An Overview and General Framework for Spatiotemporal Modeling and Applications in Transportation and Public Health

Li, Lishuai; Tsui, Kwok Leung; Zhao, Yang

DOI
10.1007/978-3-031-07155-3_8

Publication date
2022

Document Version
Final published version

Published in
Artificial Intelligence, Big Data and Data Science in Statistics

Citation (APA)
Li, L., Tsui, K. L., & Zhao, Y. (2022). An Overview and General Framework for Spatiotemporal Modeling and Applications in Transportation and Public Health. In *Artificial Intelligence, Big Data and Data Science in Statistics: Challenges and Solutions in Environmetrics, the Natural Sciences and Technology* (pp. 195-226). Springer. https://doi.org/10.1007/978-3-031-07155-3_8

An Overview and General Framework for Spatiotemporal Modeling and Applications in Transportation and Public Health

Lishuai Li, Kwok-Leung Tsui, and Yang Zhao

Abstract Spatiotemporal modeling and forecasting is an essential task for many real-world problems, especially in the field of transportation and public health. The complex and dynamic patterns with dual attributes of time and space create unique challenges for effective modeling and forecasting. With the advancement of data collection, storage, and sharing technologies, the amount of data and the types of data available for spatiotemporal modeling research in transportation and public health are rapidly increasing. Some traditional spatiotemporal methods become obsolete. There is a need to review existing methods and propose new ones to harness the power of newly available data. Therefore, in this chapter, we conduct a comprehensive survey of methods and algorithms for spatiotemporal monitoring and forecasting, focusing on applications in transportation and public health. Then, we propose a systematic framework to incorporate three different approaches: statistical methods, machine learning methods, and mechanistic simulation methods. The proposed framework is expected to help researchers in the field to better formulate spatiotemporal problems, construct appropriate models, and facilitate new developments that combine the strengths of mechanistic approaches and data-driven ones. The proposed general framework is illustrated via examples of spatiotemporal methods developed in transportation and public health.

L. Li (✉)
Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Delft, Netherlands

School of Data Science, City University of Hong Kong, Hong Kong, China
e-mail: lishuai.li@cityu.edu.hk

K.-L. Tsui
Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
e-mail: kltsui@vt.edu

Y. Zhao
School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China
e-mail: zhaoy393@mail.sysu.edu.cn

A. Steland, K.-L. Tsui (eds.), *Artificial Intelligence, Big Data and Data Science in Statistics*, https://doi.org/10.1007/978-3-031-07155-3_8

Keywords Spatiotemporal modeling · Transporation · Public health · Statistical learning · Machine learning · Simulation

1 Introduction

A wide variety of events in the real-world problems are featured by spatiotemporal dynamics, such as traffic flows, population migration, infectious diseases transmission, diffusion of air pollutants, power grids failure, etc. The ubiquitous events with dual attributes of time and space have created challenges for effective forecasting methods to predict future changes in a timely manner. The objective of spatiotemporal monitoring and forecasting is to analyze such patterns in events at both temporal and spatial dimensions and to understand and predict current and future developments. A large number of approaches under various applications have been developed in the literature, including statistical methods, machine learning/deep learning methods, and simulation models. The methods and algorithms related to spatiotemporal modeling are overabundant, yet there is no systematic approach for researchers and practitioners to choose what to use, why and when the method(s) works for what kind of real-world problems.

Developing a generic framework for any spatiotemporal problems is overly ambitious; in this chapter, we focus on spatiotemporal modeling in two specific applications: transportation and public health, particularly infectious disease transmission. Both have significant societal and economic impact, and there are common characteristics in transportation flow and infectious disease spreading in terms of dynamics and modeling methods used. In transportation systems, congestion and delays feature typical spatiotemporal patterns and cause significant economic and environment costs. For road traffic, congestion cost the US economy nearly 87 billion dollars in 2018 [29]. For air traffic, domestic flight delays were found to cost the US economy 33 billion dollars in 2019 [27]. In public health, the outbreaks and prevalence of infectious disease can be highly life-threatening. Up to July in year 2021, the COVID-19 pandemic has resulted in around 4.2 million deaths worldwide. In year 2009, H1N1 pandemic resulted in between 151,700 and 575,400 deaths worldwide. It is crucial to efficiently and accurately detect and forecast the event occurrence patterns in these problems.

