-days delayed daily new positive sequence, and note that the heterogeneous vertical axes amongst these plots. Figure 5 plots the Viterbi sequences for those countries. We note that the dynamically quantized state sequence in each plot in Figure 5 reproduces, with good accuracy, the corresponding trends in the plot of Figure 4. Several countries in Figure 5 show surprising results, for example, the three Baltic countries (Latvia, Lithuania, Estonia) stay at the highest severity level after transitioning from the lowest to the highest severity level.

We set the number of hidden states of our HMM model to seven, not only because we want that the hidden states to correspond to COVID-19 severity levels, but also to ensure the aperiodicity and irreducibility of the Markov chain. The left plot in Figure 6 shows the state transition diagram of the learned HMM model. Use of training data from multiple nations to learn the HMM parameters result in

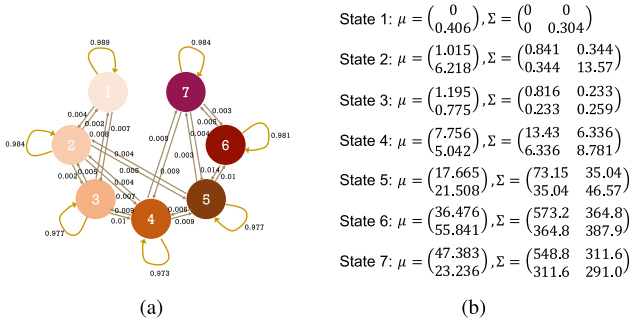


FIGURE 6. (a) State transition diagram of the overall HMM model. (b) Emission matrix of the overall HMM model.

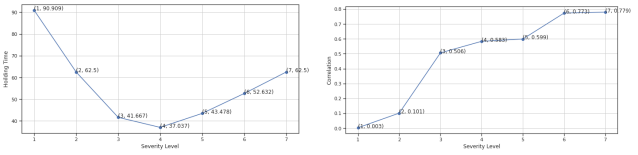


FIGURE 7. (a) Holding time in each hidden state of the overall HMM model. (b) Correlation coefficient between daily deaths and daily positive cases.

a diffused initial state probability vector. Indeed, the initial state probabilities (see Eq. 1) of the hidden chain is $\pi = [0.414, 0, 0.401, 0.137, 0, 0, 0.047]$.

The average holding time at each severity level (hidden state) can also be calculated from the model's transition matrix as $1/(1 - p_{ii})$, where p_{ii} is the diagonal element of the transition matrix. As plotted in the left figure in Figure 7, the first severity level has the largest holding time. The holding time monotonically decreases until the fourth severity level, which has the smallest holding time, and then, interestingly, it monotonically increases again up to the seventh severity level, which has the same holding time as the second severity level.

Emission matrix elements for each hidden state can be extracted as in the right plot of Figure 6. Based on the covariance matrix, we can calculate the correlation between the daily new positives and daily deaths for each severity state. It is clear from the right plot of Figure 7 that the higher the severity level, the larger the correlation. This indicates that when the severity level is high, the number of daily new positive cases on a given day in a country (or region) is a good indicator of what the number of daily deaths two weeks later will be. For countries that controlled COVID-19 well, such as Singapore, Taiwan, New Zealand and Australia, their COVID-19 contagion is typically at a lower severity level (state 1) most of the time. Consequently, the state-dependent correlation relationship between the shifted daily positive cases and deaths indicates that it is hard to predict these countries' daily deaths based on their daily positive cases because the correlation between daily deaths and daily positive cases is very low. This suggests that co-morbidities may play a substantial role in deaths at low severity level, while the number of daily positive cases can serve as reliable surrogates for the daily deaths in the future (14 days from now) at higher severity levels.

While in many instances the severity levels of the HMM model are close to the quantized versions of the corresponding raw data, there are notable exceptions that are revealing. For example, the raw data for Bulgaria, Hungary and Bosnia and Herzegovina, shown in Figure 4, reveals two waves of outbreak interleaved by a period of relative pandemic containment. The HMM analysis (Figure 5) reveals instead that for those states, the severity level remains at the highest value during the entire period.

C. ANALYSIS ON THE USA DATASETS

This section discusses the results of our analysis on the spread of COVID-19 in the 50 states and the District of Columbia in the USA. Figure 8 displays the aligned raw observation sequences of daily deaths and the 14 day right-shift of daily positive cases. The two sequences, in almost all of the states, aligned very well. This means that the 14 day time delay between daily positives and daily deaths generalizes quite well for the data from the USA. Figure 9 displays the predicted Viterbi sequences of all the 51 regions in the USA. We can see that each region has its own COVID-19 trend and the results obtained by the HMM analysis closely reflect the local evolution of the contagion. Most of the regions in the USA are either in the lower half of the severity levels or the higher half of the severity levels most of the time. Some states like California (CA), Florida (FL), Massachusetts (MA), New York (NY), Texas (TX), etc., are always in the higher half of the severity level; since these are big states, contagion severity at the national level is mainly characterized by these states. Another interesting finding is that Georgia (GA), FL, TX are pretty much always at a higher severity level, despite their death totals per capita not being as bad as NY or CA.

In order to compare the severity levels across states, we place all the Viterbi state sequences together as in Figure 10. In the figure, the x-axis is the time (day), the y-axis is the region, and each position (x, y) represents the predicted HMM state of the corresponding region based on the vector observation at that specific time. This means each line represents a region's quantized sequence that reflects the severity of COVID-19 in this region over time. We can see that for some states like CA, FL, NY, and TX, although the spread of COVID-19 was locally lessening, they were still at the most severe level compared to other states. Also, almost all the states were still at a relatively high severity level at the time that the datasets stopped updating, which is March 7, 2021.

To further explore how the COVID-19 spread in the USA, we aggregate HMM state sequences of each state by month. The path of transmission of COVID-19 over time can be inferred by the analysis of the monthly-aggregated data. If we split time into months as shown in Figure 11, we see that

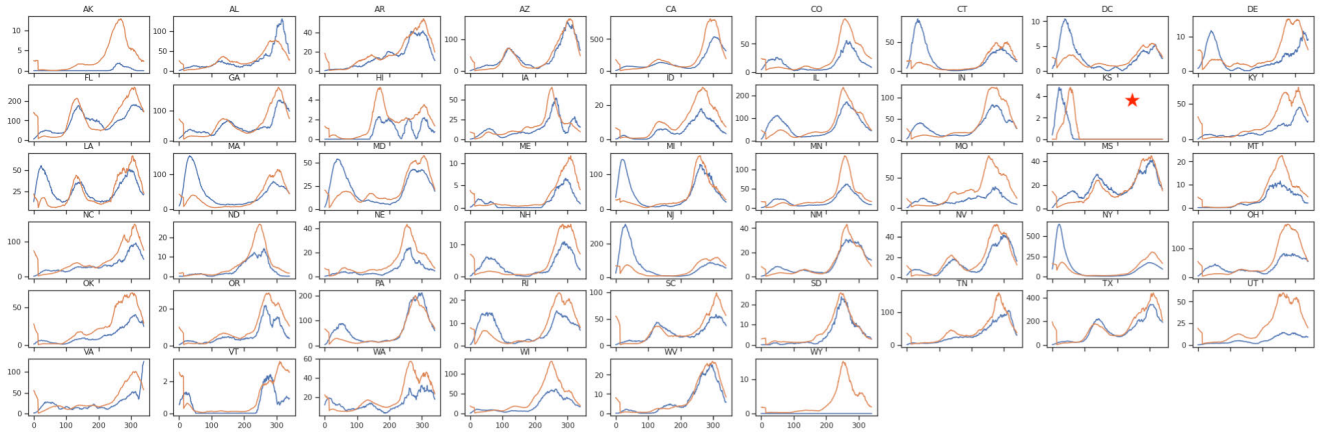


FIGURE 8. Scaled daily deaths (blue line) and right-shifted daily new positive cases (orange line) of 50 states and DC in the USA. Note the different ranges of values on the vertical axis.

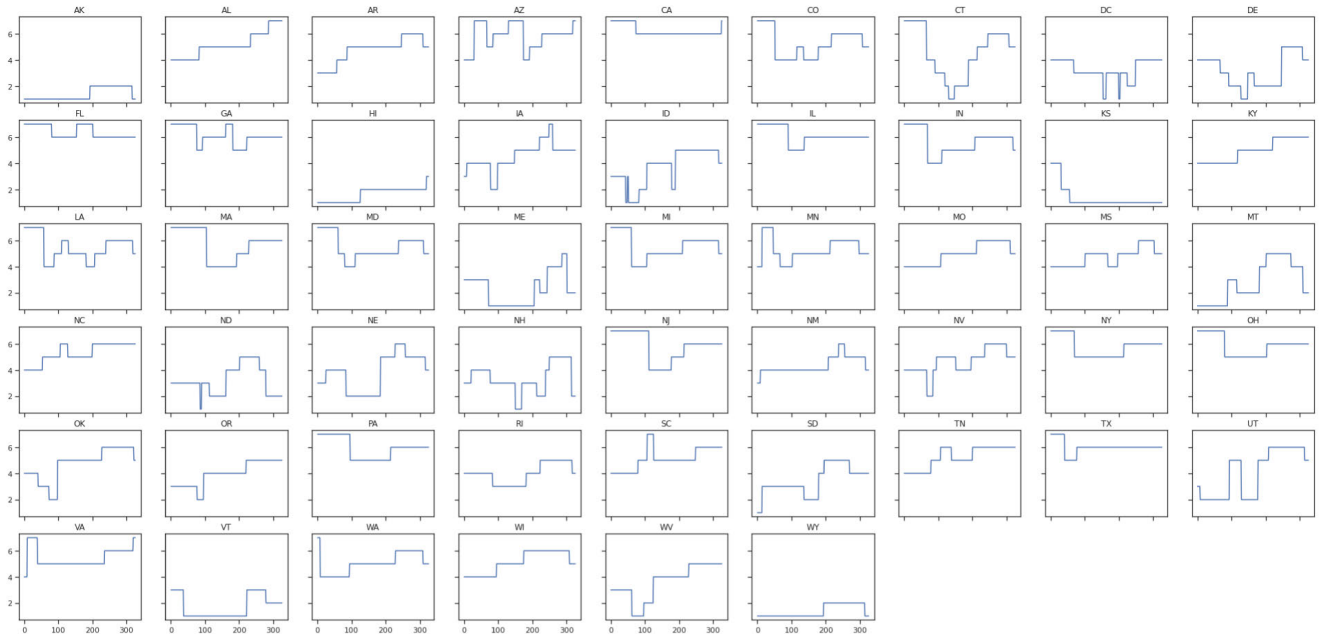


FIGURE 9. Predicted HMM Viterbi sequences of 50 states and DC in the USA.

northeastern and some southern and western states like CA, TX, Louisiana (LA), FL, and GA were firstly at a very severe level. As time progressed, the eastern part of the USA became less critical, the middle part of the USA changed to the middle severity level, while the southern states were still more severe compared with other states. Then, starting in October 2020, all the states in the USA started to get worse. Until the end of 2020, almost all states are in the fifth or sixth severity level.