With the deployment of wireless sensors, decreasing cost of data transmission and storage, the data available for studying these spatiotemporal problems become increasingly large and diverse. This brings both opportunities and challenges for spatiotemporal modeling in transportation and public health. Existing spatiotemporal methods are not readily available to harness the power of emerging data for real-life situations. Existing studies generally adopt three approaches: (1) statistical models that consider mathematical description of the physical processes of related variable, (2) machine learning/deep learning approaches that utilize a complex model structure to estimate the future event occurrences, and (3) mechanistic/simulation approaches that built upon domain knowledge. For the statistical approaches, many models are limited to strong probability and distribution assumptions, which are not always valid for data from wireless sensors, text data, or data generated from human behaviors. They also have limited capability to model the spatiotemporal structure of multivariate data and are lack of flexibility to incorporate external factors that influence the spatiotemporal patterns. These limitations have prohibited broad applications of statistical methods involving high-dimensional variables, especially in complex socio-technical problems. For machine learning and deep learning methods, experience-based model construction, feature engineering, and parameter tuning have made it difficult for researchers and practitioners to generalize and implement for a wide range of applications. In addition, model interpretability is another major shortcoming of deep learning methods. The lack of a clear understanding of the model and the meaning of its results limits its deployment and impact in real world. At last, the mechanistic/simulation approaches are strong in explaining the underlying mechanism of the spatiotemporal dynamics, yet few of the traditional mechanistic/simulation approaches can harness the power of big data. To our best knowledge, few papers address both mechanistic and data-driven approaches.

Motivated by the above challenges, we review methods and algorithms for spatiotemporal monitoring and forecasting in transportation and public health applications. We propose a systematic framework to incorporate statistical methods, machine learning methods, and mechanistic simulation methods. The framework is not meant to solve any specific spatiotemporal problems, but rather to structure the problems, construct appropriate spatiotemporal models, and facilitate new developments that combine the strengths of mechanistic approaches and data-driven ones.

More specifically, we plan to achieve the following objectives in this chapter.

- Summarize existing spatiotemporal models and compare statistical learning, machine learning, and simulation methods. The goal is to understand what methods have been developed for what kind of problems, infer how they can complement each other, and suggest what new models can be developed to address problems in transportation and public health.
- Propose a systematic framework that integrates statistical models, machine learning methods, and simulation approaches for modeling spatiotemporal problems, focusing on applications in transportation and public health. The focus is on how to handle typical categories of spatiotemporal problems and what key steps are involved in spatiotemporal modeling.
- Illustrate the proposed method and strategy with examples in transportation and public health applications. A number of example methods are shown with different focuses and application areas.

This chapter is expected to be useful to researchers and practitioners to meet the increasing demand and challenges for spatiotemporal monitoring and forecasting in transportation and public health applications. The comprehensive analysis of spatiotemporal methods helps us to understand which methods work for what problems in real life. It also contributes to the development of robust methods with meaningful interpretability for spatiotemporal monitoring and forecasting problems in transportation and public health. New methods proposed in this framework are expected to incorporate multiple data sources with various data structure, reveal the inherent evolution of target event occurrences, and deliver accurate predictions of future changes for transportation and public health problems.

2 Literature Review

We review existing general methods for spatiotemporal modeling and forecasting and conduct an in-depth survey of methods to address spatiotemporal problems in the field of transportation and public health. Most of the existing methods for spatiotemporal modeling in general take the data-driven approach, including statistical approaches and machine learning/deep learning-based approaches. The survey of methods used in transportation and public health provided us new perspectives, particularity on the value of the mechanistic simulation approaches on spatiotemporal modeling.