Note that Kansas (KS), which is white in Figure 11, had data availability issues; both daily positive cases and deaths are not available after June 2020. See also the raw data sequences in Figure 8.

D. ANALYSIS ON THE ITALIAN DATASETS

This section shows the results of our analysis on the spread of COVID-19 in Italy. Figure 12 displays the two aligned raw observation sequences: daily deaths with time shifts of 14 days (to the right) and daily positive cases. Almost all the regions have three waves for both the daily positive cases and daily deaths. For daily positives, the first wave has the lowest peak and the second has the highest peak, while for daily deaths, some regions' first peaks are greater than the second peaks. Figures 13 and 14 display the predicted HMM state sequence for the 21 regions in Italy. We note that all the regions have a valley in the middle, corresponding to the interval between the first and second waves. Most of the

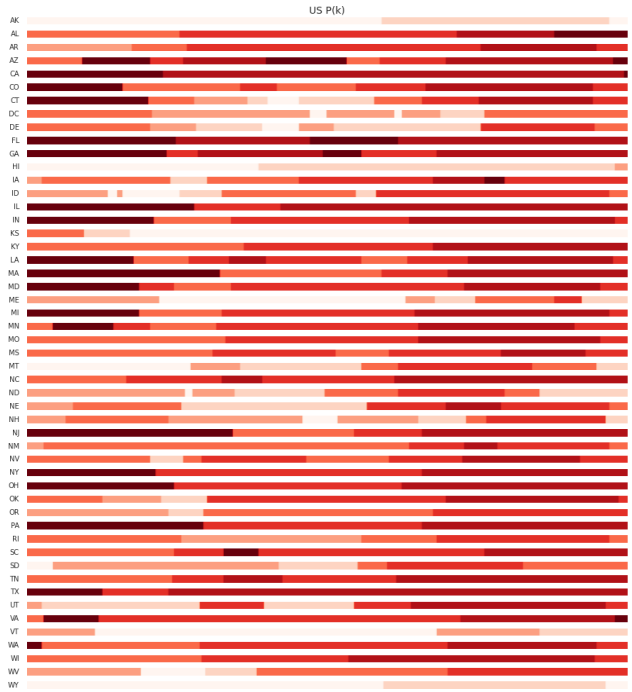


FIGURE 10. Predicted HMM state sequence for various states in the USA and DC.

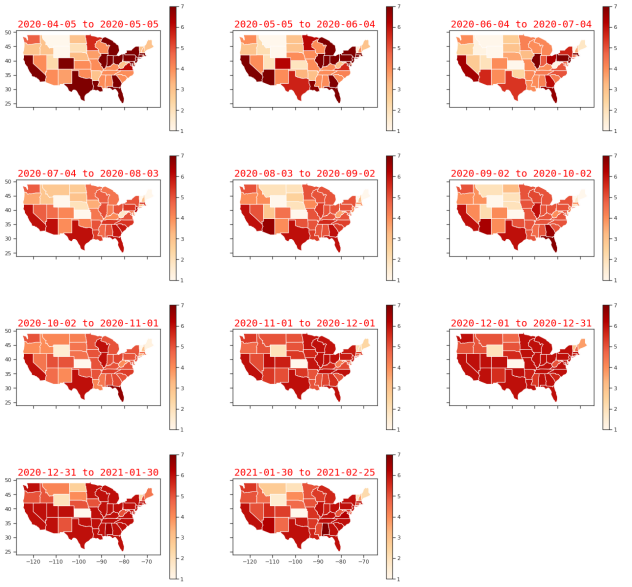


FIGURE 11. Monthly grouping results of HMM state sequences for various states in the USA.

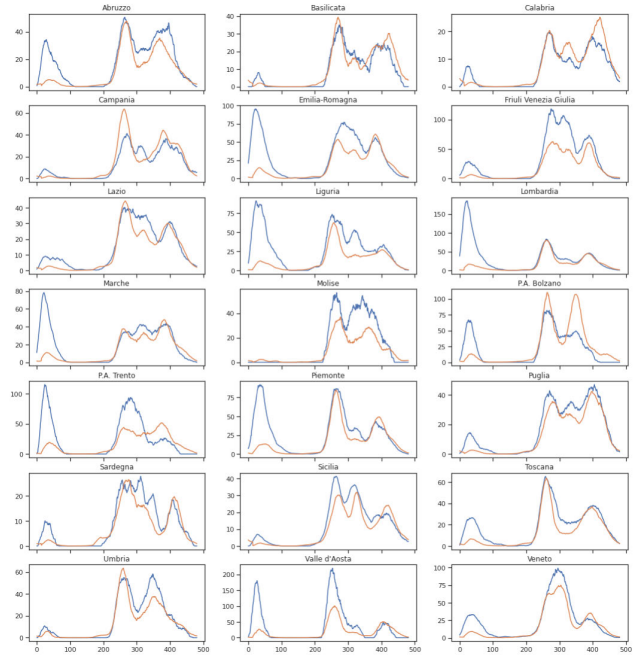


FIGURE 12. Scaled daily deaths (blue line) and right-shifted daily new positive cases (orange line) of 21 regions in Italy. Note the different ranges of values on the vertical axis.

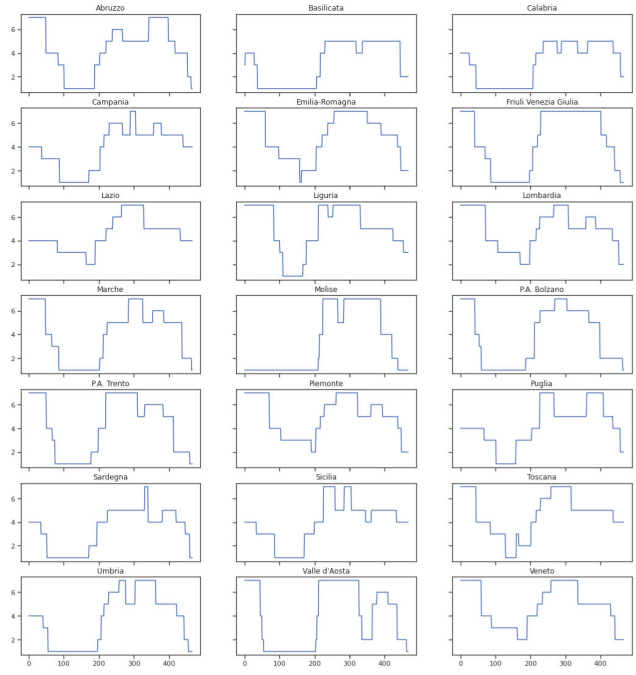


FIGURE 13. Predicted HMM Viterbi sequences of 21 regions in Italy.

regions were at the same severity level before and after the valley; this might be because the daily deaths dominate the severity of the disease at the beginning; this is an indication of the necessity of including the deaths in the observation sequence when learning the HMM models. In Italy, it appears that, when the COVID-19 situation is severe, its intensity appears to be worse in the north than the south.

Figure 15 shows the results of monthly-aggregated data, for all the regions in Italy. We note that the northern part of Italy experienced the highest severity level in the beginning. Then, all regions started to get better until the middle of July because almost all regions were in the lowest severity level; the northern regions were still a little bit worse than the southern regions. This situation changes after the middle

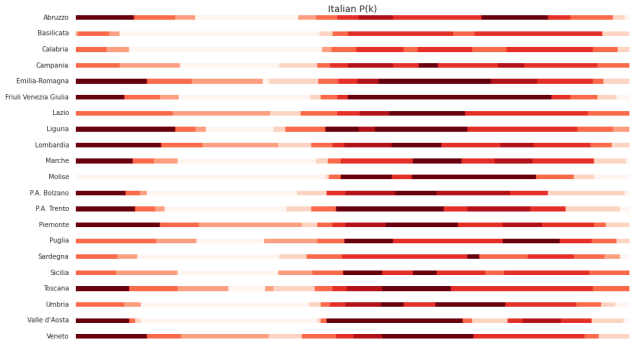


FIGURE 14. Predicted HMM state sequence of regions in Italy.

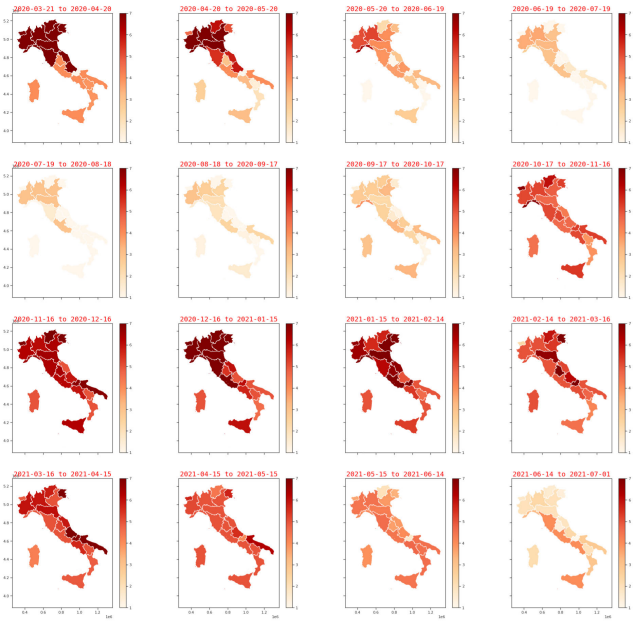
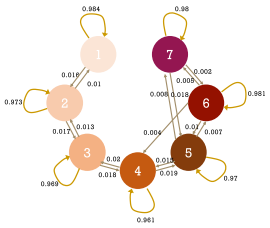


FIGURE 15. Monthly grouping results of HMM state sequences in various regions in Italy.

of October when all regions again trended alarmingly. Until the middle of December, almost all regions were at higher severity levels, especially the northern part of Italy. At the beginning of 2021, the western part of Italy started to get better to the middle severity level. By June 2021, almost all regions were at the first or the second severity level, with the western coastal regions as the exceptions at the middle level.

VI. DISCUSSION

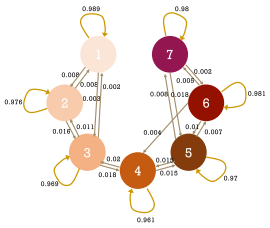
HMMs provide a useful theoretical machinery to learn a probabilistic model from data. They can quantize the raw sequences to several discrete levels so that the change of the severity of a country or a region over time is explainable by looking at the quantized state sequences. This means HMM can characterize the overall pandemic evolution in a region or nation or the entire world using scalar or vector sequences of data.



(a) 1-1

State 1: $\mu = \begin{pmatrix} 0 \\ 0.182 \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.038 \end{pmatrix}$
State 2: $\mu = \begin{pmatrix} 0.019 \\ 1.036 \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.364 \end{pmatrix}$
State 3: $\mu = \begin{pmatrix} 0.060 \\ 3.934 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.002 & -0.04 \\ -0.04 & 4.607 \end{pmatrix}$
State 4: $\mu = \begin{pmatrix} 0.171 \\ 9.665 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.010 & -0.21 \\ -0.21 & 17.29 \end{pmatrix}$
State 5: $\mu = \begin{pmatrix} 0.434 \\ 21.71 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.02 & -0.003 \\ -0.003 & 64.50 \end{pmatrix}$
State 6: $\mu = \begin{pmatrix} 0.651 \\ 58.36 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.216 & 8.399 \\ 8.399 & 436.7 \end{pmatrix}$
State 7: $\mu = \begin{pmatrix} 1.162 \\ 36.15 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.171 & 4.467 \\ 4.467 & 360.4 \end{pmatrix}$

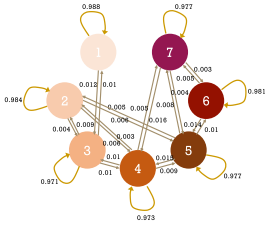
(b) 1-1



(c) 10-1

State 1: $\mu = \begin{pmatrix} 0 \\ 0.399 \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.29 \end{pmatrix}$
State 2: $\mu = \begin{pmatrix} 0.217 \\ 0.852 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.03 & 0.03 \\ 0.03 & 0.328 \end{pmatrix}$
State 3: $\mu = \begin{pmatrix} 0.590 \\ 3.876 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.210 & -0.43 \\ -0.43 & 4.446 \end{pmatrix}$
State 4: $\mu = \begin{pmatrix} 1.703 \\ 9.636 \end{pmatrix}, \Sigma = \begin{pmatrix} 1.033 & -2.07 \\ -2.07 & 17.08 \end{pmatrix}$
State 5: $\mu = \begin{pmatrix} 4.335 \\ 21.71 \end{pmatrix}, \Sigma = \begin{pmatrix} 2.244 & -0.001 \\ -0.001 & 64.38 \end{pmatrix}$
State 6: $\mu = \begin{pmatrix} 6.507 \\ 58.32 \end{pmatrix}, \Sigma = \begin{pmatrix} 21.57 & 84.01 \\ 84.01 & 436.9 \end{pmatrix}$
State 7: $\mu = \begin{pmatrix} 11.62 \\ 36.14 \end{pmatrix}, \Sigma = \begin{pmatrix} 17.13 & 44.72 \\ 44.72 & 360.7 \end{pmatrix}$

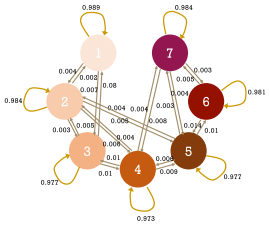
(d) 10-1



(e) 20-1

State 1: $\mu = \begin{pmatrix} 0 \\ 0.387 \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.268 \end{pmatrix}$
State 2: $\mu = \begin{pmatrix} 0.391 \\ 6.567 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.133 & 0.195 \\ 0.195 & 13.34 \end{pmatrix}$
State 3: $\mu = \begin{pmatrix} 0.465 \\ 0.839 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.199 & 0.08 \\ 0.08 & 0.314 \end{pmatrix}$
State 4: $\mu = \begin{pmatrix} 3.007 \\ 4.689 \end{pmatrix}, \Sigma = \begin{pmatrix} 2.154 & 2.445 \\ 2.445 & 7.446 \end{pmatrix}$
State 5: $\mu = \begin{pmatrix} 6.916 \\ 20.82 \end{pmatrix}, \Sigma = \begin{pmatrix} 11.50 & 14.89 \\ 14.89 & 47.59 \end{pmatrix}$
State 6: $\mu = \begin{pmatrix} 14.30 \\ 55.01 \end{pmatrix}, \Sigma = \begin{pmatrix} 91.17 & 148.1 \\ 148.1 & 395.1 \end{pmatrix}$
State 7: $\mu = \begin{pmatrix} 18.75 \\ 23.11 \end{pmatrix}, \Sigma = \begin{pmatrix} 88.70 & 124.6 \\ 124.6 & 288.7 \end{pmatrix}$

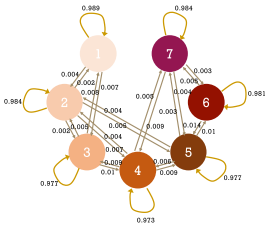
(f) 20-1



(g) 40-1

State 1: $\mu = \begin{pmatrix} 0 \\ 0.403 \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.299 \end{pmatrix}$
State 2: $\mu = \begin{pmatrix} 0.772 \\ 6.340 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.507 & 0.376 \\ 0.376 & 13.65 \end{pmatrix}$
State 3: $\mu = \begin{pmatrix} 0.96 \\ 0.799 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.515 & 0.192 \\ 0.192 & 0.282 \end{pmatrix}$
State 4: $\mu = \begin{pmatrix} 6.120 \\ 5.001 \end{pmatrix}, \Sigma = \begin{pmatrix} 8.663 & 5.025 \\ 5.025 & 8.615 \end{pmatrix}$
State 5: $\mu = \begin{pmatrix} 14.11 \\ 21.46 \end{pmatrix}, \Sigma = \begin{pmatrix} 46.77 & 28.20 \\ 28.20 & 46.74 \end{pmatrix}$
State 6: $\mu = \begin{pmatrix} 29.15 \\ 55.79 \end{pmatrix}, \Sigma = \begin{pmatrix} 366.9 & 292.2 \\ 292.2 & 388.5 \end{pmatrix}$
State 7: $\mu = \begin{pmatrix} 37.86 \\ 23.20 \end{pmatrix}, \Sigma = \begin{pmatrix} 351.6 & 249.6 \\ 249.6 & 291.1 \end{pmatrix}$

(h) 40-1



(i) 50-1

State 1: $\mu = \begin{pmatrix} 0 \\ 0.406 \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 0.304 \end{pmatrix}$
State 2: $\mu = \begin{pmatrix} 1.015 \\ 6.218 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.841 & 0.344 \\ 0.344 & 13.57 \end{pmatrix}$
State 3: $\mu = \begin{pmatrix} 1.195 \\ 0.775 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.816 & 0.233 \\ 0.233 & 0.259 \end{pmatrix}$
State 4: $\mu = \begin{pmatrix} 7.756 \\ 5.042 \end{pmatrix}, \Sigma = \begin{pmatrix} 13.43 & 6.336 \\ 6.336 & 8.781 \end{pmatrix}$
State 5: $\mu = \begin{pmatrix} 17.665 \\ 21.508 \end{pmatrix}, \Sigma = \begin{pmatrix} 73.15 & 35.04 \\ 35.04 & 46.57 \end{pmatrix}$
State 6: $\mu = \begin{pmatrix} 36.476 \\ 55.841 \end{pmatrix}, \Sigma = \begin{pmatrix} 573.2 & 364.8 \\ 364.8 & 387.9 \end{pmatrix}$
State 7: $\mu = \begin{pmatrix} 47.383 \\ 23.236 \end{pmatrix}, \Sigma = \begin{pmatrix} 548.8 & 311.6 \\ 311.6 & 291.0 \end{pmatrix}$

(j) 50-1

FIGURE 16. State transition diagram and emission matrix of models using vector observation ([death, shifted positive]) sequences with different scaling factors for daily deaths sequences. Daily positive sequences always have $\alpha = 1$. "1-1" means daily deaths has $\alpha = 1$; "10-1" means daily deaths has $\alpha = 10$, etc.

A. DISCUSSION ON THE NORMALIZATION OF OBSERVATION SEQUENCES

In the experiments, the vector observation for learning the parameters of the HMM is composed by normalized daily

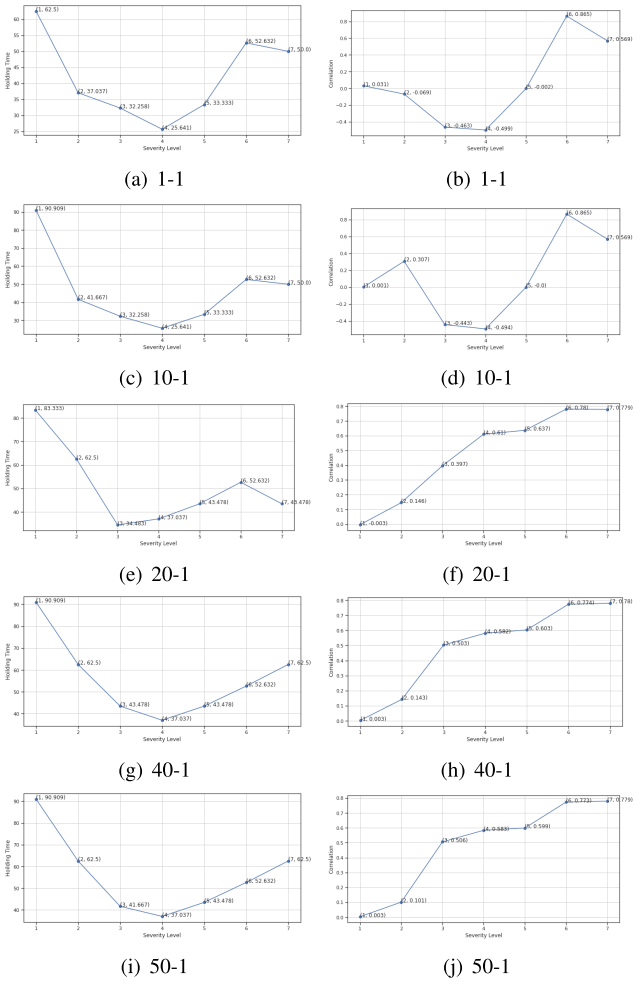


FIGURE 17. Holding time plots (a,b,d,f,h,j) and correlation plots (c,e,g,i,k) with daily death sequences normalized to different scales.