2.1 Statistical Approaches to Spatiotemporal Modeling

Time series forecasting has fundamental importance to various practical domains [26, 64, 88]. Autoregressive integrated moving average (ARIMA) and its family is the most general class of models for forecasting time series. Hamed et al. [35] applied ARIMA model for short-term forecasting by using traffic volume data of urban arterials. Williams and Hoel [118] presented the theoretical basis for modeling univariate data streams as seasonal autoregressive integrated moving average (SARIMA) processes. ARIMA and SARIMA work well in specific conditions, but they are limited by the assumption of "stationary" data. Motivated by the superior capability to cast the regression problem of a Kalman Filter (KF) [44, 45], numerous KF-based prediction studies began to emerge [34, 76, 122]. However, traditional time series analysis methods cannot consider the spatiotemporal correlation.

Statistical spatiotemporal models are often constructed by combining time series models with variogram-based models. Popular time series approaches include autoregressive moving average models [11] for stationary data and state-space models [117] for non-stationary components. In the spatial setting, the early work often involves kriging-based models. Spatial processes consider the correlation depends on location as well as distance. For example, the STARMA [84] and STARMAX [98] models were constructed by adding spatial covariance matrices to standard vector autoregressive moving average models. However, they are limited to stationary temporal processes. Stroud et al. [99] developed a statistical model for non-stationary spatiotemporal data. The model is cast in a Gaussian state-space framework and can include temporal components such as trends, seasonal effects, and autoregressions. Some other methods [25, 56] use the vector autoregressive models (VARs) for spatiotemporal data, in which the Y variable is a vector of observations at difference sites at time t, and the coefficient matrices are carefully constructed to model the spatiotemporal autoregressive relationships.

More generally, regression analysis is widely used for prediction and forecasting [1, 6, 47, 70, 73, 96]. The major advantage of regression models is that they can be used to capture important relationships between the forecast variable of interest and the predictor variables. By modeling the spatial as well as the temporal dependence of the errors, [79] applied spatial–temporal regression with 14 variables to forecast real estate prices. Yang et al. [128] improved the accuracy of flu activity predictions by establishing an autoregressive model of Google search data as an external explanatory variable. Lu et al. [66] adapted a multivariate dynamic regression method integrating Google searches, Twitter posts, electronic health records, and a crowd-sourced influenza reporting system to forecast influenza activity. To estimate regional activity, [74] proposed a two-step augmented regression model that efficiently combines publicly available Google search data at different resolutions (national and regional) and spatial dependence of influenza transmission.

Probabilistic graphical models (PGMs) are a powerful framework that bring together graph theory and probability theory. Considering the uncertainty of the noisy data and simplifying the complexity of the real world are the main advantages of PGMs, [61] proposed dynamic cost predictions for a trip planner by using a spatiotemporal Markov random field (STMRF). Hoang et al. [40] proposed a Gaussian Markov random field-based model to forecast citywide crowd flows based on traffic big data. However, PGM is highly computationally complex at the training stage of the algorithm, making it very difficult to retrain the model when newer data become available.

A related topic to spatiotemporal modeling is change detection of spatiotemporal processes, yet it is not the focus of this chapter. Many methods have been proposed for change detection in monitoring industrial processes [85, 86], remote sensing with digital aerial imagery [2, 9], and other applications [92].

2.2 Machine Learning/Deep Learning-Based Approaches to Spatiotemporal Modeling

Machine learning/deep learning methods have advanced in many application fields over the past years, especially under the big data environment. Compared with conventional statistical methods, machine learning and deep learning methods have more flexibility in handling data with complex structure, such as graphs and networks. The well-known machine learning methods for forecasting include k-nearest neighbors (KNN) algorithm [14, 67, 136], support vector machine (SVM) [41, 135], and neural networks (NNs) [94, 112]. To evaluate a wider range of machine learning methods, [42] implemented a network of stacked sparse auto-encoders to detect and predict event occurrence. Besides, some researchers developed various extensions of SVM [21, 91, 114, 127] on short-term forecasting problem. In addition to these machine learning approaches, deep learning has attracted many attention and shown superior performance in spatiotemporal modeling.