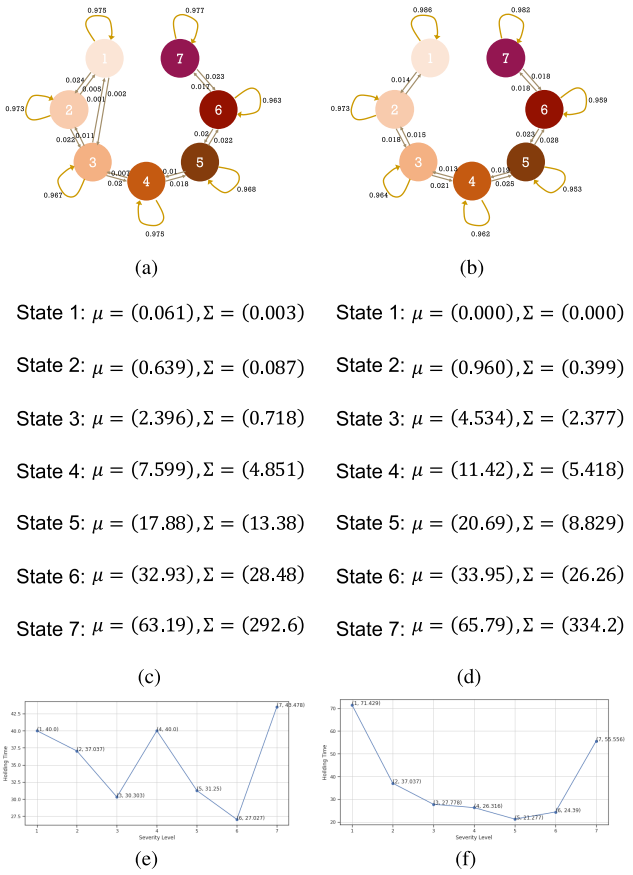


FIGURE 18. *Left*: State transition diagram, emission matrix and holding time plot when observation sequence is daily new positive cases; *Right*: State transition diagram, emission matrix and holding time plot when observation sequence is daily deaths.

positive cases and normalized daily deaths. Namely, the normalization equation of the observation sequences is:

$$\widetilde{S}(t) = \frac{\alpha \times 100,000 \times S(t)}{\text{population}}. \tag{7}$$

where $S(t)$ is the single observation sequence (daily deaths or daily new positive), α is the scaling factor in the normalization. In our experiments, the value of α for the sequence of new daily positive cases is set to 1 (normalized to a population of 100, 000), while it is set to 50 for the sequence of daily deaths (normalized to a population over 5, 000, 000). Assuming a nominal value of 2% deaths over the positive population, this choice allows us to obtain normalized sequences with the same dynamic range.

In this part, we discuss the importance of keeping the daily deaths and daily positive cases to the same dynamic range. Figure 16 shows the state transition diagram and emission matrix when building models based on vector observation sequences of deaths and shifted positive cases using different normalization scales for daily deaths. Daily positive case sequences are set to $\alpha = 1$, but daily death sequences have α set to 1, 10, 20, 40, and 50, respectively. The state transition diagram when the observation sequence comprises only daily deaths (Figure 18 b) suggests that it is a birth-death process where each state can only transit to either its next state or its previous state; in the vector observation case, the transition matrix becomes more fully populated. This suggests that daily deaths make only a small contribution to the models. As the scaling factor for daily deaths becomes larger, this feature becomes increasingly important and, at $\alpha = 50$, the two sequences are approximately on the same scale contributing equally to the model structure. With this value of α, the states in the Markov chain can transit to their next two or the previous two states (except for state 1). Moreover, Figure 17 shows the holding time plots and correlation coefficient plots. Figure 17 (a) and (c) show that there exists a negative correlation coefficient between the daily positive cases and daily deaths under some severity levels, which is contrary to our common sense. Both (a) and (c) are under the condition that the scaling factor α for daily deaths is very small, i.e., the scales of the two sequences are significantly different. As the scaling factor of daily death sequences increases, the trend – the higher the severity level, the greater the correlation between daily positive and daily

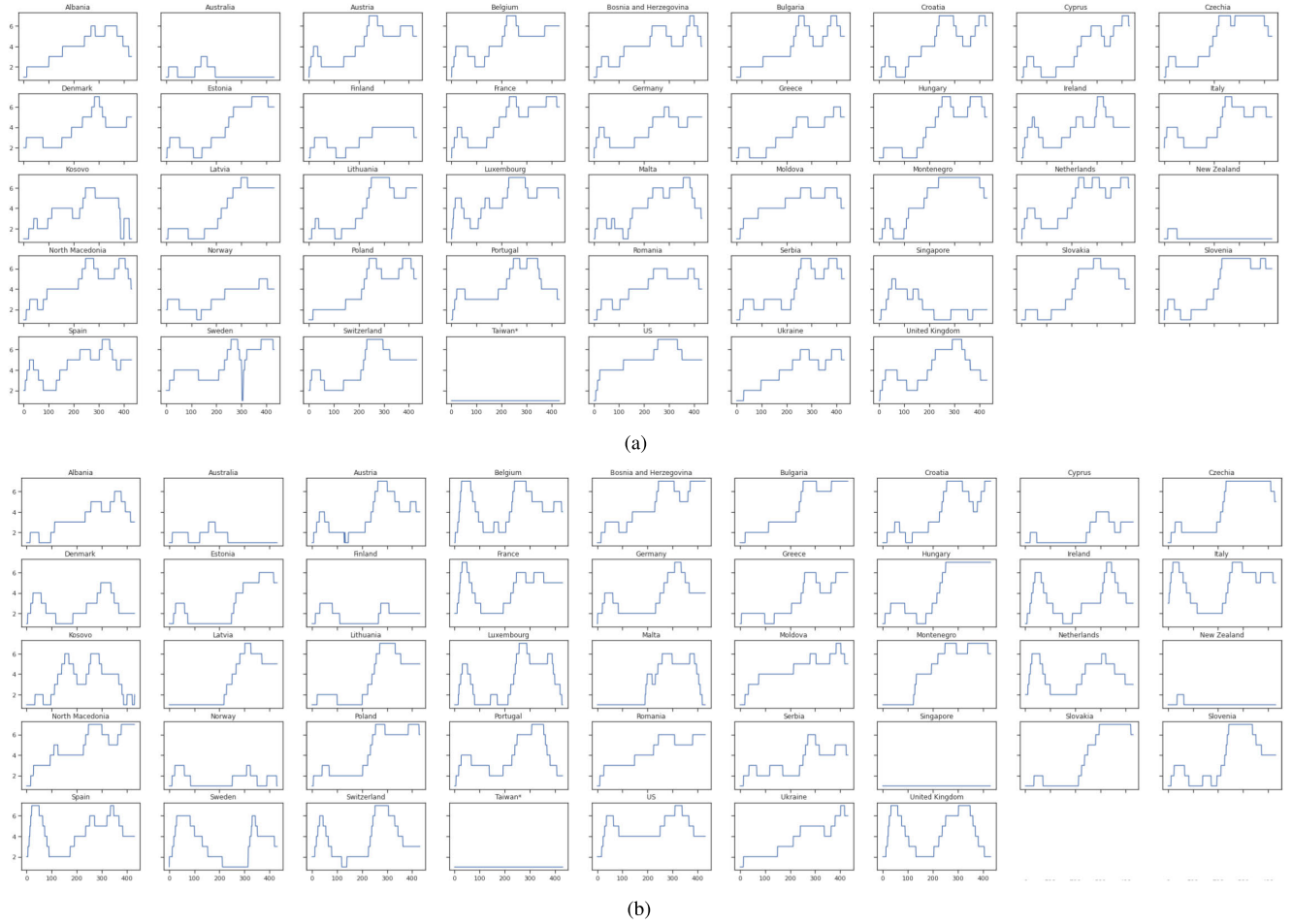


FIGURE 19. HMM Viterbi sequences of the countries used in training data when observation sequence being (a) Daily new positive cases; (b) Daily new deaths.

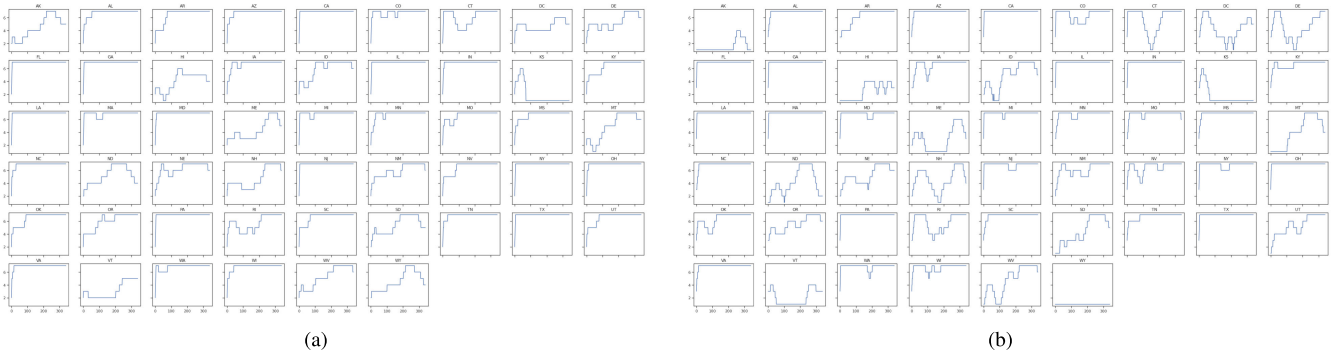


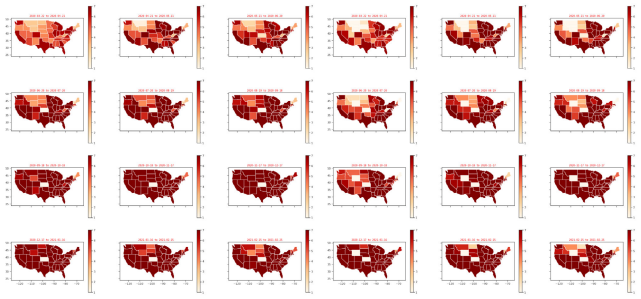
FIGURE 20. Predicted HMM Viterbi sequences of 50 states and DC in the USA under single observation sequence: (a) Daily new positive cases; (b) Daily deaths.

deaths – gradually emerges. These empirical observations demonstrate that scaling different observation sequences to the same dynamic scale is essential.

B. COMPARING SINGLE-NATION MODEL WITH MULTI-NATION MODEL

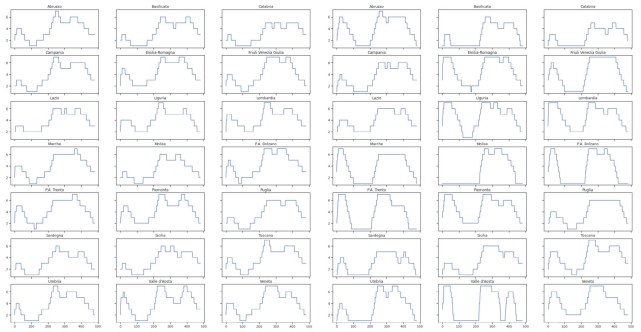
Our computational results demonstrate that a single HMM model based on data from multiple nations can be used to classify the severity level of an epidemic in a region by using normalized observation sequences that account for population size.

There are several advantages in building an across-the-nations model instead of individual models for each country. First, the multi-nation model includes diverse information in the training process. It is more generative and more broadly applicable than a single-nation model, and the predictions



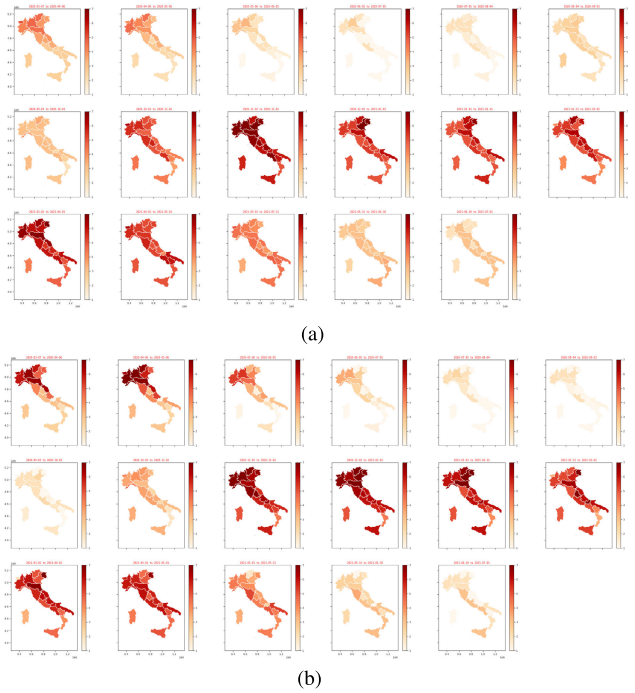
(a) (b)

FIGURE 21. Monthly grouping results of HMM state sequences for various states in the USA under single observation sequence: (a) Daily new positive cases; (b) Daily deaths.



(a) (b)

FIGURE 22. Predicted HMM Viterbi sequences of 21 regions in Italy under single observation sequence: (a) Daily new positive cases; (b) Daily deaths.



(a) (b)

FIGURE 23. Monthly grouping results of HMM state sequences in various regions in Italy under single observation sequence: (a) Daily new positive cases; (b) Daily deaths.

should be more smoother. Second, to comply with the aperiodicity and irreducibility constraints, the optimal number of hidden states differs according to different nations' data. This makes the single-nation model hard to generalize to other countries. For example, the optimal number of hidden states for the HMM model built on Italian data is three, while the optimal number of hidden states for the HMM model built on the data from the USA is nine. Consequently, using the Italian model to predict each USA state's Viterbi sequence would not be accurate or even meaningful. Third, regarding the correlation between daily positive cases and daily deaths, a clear pattern exists in the multi-nation model, which makes the multi-nation model more explainable. In summary, the multi-nation training model is able to exploit common features of the pandemic evolution that are not necessarily observed in the data of individual nations. This, we believe, is a distinct point of strength of the developed approach.

There are also disadvantages to the multi-nation model. Chief among these is that a good model of the heterogeneity of the observation streams seem to require a large number of states; and the increase in these, despite more data, seems to outpace the (asymptotic) applicability of information-based criteria – such as AIC and BIC – to estimate the best model complexity. Consequently, we chose seven to be a suitable number of states granularity to make the model physically meaningful, because the states can be directly mapped to different levels of severity.

Details of single-nation models on Italian data and USA data respectively are shown in Appendix B.

C. COMPARING SINGLE OBSERVATION MODELS WITH VECTOR OBSERVATION MODEL

In this subsection, we compare the HMM models learned from a single observation sequence, i.e., based on daily new positive cases only or based on daily deaths only, with models learned from vector observation sequences.

For the reasons explained in the previous subsection, we keep the number of hidden states as 7 for both the single and vector observation sequence models.

Figure 18 shows the state transition diagrams, emission matrices, and holding time plots for the single observation sequence case. As shown, the state transition diagrams are substantially different from the vector observation model. There is no clear pattern in the left of Figure 18, but the right of Figure 18 shows a trend similar to the holding time plot in Figure 7. Figure 19 shows the HMM Viterbi sequences of the countries used in the training data when the observation sequences are univariate. Different patterns exist for the same country when the scalar observation sequences (daily positive versus daily deaths) are somewhat different. Figure 20, 21, 22 and 23 show predicted HMM state sequences of the geographic states in the USA and the Italian regions, respectively. Similar to Figure 19, different patterns exist under different observation sequences, which suggests

the validity of integrated daily deaths and daily new positive cases as vector observations.

D. FUTURE WORK

There are a number of limitations to the proposed model that we plan to address in future studies. Firstly, HMMs have specific assumptions on the state transitions, observation distributions, and also duration distributions. It implies self-transition of a non-absorbing state with non-zero probability, so that the state duration is implicitly a geometric distribution [41]. Hidden semi-Markov model (HSMMs) [42], extensions of HMMs, might be more suitable as these allow each state to have tunable duration and sojourn time statistics. Secondly, the HMM models considered here do not include spatial correlations. Coupled HMMs/HSMMs could overcome this limitation. Finally, use of multi-modal data (e.g., mobility, air travel, cell phone data) could further enhance the utility of these models.

VII. CONCLUSION

In this paper, we used vector observation sequences, such as daily infections and daily deaths, to characterize the severity of COVID-19 using HMM models. The learned HMM models based on data from multiple countries, including the USA, several European countries and countries with strict control policies, are used to study the distributions of the spread of COVID-19 in small geographic regions, and were applied to the United States and Italy. Aggregation of the HMM state sequences over a fixed time period is also conducted to glean the COVID-19 disease's spreading patterns. Results are consistent with what is observed in the USA and Italy. Also, we compared the results from multi-nation models with results from single-nation models, e.g., models learned from only the USA data and only the Italian data, and we also compared the results from vector observation sequences with results from single observation sequences, such as daily infections or daily deaths. This analysis corroborates the approach with multiple-nation training set and vector observations. The developed approach has unique features and provides new insights:

- We obtain results hardly obtainable by other methods or simple inspection of the raw data. For instance, while in many instances the severity levels of the HMM model are close to the quantized versions of the corresponding raw data, there are notable exceptions that are revealing. Just to make an example, for Bulgaria, Hungary and Bosnia and Herzegovina, the raw data in Figure 4 shows two waves of outbreak interleaved by a period of relative pandemic containment. The HMM analysis (Figure 5) reveals instead that for those states, the severity level remains at the highest value during all the period.
- It is generally believed that the number of new positives is a good indicator of the number of daily deaths that will be observed after some time. We have shown that such prediction is reliable only when the severity level is sufficiently high. In addition, when this happens, the prediction delay is very stable: it amounts to two weeks for essentially all the considered regions.
- At a higher level, the developed approach allows us to make predictions by integrating different observation modalities (in this paper we considered the daily infections and daily deaths) beforehand, thus boosting the system performance with respect to alternative "post-prediction" data fusion strategies.

In summary, we believe that HMM models and the inferences from them can serve as a visualization tool and as a proactive decision support system to policy makers.

APPENDIX A
HMM RECURSIONS, INFERENCE AND LEARNING

Forward-Backward Recursions [30], [43] are used to calculate the model likelihood given the model parameters and observed data. This is basically an evaluation problem with the goal of estimating $p(\{\mathbf{x}\}_{t=1}^{T}|\boldsymbol{\lambda})$, where $\boldsymbol{\lambda} = (N, \boldsymbol{\pi}, A, \{b_j(\mathbf{x})\})$. By the total probability theorem, this is the sum over all possible state sequences of the joint probability density-probability mass function (pdf-pmf) of the observation sequence and the state sequence. Using the Bayes' rule and the conditional independence of a Markov process and of the emission distribution, the computation of the probability $P(z_t|\mathbf{x}_{1:T})$ can be decomposed as in (8) below.

$$\begin{aligned} P\left(z_t \mid \mathbf{x}_{1:T}\right) &= \frac{P\left(z_t, \mathbf{x}_{1:T}\right)}{p(\mathbf{x}_{1:T})} \propto P\left(z_t, \mathbf{x}_{1:T}\right) \\ &= P\left(z_t, \mathbf{x}_{1:t}, \mathbf{x}_{t+1:T}\right) \\ &= P\left(z_t, \mathbf{x}_{1:t}\right) p\left(\mathbf{x}_{t+1:T} \mid z_t, \mathbf{x}_{1:t}\right) \\ &= P\left(z_t, \mathbf{x}_{1:t}\right) p\left(\mathbf{x}_{t+1:T} \mid z_t\right). \end{aligned} \tag{8}$$

Then, the probability that the observation sequence $\mathbf{x}_{1:t}$ ends up in state z_t, denoted by $P(z_t, \mathbf{x}_{1:t})$, can be computed by forward recursion as in (9) below.

$$\begin{aligned} \alpha_t\left(z_t\right) &= P\left(\mathbf{x}_{1:t}, z_t\right) = p\left(\mathbf{x}_t \mid z_t\right) P\left(\mathbf{x}_{1:t-1}, z_t\right) \\ &= p\left(\mathbf{x}_t \mid z_t\right) \sum_{z_{t-1}} P\left(z_t \mid z_{t-1}\right) P\left(\mathbf{x}_{1:t-1}, z_{t-1}\right) \\ &= p\left(\mathbf{x}_t \mid z_t\right) \sum_{z_{t-1}} P\left(z_t \mid z_{t-1}\right) \alpha_{t-1}\left(z_{t-1}\right). \end{aligned} \tag{9}$$

Here, the initial condition is $\alpha_1(z_1) = P(z_1) = \pi_{z_1}$. The conditional probabilities, $P(z_t \mid z_{t-1})$, are obtained from the transition matrix A, and the likelihood function $p(\mathbf{x}_t \mid z_t)$ comes from the emission density $b_{z_t}(\mathbf{x})$.