Convolutional Neural Networks (CNNs) CNNs have been widely used in mining spatiotemporal data because it is effective in capturing the spatial correlations in the data [57]. Especially, for data types of spatial maps and spatial rasters, which can be represented as a two-dimensional matrix, CNN is well suited to learn the spatial features [17, 24, 51, 68, 69, 75]. Spatiotemporal data are sometimes represented as a tensor or a sequence of tensors, and three-dimensional CNNs can be used to learn the complex spatial and temporal dependencies of the data [16, 53].

Recurrent Neural Networks (RNNs) RNNs have been well recognized for sequence learning tasks [103]. Incorporating long short-term memory (LSTM) or gated recurrent unit (GRU) enables RNNs to capture the long-term temporal dependency of time series. RNN and LSTM are increasingly used in time series prediction. For example, RNNs are applied for future weather forecasting where the weather variables are modeled as time series [18]. Volkova et al. [113] evaluated the predictive power of neural network architectures based on LSTM units and demonstrate its capability of nowcasting and forecasting ILI dynamics. However, these algorithms cannot capture the spatial features.

Hybrid Models of CNN+RNN New hybrid models that combine CNN and RNN have been proposed to extract spatiotemporal dependencies simultaneously in spatiotemporal forecasting models. The basic idea is to structure the input as a sequence of image-like matrices, and then a hybrid model that combines CNN and RNN can be used, where CNN extracts the spatial relationships embedded in the matrices and RNN learns the temporal pattern from the sequences. For example, [121, 131] proposed the structures with the combination of CNN and LSTM for spatiotemporal forecasting. Instead of simply stacking the architectures of CNN and RNN, by extending the fully connected LSTM (FC-LSTM) to have convolutional structures in both the input-to-state and state-to-state transitions, [93] proposed the convolutional LSTM (ConvLSTM) model for the precipitation nowcasting problem. ConvLSTM was then used in the spatiotemporal forecasting on transportation applications [3, 48]. The hybrid architectures show good performance on extracting spatiotemporal dependencies and correlations in forecasting. But the training procedure may become time consuming as the size of dataset increases because the complexity of RNNs is determined by the size of data sequences. Furthermore, [137] proposed a spatiotemporal residual network for forecasting crowd flow in each regular region of a city, yet it cannot be adapted to deal with irregular regions.

Graph Neural Networks (GNNs) and the Hybrid Models of GNNs+RNNs CNNs are commonly applied for dealing with Euclidean data such as images, regular grids, etc. However, spatial features based on the topological structure of a network or a graph have strong effects on modeling graph-structured data. Graph Convolutional Networks (GCNs) were widely used to capture network-based spatial dependencies as GCNs can handle arbitrary graph-structured data. Zhao et al. [138] developed a spatiotemporal neural network named Temporal Graph Convolutional Network for forecasting problem, which combines the GCN with GRU. However, the experimental results show that it has difficulty to capture the sudden changes of events occurrence. Sun et al. [100] proposed a novel multi-view deep learning model, named Multi-View Graph Convolutional Network, to predict the inflow and outflow in each irregular region of a city. Besides, another graph neural network, Diffusion Convolutional Neural Networks (DCNNs) were also developed for graph-structured data [5]. Later on, [60] proposed the diffusion convolutional recurrent neural network (DCRNN) to model the traffic flow as a diffusion process on a directed graph and incorporate both spatial and temporal dependency in the traffic flow for traffic forecasting. More recently, attention mechanism was widely applied into temporal and graph-structured spatial dependencies' extractions. Spatiotemporal graph attention models were proposed for spatiotemporal forecasts of traffic states [80, 116].