In the same vein, $p(\mathbf{x}_{t+1:T} \mid z_t)$ can be computed using the backward recursion as in (10) below.

$$\begin{aligned} \beta_t\left(z_t\right) &= p\left(\mathbf{x}_{t+1:T} \mid z_t\right) = \sum_{z_{t+1}} P\left(z_{t+1}, \mathbf{x}_{t+1:T} \mid z_t\right) \\ &= \sum_{z_{t+1}} p\left(\mathbf{x}_{t+2:T} \mid z_{t+1}\right) p\left(\mathbf{x}_{t+1} \mid z_{t+1}\right) P\left(z_{t+1} \mid z_t\right) \\ &= \sum_{z_{t+1}} \beta_{t+1}\left(z_{t+1}\right) p\left(\mathbf{x}_{t+1} \mid z_{t+1}\right) P\left(z_{t+1} \mid z_t\right). \end{aligned} \tag{10}$$

Here, the terminal condition is $\beta_T(\mathbf{x}_T) = 1$. $P(z_{t+1} \mid z_t)$ can be obtained from the transition matrix A, and

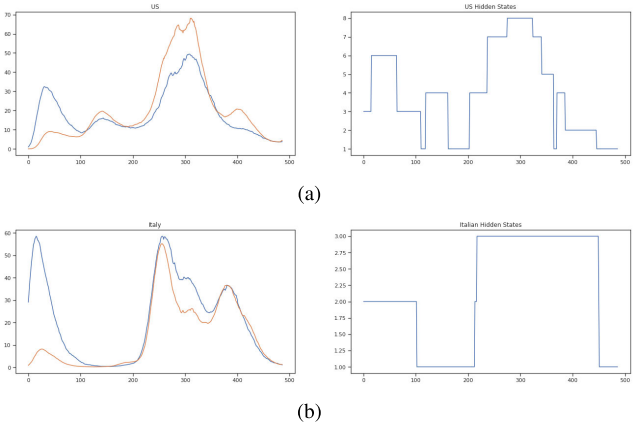


FIGURE 24. Upper left: Scaled daily deaths (blue line) and right-shifted daily new positive cases (orange line) of the USA; Upper right: HMM Viterbi sequences of the USA; Lower left: Scaled daily deaths (blue line) and right-shifted daily new positive cases (orange line) of Italy; Lower right: HMM Viterbi sequences of Italy.

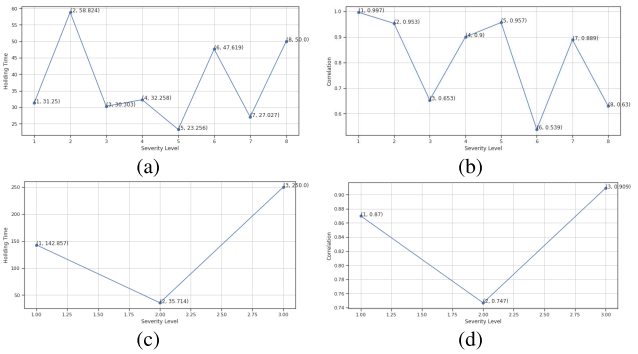


FIGURE 25. (a) Holding time plots of the USA model; (b) Correlation plots of the USA model; (c) Holding time plots of the Italian model; (d) Correlation plots of the Italian Model.

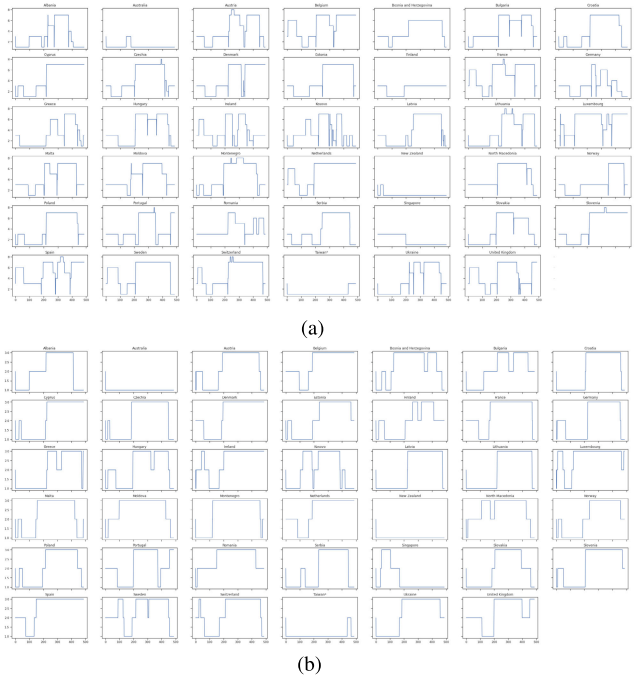


FIGURE 26. Predicted HMM Viterbi sequences of several nations under single-nation models: (a) under the USA model; (b) under the Italian model.

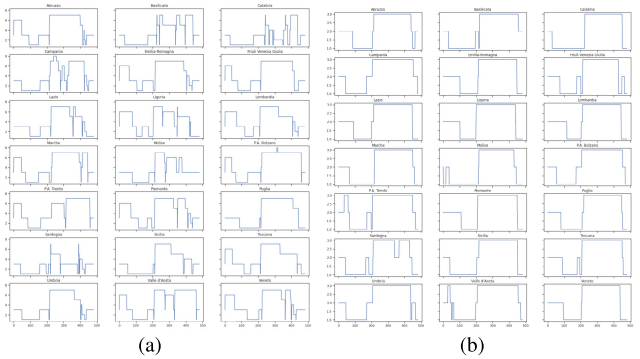


FIGURE 27. Predicted HMM Viterbi sequences of 21 regions in Italy under: (a) the USA model; (b) the Italian model.

$p(\mathbf{x}_{t+1} \mid z_{t+1})$ can be obtained from the emission density $b_{z_{t+1}}(\mathbf{x})$.

Then, the probability of each hidden state at any time t can be calculated when a sequence of observations is available using the fact that $P(z_t \mid \mathbf{x}_{1:T}) \propto \alpha_t(z_t)\,\beta_t(z_t)$, and normalizing the probabilities to sum to 1.

Baum-Welch Algorithm [43] is used to estimate the HMM parameters, given the observed data. This is a parameter estimation problem, and Expectation–Maximization (EM) algorithm is adopted to solve this learning problem [44]. If a Gaussian output probability distribution $b_j(\mathbf{x}) = N(\mathbf{x} : \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ is assumed, the parameters that need to be estimated are the transition probabilities $\{a_{ij}\}$, and two Gaussian parameters for state z_j: $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$.

Then each iteration has two steps: Given the current estimates of HMM parameters, the E-step recursively computes the forward probabilities $\alpha_t(z_j)$, backward probabilities $\beta_t(z_j)$ and the state occupancy probabilities. The M-step re-estimates the HMM parameters based on the estimated state occupancy probabilities. The M-step computes the $\{\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j\}$, $\{a_{ij}\}$ and $\{\pi_i\}$ as in (11):

$$\begin{aligned}
\hat{\boldsymbol{\mu}}_j &= \frac{\sum_{t=1}^{T} \gamma_t(z_j)\,\mathbf{x}_t}{\sum_{t=1}^{T} \gamma_t(z_j)} \\
\hat{\boldsymbol{\Sigma}}_j &= \frac{\sum_{t=1}^{T} \gamma_t(z_j)\,(\mathbf{x}_t - \hat{\boldsymbol{\mu}}_j)(\mathbf{x}_t - \hat{\boldsymbol{\mu}}_j)^T}{\sum_{t=1}^{T} \gamma_t(z_j)} \\
\hat{a}_{ij} &= \frac{\sum_{t=1}^{T-1} \xi_t(z_i, z_j)}{\sum_{k=1}^{N} \sum_{t=1}^{T-1} \xi_t(z_i, z_k)} \\
\hat{\pi}_i &= \frac{\gamma_1(z_i)}{\sum_{k=1}^{N} \gamma_1(z_k)}
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
\xi_t(z_i, z_j) &= \frac{\alpha_t(z_i)\,a_{ij} b_j(\mathbf{x}_{t+1})\,\beta_{t+1}(z_j)}{\alpha_T(z_T)}, \\
\gamma_t(z_j) &= \frac{1}{\alpha_T(z_T)} \alpha_t(z_j)\beta_t(z_j).
\end{aligned}$$

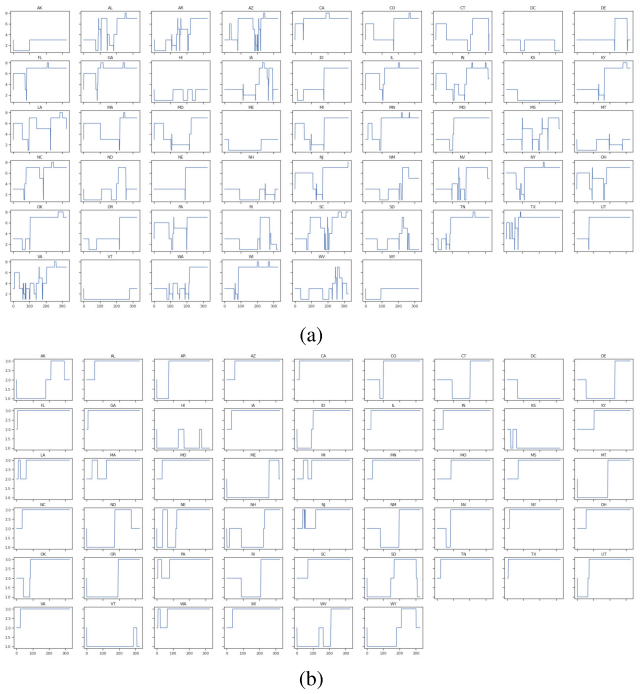


FIGURE 28. Predicted HMM Viterbi sequences of 50 states and DC in the USA under: (a) the USA model; (b) the Italian model.