2.3 Spatiotemporal Modeling in Transportation

In the field of transportation, spatiotemporal modeling is often used for the modeling and forecasting of (1) **traffic conditions** (including flow speed, volume, congestion level, etc.) and (2) **travel demand**.

Spatiotemporal modeling and forecasting for **traffic conditions** is a fast evolving field in recent years. Many papers have been published using machine learning or deep learning-based approach to do traffic condition forecasting. Ermagun and Levinson [25] provides a comprehensive literature review. The output of these models includes traffic flow [142], traffic speed [31], travel time [89], relative velocity [46], etc. The forecast time horizon is normally short term, e.g., a few minutes to 1 h. The modeling techniques are primarily data driven, ranging from statistical approaches [20] to machine learning/deep learning methods [130]. Many efforts have been made on developing effective methods to model the spatial dependency, the temporal dependency, and their dynamics for traffic condition forecasting. Regarding the spatial dependency in traffic condition forecasting, it can be coded as regions or positions in Euclidean coordinates without network structures. For example, [137] developed a deep learning-based approach, called ST-ResNet, to collectively forecast two types of crowd flows (i.e., inflow and outflow) in each and every region of a city. Yu et al. [131] proposed a network grid representation method for traffic speed forecasting on a transportation network. The network-wide traffic speeds are transformed into a series of images and fed into a deep architecture combining both convolutional neural networks (DCNNs) and long short-term memory (LSTM) neural networks for traffic forecasting. Meanwhile, networks are very common in transportation, such as highways, railroads, and airways. Some recent developments in traffic condition forecasting utilize the graph/network structure in the deep learning framework and show significant improvement in terms of forecasting accuracy [22, 140]. He et al. [36] focused on exploring the influencing factors on forecasting urban rail transit (URT) ridership. In this chapter, the authors proposed an approach based on spatial models considering spatial autocorrelation of variables, which outperform the traditional global regression model, OLS, in terms of model fitting and spatial explanatory power. He et al. [37] made an effort to incorporate multiple factors, including spatial factors (distance and network topology), temporal factors (e.g., period and trend), and external factors (e.g., land use and socioeconomics) to estimate metro ridership based on general estimating equation (GEE) models. A following study investigated local model selection in ridership prediction [38]. In this study, an adapted geographically weighted LASSO (Ada-GWL) framework is proposed for modeling subway ridership, which involves regression coefficient shrinkage and local model selection. It takes subway network layout into account and adopts network-based distance instead of Euclidean-based distance [38].

Similarly, in air transportation, many methods have been developed to monitor and forecast traffic flow, travel time, congestion level, and delay time. In the past, statistical methods or probabilistic approaches are used to analyze factors that influenced flight delays and estimate delay distributions [111, 123]. Several machine learning methods have been used to predict delays, e.g., k-nearest neighbors, neural networks, support vector machine, fuzzy logic, and tree-based methods, yet did not explicitly utilize the spatiotemporal patterns [19, 83, 87, 97]. With the vast volume of commercial aviation system data being collected, classic methods are not sufficient to incorporate the complexity involved in the real-world operational data collected from multiple sources. For instance, tracking of aircraft position becomes available for many areas in the world, which contains rich spatiotemporal information needed for delay predictions. However, how to use and model this kind of raw position data, as well as incorporating external data sources (e.g., weather, airline records), is still an open research question [133]. More recently, Kim et al. proposed a recurrent neural network (RNN) to predict the flight delays of an individual airport with day-to-day sequences [50].

Different from the data-driven approaches, various simulation models have been developed and commonly used for traffic condition forecasting in research as well as in practice. The simulation models are built upon system-level abstractions of real world (e.g., queue theory), component-level abstractions (agent-based), or a combination of both (e.g., delay propagation). The purpose of using simulation-based approaches for spatiotemporal modeling is to understand the "physics" of transportation systems, design for "optimality," and manage operations in real time. Some simulation models are based on representations of driver behavior, e.g., car following, gap acceptance, and lane choice. These are considered as microscopic models or agent-based simulations. Another type of simulation models is based on macroscopic traffic flow theory developed during the 1950s [62, 90]. The objective of these models is to represent temporal congestion phenomena over a road network based on system-level equations instead of individual vehicle level.