Viterbi Algorithm [45] is used to estimate the optimal sequence of hidden states, given the model parameters and the observed data, i.e., inferring the evolution of hidden state sequence. Letting $z_0 = \emptyset$, the maximum a posterior estimate can be computed using (12) below:

$$\begin{aligned}\mathbf{z}^T &= \arg\max_{z_{1:T}} P\left(z_{1:T} \mid \mathbf{x}_{1:T}\right) = \arg\max_{z_{1:T}} P\left(\mathbf{x}_{1:T}, z_{1:T}\right) \\ &= \arg\max_{z_{1:T}} \prod_{t=1}^{T} \left[p\left(\mathbf{x}_t \mid z_t\right) P\left(z_t \mid z_{t-1}\right)\right] \\ &= \arg\max_{z_{1:T}} \left[\sum_{t=1}^{T} \left\{\ln p\left(\mathbf{x}_t \mid z_t\right) + \ln P\left(z_t \mid z_{t-1}\right)\right\}\right]. \quad (12)\end{aligned}$$

Forward dynamic programming is leveraged to recursively find the probability of the most likely hidden state sequence, as shown in (13):

$$\begin{aligned}\omega\left(z_t\right) &= \max_{z_{1:t-1}} P\left(\mathbf{x}_{1:t}, z_{1:t}\right) \\ &= \max_{z_{1:t-1}} \left[\sum_{n=1}^{t} \left\{\ln p\left(\mathbf{x}_n \mid z_n\right) + \ln P\left(z_n \mid z_{n-1}\right)\right\}\right] \\ &= \ln p\left(\mathbf{x}_t \mid z_t\right) \\ &\quad + \max_{z_{0:t-1}} \left[\sum_{n=1}^{t-1} \ln p\left(\mathbf{x}_n \mid z_n\right) + \sum_{n=1}^{t} \ln P\left(z_n \mid z_{n-1}\right)\right] \\ &= \ln p\left(\mathbf{x}_t \mid z_t\right) + \max_{z_{t-1}} \left[\ln P\left(z_t \mid z_{t-1}\right) + \omega\left(z_{t-1}\right)\right]. \end{aligned} \quad (13)$$

The initial condition is $\omega\left(z_1\right) = \ln P\left(z_1\right) = \ln \pi_{z_1}$. The most likely sequence can be obtained by backtracking as in (14), where $t = T-1, \ldots, 1$:

$$\begin{aligned} z_T^* &= \arg\max_{z_T} \omega\left(z_T\right) \\ z_t^* &= \arg\max_{z_t} \left[\ln P\left(z_{t+1}^* \mid z_t\right) + \omega\left(z_t\right)\right]. \end{aligned} \quad (14)$$

Given an observed data sequence for a country (say USA or Italy), we learn the HMM model parameters using the Baum-Welch algorithm and apply it to individual state (region) data to infer their severity states over time via the Viterbi algorithm. The USA's model, for example, can be used to assess the severity levels of the disease in other countries.

APPENDIX B
SINGLE-NATION MODELS

Detail of single-nation models on Italian data and USA's data are shown in this appendix. Figure 24 shows the scaled daily deaths and right-shifted daily new positive cases of the USA and Italy, respectively. As shown, the USA model has nine hidden states, and the Italian model has three hidden states. Figure 25 shows the correlation between daily deaths and daily new positive cases under the USA and the Italian models, respectively. As shown, there is no clear pattern in the correlation plots. Figure 26 shows the predicted HMM Viterbi sequences of nations, which were used as training data under multi-nation models, under the two single-nation models. Figure 27 and 28 show predicted HMM Viterbi sequences of Italian regions and the USA states under the single-nation models, respectively. As discussed in the main part of the paper, Viterbi state predictions of the single nation model are noisy, while those from multi-national model are smoother.

REFERENCES

[1] G. A. Churchill, "Stochastic models for heterogeneous DNA sequences," *Bull. Math. Biol.*, vol. 51, no. 1, pp. 79–94, 1989.
[2] T. Koski, *Hidden Markov Models for Bioinformatics*, vol. 2. Norwell, MA, USA: Kluwer, 2001.
[3] R. C. Elston and J. Stewart, "A general model for the genetic analysis of pedigree data," *Hum. Heredity*, vol. 21, no. 6, pp. 523–542, 1971.
[4] E. S. Lander and P. Green, "Construction of multilocus genetic linkage maps in humans," *Proc. Nat. Acad. Sci. USA*, vol. 84, no. 8, pp. 2363–2367, Apr. 1987.
[5] A. Krogh, M. Brown, I. S. Mian, K. Sjölander, and D. Haussler, "Hidden Markov models in computational biology: Applications to protein modeling," *J. Mol. Biol.*, vol. 235, no. 5, pp. 1501–1531, 1994.
[6] R. Hughey and A. Krogh, "Hidden Markov models for sequence analysis: Extension and analysis of the basic method," *Bioinformatics*, vol. 12, no. 2, pp. 95–107, 1996.
[7] B.-J. Yoon, "Hidden Markov models and their applications in biological sequence analysis," *Current Genomics*, vol. 10, no. 6, pp. 402–415, 2009.
[8] S. R. Eddy, "Hidden Markov models," *Current Opinion Struct. Biol.*, vol. 6, no. 6, pp. 361–365, 1996.
[9] L. Davis, *Basic Methods in Molecular Biology*. Amsterdam, The Netherlands: Elsevier, 2012.
[10] Y. L. Strat and F. Carrat, "Monitoring epidemiologic surveillance data using hidden Markov models," *Statist. Med.*, vol. 18, no. 24, pp. 3463–3478, Dec. 1999.
[11] R. E. Watkins, S. Eagleson, B. Veenendaal, G. Wright, and A. J. Plant, "Disease surveillance using a hidden Markov model," *BMC Med. Inform. Decis. Making*, vol. 9, no. 1, pp. 1–12, Dec. 2009.
[12] R. J. Cook and J. F. Lawless, "Statistical issues in modeling chronic disease in cohort studies," *Statist. Biosci.*, vol. 6, no. 1, pp. 127–161, May 2014.
[13] B. Cooper, "The analysis of hospital infection data using hidden Markov models," *Biostatistics*, vol. 5, no. 2, pp. 223–237, Apr. 2004.
[14] P. J. Green and S. Richardson, "Hidden Markov models and disease mapping," *J. Amer. Stat. Assoc.*, vol. 97, no. 460, pp. 1055–1070, 2002.

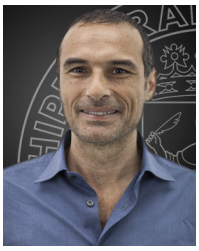
[15] K. Biswas, A. Khaleque, and P. Sen, "Covid-19 spread: Reproduction of data and prediction using a SIR model on Euclidean network," 2020, *arXiv:2003.07063*. [Online]. Available: http://arxiv.org/abs/2003.07063
[16] A. L. Ziff and R. M. Ziff, "Fractal kinetics of COVID-19 pandemic," *Int. J. Educ. Excellence*, vol. 6, no. 1, pp. 43–69, 2020. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.02.16.20023820v2, doi: 10.18562/IJEE.053.
[17] C. Eksin, K. Paarporn, and J. S. Weitz, "Systematic biases in disease forecasting—The role of behavior change," *Epidemics*, vol. 27, pp. 96–105, Jun. 2019.
[18] S. Das, "Prediction of COVID-19 disease progression in India: Under the effect of national lockdown," 2020, *arXiv:2004.03147*. [Online]. Available: http://arxiv.org/abs/2004.03147
[19] D. Gaglione, P. Braca, L. M. Millefiori, G. Soldi, N. Forti, S. Marano, P. K. Willett, and K. R. Pattipati, "Adaptive Bayesian learning and forecasting of epidemic evolution—Data analysis of the COVID-19 outbreak," *IEEE Access*, vol. 8, pp. 175244–175264, 2020.
[20] B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, and J. Wu, "Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions," *J. Clin. Med.*, vol. 9, no. 2, p. 462, Feb. 2020.
[21] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study," *Lancet*, vol. 395, no. 10225, pp. 689–697, Mar. 2020.
[22] Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, and Z. Mai, "Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions," *J. Thoracic Disease*, vol. 12, no. 3, p. 165, 2020.
[23] P. Braca, D. Gaglione, S. Marano, L. M. Millefiori, P. Willett, and K. Pattipati, "Decision support for the quickest detection of critical COVID-19 phases," *Sci. Rep.*, vol. 11, no. 1, p. 8558, Dec. 2021.
[24] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, and A. Vespignanid, "The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak," *Science*, vol. 368, no. 6489, pp. 395–400, Apr. 2019.
[25] Y. Zhang *et al.*, "Prediction of the COVID-19 outbreak in China based on a new stochastic dynamic model," *Sci. Rep.*, vol. 10, p. 21522, 2020. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.03.10.20033803v1, doi: 10.1038/s41598-020-76630-0.
[26] P. Braca, D. Gaglione, S. Marano, L. M. Millefiori, P. Willett, and K. R. Pattipati, "Quickest detection of COVID-19 pandemic onset," *IEEE Signal Process. Lett.*, vol. 28, pp. 683–687, 2021.
[27] G. Soldi, N. Forti, D. Gaglione, P. Braca, L. M. Millefiori, S. Marano, P. Willett, and K. Pattipati, "Quickest detection and forecast of pandemic outbreaks: Analysis of COVID-19 waves," *IEEE Commun. Mag.*, Sep. 2021. [Online]. Available: https://arxiv.org/pdf/2101.04620
[28] T. T. Le, F. Chatelain, and C. Berenguer, "Hidden Markov models for diagnostics and prognostics of systems under multiple deterioration modes," in *Proc. 24th Eur. Saf. Rel. Conf. (ESREL)*. Boca Raton, FL, USA: CRC Press, 2014, pp. 1197–1204.
[29] M. Bjerkeseth, "Using hidden Markov models for fault diagnostics and prognosis in condition based maintenance systems," M.S. thesis, Dept. Inf. Commun. Technol., Univ. Agder, Kristiansand, Norway, 2010.
[30] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[31] H. Liu, S. Shah, and W. Jiang, "On-line outlier detection and data cleaning," *Comput. Chem. Eng.*, vol. 28, no. 9, pp. 1635–1647, Aug. 2004.
[32] E. W. Grafarend, *Linear Nonlinear Models: Fixed Effects, Random Effects, and Mixed Models*. Berlin, Germany: Walter de Gruyter, 2006.
[33] L. Davies and U. Gather, "The identification of multiple outliers," *J. Amer. Stat. Assoc.*, vol. 88, no. 423, pp. 782–792, 1993.
[34] R. K. Pearson, "Outliers in process modeling and identification," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 1, pp. 55–63, Jan. 2002.
[35] F. R. Hampel, "The influence curve and its role in robust estimation," *J. Amer. Statist. Assoc.*, vol. 69, no. 346, pp. 383–393, 1974.
[36] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, 2020.
[37] N. Medicine. *How to Calculate COVID-19 Stats for Your Area*. Accessed: Jan. 12, 2021. [Online]. Available: https://www.nebraskamed.com/COVID/how-to-calculate-covid-19-stats-for-your-area#positive_cases_per_capita
[38] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
[39] R. N. Bracewell, *The Fourier Transform and its Applications*, vol. 31999. New York, NY, USA: McGraw-Hill, 1986.
[40] P. A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation*. Hoboken, NJ, USA: Wiley, 2017.
[41] Z. Ghahramani, "An introduction to hidden Markov models and Bayesian networks," in *Hidden Markov Models: Applications in Computer Vision*. Singapore: World Scientific, 2001, pp. 9–41.
[42] J. D. Ferguson, "Variable duration models for speech," in *Proc. Symp. Appl. Hidden Markov Models Text Speech*. Princeton, NJ, USA: Institute for Defense Analyses, Oct. 1980, pp. 143–179.
[43] L. E. Baum, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," *Inequalities*, vol. 3, no. 1, pp. 1–8, 1972.
[44] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. B, Methodol.*, vol. 39, no. 1, pp. 1–22, 1977.
[45] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

SHANGLIN ZHOU (Member, IEEE) received the B.Sc. degree in statistics from Minzu University of China, Beijing, China, in 2015, and the M.Sc. degree in statistics from the University of Connecticut, Storrs, in 2017, where she is currently pursuing the Ph.D. degree in computer science.

Since 2019, she has been a Research Assistant with the University of Connecticut. Her main research interests include machine learning, deep learning model compression, and the application of computer vision.

PAOLO BRACA (Senior Member, IEEE) received the Laurea degree *(summa cum laude)* in electronic engineering and the Ph.D. degree (Hons.) in information engineering from the University of Salerno, Italy, in 2006 and 2010, respectively.

In 2009, he was a Visiting Scholar with the Department of Electronics and Communication Engineering (ECE), University of Connecticut. From 2010 to 2011, he was a Postdoctoral Associate with the University of Salerno. In 2011, he joined NATO STO CMRE, where he is currently a Senior Scientist and the Project Manager. Furthermore, he leaded a number of research projects funded by the EU Horizon 2020 Program, the U.S. Office of Naval Research (ONR), and other institutions. He has coauthored more than 100 publications in international scientific journals and conference proceedings. He conducts research in the general area of statistical signal processing with emphasis on detection and estimation theory, wireless sensor networks, multi-agent algorithms, target tracking and data fusion, adaptation and learning over graphs, radar (sonar) signal processing, and machine learning. He is in the technical committee of the major international conferences in the field of signal processing and data fusion. He was awarded with the National Scientific Qualification to function as an Associate Professor and a Full Professor at Italian Universities, in 2017 and 2018, respectively. He was a recipient of the Best Student Paper Award (first Runner-Up) at FUSION conference in 2009, and the NATO STO Scientific Achievement Award (SAA) 2017 for its contribution to the "Development and Demonstration of Networked Autonomous ASW." He has coauthored the paper received the Best Paper Award (first Runner-Up) at the SSPD conference, in 2019. He was also a recipient of NATO STO SAA, in 2020, as the Team Leader for the "Advances in Artificial Intelligence and Information Fusion for Maritime Situational Awareness." In 2017, he was a Lead Guest Editor of the Special Issue "Sonar Multi-Sensor Applications and Techniques" in *IET-RSN*. He is currently serving as AE for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2016–present), IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS (2015–present), *IET Radar, Sonar & Navigation* (2018–present), as Area Editor for *ISIF Journal of Advances in Information Fusion* (2019–present). He was an AE for the *IEEE Signal Processing Magazine e-Newsletter* (2014–2016), *EURASIP Journal of Advances in Signal Processing* (2015–2019), *ISIF Journal of Advances in Information Fusion* (2014–2019).

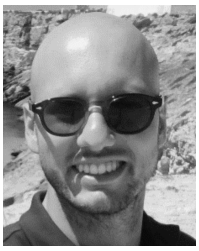


STEFANO MARANO (Senior Member, IEEE) received the Laurea degree *(summa cum laude)* in electronic engineering and the Ph.D. degree in electronic engineering and computer science from the University of Naples, Italy, in 1993 and 1997, respectively.

He has held visiting positions with the Department of Physics, College of Cardiff, University of Wales, and the Department of Electrical and Computer Engineering, University of California at San Diego, in 1996 and 2013, respectively. Since 1999, he has been with the University of Salerno, Italy, where he is currently a Professor with DIEM. His research interests include statistical signal processing with emphasis on distributed inference, sensor networks, and information theory. He was in the Organizing Committee of the Ninth International Conference on Information Fusion (FUSION 2006), and the 2008 IEEE Radar Conference (RADARCON 2008). He was awarded the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION 1999 Best Paper Award for his work on stochastic modeling of electromagnetic propagation in urban environments. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, from 2010 to 2014, and an Associate Editor and a Technical Editor for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, from 2009 to 2016.

PETER WILLETT has been a Faculty Member with the Department of Electrical and Computer Engineering, University of Connecticut, since 1986. Since 1998, he has also been a Professor. His primary research interests include statistical signal processing, detection, machine learning, communications, data fusion, and tracking. He is the Chief Editor of *IEEE Aerospace and Electronic Systems Magazine*, for the period 2018–2021.

LEONARDO M. MILLEFIORI (Member, IEEE) received the B.Sc. degree in aerospace information engineering and the M.Sc. degree *(summa cum laude)* in communication engineering with a focus on radar systems and remote sensing from Sapienza University of Rome, Italy, in 2010 and 2013, respectively. He was a Visiting Researcher with NATO Science and Technology Organization Center for Maritime Research and Experimentation (CMRE), La Spezia, where he joined the Research Department as a Research Scientist, in 2014. His research interests include signal processing, statistical and machine learning, target motion modeling, and target tracking and data fusion. He was a recipient of NATO STO Scientific Achievement Award, in 2020, as one of the Team Leaders of the Data Knowledge and Operational Effectiveness (DKOE) Research Group, CMRE for "Advances in Artificial Intelligence and Information Fusion for Maritime Situational Awareness."



DOMENICO GAGLIONE (Member, IEEE) received the B.Sc. and M.Sc. degrees *(summa cum laude)* in telecommunications engineering from the Università degli Studi di Napoli "Federico II," Naples, Italy, in 2011 and 2013, respectively, and the Ph.D. degree from the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K., in 2017.

Since 2016, he has been a Research Assistant with the University of Strathclyde. He joined NATO STO Centre for Maritime Research and Experimentation (CMRE), La Spezia, Italy, in 2017, as a Junior Scientist. His research interests include statistical signal processing with emphasis on state estimation, data fusion, and multi-sensor multi-target tracking. He was a recipient of the Best Student Paper Award at the 2015 IEEE International Radar Conference (RadarCon), Arlington, VA, USA, and the NATO STO Scientific Achievement Award in 2020, as a member of the Data Knowledge and Operational Effectiveness (DKOE) Research Group, CMRE for "Advances in Artificial Intelligence and Information Fusion for Maritime Situational Awareness."

KRISHNA R. PATTIPATI (Life Fellow, IEEE) received the B.Tech. degree (Hons.) in electrical engineering from the Indian Institute of Technology Kharagpur, Kharagpur, in 1975, and the M.S. and Ph.D. degrees in systems engineering from UCONN, Storrs, in 1977 and 1980, respectively. He was with ALPHATECH Inc., Burlington, MA, USA, from 1980 to 1986.

He has been with the Department of Electrical and Computer Engineering, UCONN, since 1986, where is currently the Board of Trustees Distinguished Professor and the UTC Chair Professor of systems engineering. He is also the Cofounder of Qualtech Systems, Inc., a firm specializing in advanced integrated diagnostics software tools, namely TEAMS, TEAMS-RT, TEAMS-RDS, TEAMATE, and PackNGo, and serves on the Board of Aptima, Inc. His research interests include the application of systems theory, optimization, and inference techniques to agile planning, anomaly detection, diagnostics, and prognostics. He has published over 500 scholarly journal articles and conference papers in these areas. He is an Elected Fellow of IEEE for his contributions to discrete-optimization algorithms for large-scale systems and team decision-making and of the Connecticut Academy of Science and Engineering. He was selected by the IEEE Systems, Man, and Cybernetics (SMC) Society as the Outstanding Young Engineer of 1984 and received the Centennial Key to the Future Award. He was a co-recipient of Andrew P. Sage Award for the Best SMC Transactions Paper, in 1999, the Barry Carlton Award for the Best AES Transactions Paper, in 2000, the 2002 and 2008 NASA Space Act Awards for "A Comprehensive Toolset for Model-based Health Monitoring and Diagnosis," and "Real-time Update of Fault-Test Dependencies of Dynamic Systems: A Comprehensive Toolset for Model-Based Health Monitoring and Diagnostics," the 2005 School of Engineering Outstanding Teaching Award, and the 2003 AAUP Research Excellence Award at UCONN. He has served as the Editor-in-Chief for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, from 1998 to 2001.