Regarding the **travel demand** modeling and forecasting, the conventional focus is on aggregate forecasts for transportation facilities, districts, cities, and regions, and it is normally for long-term forecasts, such as forecasting the travel demand for the next few year(s) [23, 78]. The basic principle used to construct travel demand models is similar to the one used in standard microeconomics, which is utility maximization. The travel activity is a reflection of explicit preferences and limitations [78]. Individuals are defined by socioeconomic variables. In the conventional models, the time is discretized into intervals and space into zones, the depiction of travel patterns can be trip-based, tour-based, or activity schedule-based, and the forecasting methods are regression-based [10]. More recently, there are a lot of developments in simulation-based approaches for travel demand modeling and forecasting [52, 129, 134].

2.4 Spatiotemporal Modeling in Public Health

In public health, considerable attention has been paid to spatiotemporal modeling approaches for real-time tracking, timely detection and early intervention of disease outbreaks and other public health-related events. Correspondingly, related papers can be broadly grouped into three categories: forecasting, surveillance, and simulation.

Forecasting The outbreaks and prevalence of infectious disease can be highly life-threatening. Therefore, the focus of papers in this category is on predictive modeling and forecasting of the spread of infectious disease and other public health-related events. One of the earliest key studies in this category is often referred to as "Google Flu" [33]. The authors proposed the first method to use Google search queries to track influenza-like illness and detect influenza epidemics in areas with a large population of web search users. Later, [128] proposed an influenza tracking model, ARGO (AutoRegression with GOogle search data), with significantly improved performance of Google search-based real-time tracking than other existing models for influenza epidemics at the national level of the USA, including Google Flu Trends. This first ARGO model only considers the temporal trends of influenza epidemics. To overcome this limitation, the ARGO2 (2-step Augmented Regression with GOogle data) was proposed to model both spatial and temporal trends [74]. ARGO2 can efficiently combine publicly available Google search data at different resolutions (national and regional) with traditional influenza surveillance data from the Centers for Disease Control and Prevention (CDC) for accurate, real-time regional tracking of influenza [74]. Along this line of research, there are many recent research studies, such as forecasting influenza in Hong Kong with Google search queries and statistical model fusion [64, 125], using Baidu index to nowcast hand–foot–mouth disease in China [139], a hybrid autoregressive integrated moving average–linear regression (ARIMA–LR) approach for forecasting patient visits in emergency department [124], and personalized health monitoring of elderly wellness at community level [132].

Surveillance The objective of public health surveillance is to systematically collect, analyze, and interpret public health data (chronic or infectious diseases) in order to understand trends and detect changes in disease incidence and death rates and for planning, implementation, and evaluation of public health practices. The focus is on the accurate detection of the time or/and location(s) of changes in the occurrence rate as soon as possible. Several review papers examined existing methods for public health surveillance and discussed research opportunities and challenges [107–110, 120]. The most common existing disease spread monitoring methods can be categorized into temporal, spatial, and spatiotemporal surveillance techniques. Most basic methods such as SPC, regression, time series, and forecast-based methods were originally developed as temporal approaches. On the other hand, popular public health surveillance methods such as scan statistics were originally developed as spatial approaches [54] and later extended as temporal and spatiotemporal approaches [55]. Most spatial surveillance techniques rely on existing statistical clustering methods. Many techniques have been developed to expand those models to spatiotemporal methods that also search for clusters in time. Despite many independent implementations of surveillance systems have been deployed across different disciplines, such as ESSENCE, Google Flu Trends, and Global Microbial, the ability to accurately detect infectious disease outbreaks and pandemics is still in its nascent stages. Current surveillance systems lack the means to integrate disparate data sources, although recently proposed methods for multivariate surveillance hold promise for deployment in future systems to provide accurate prediction for infectious disease outbreaks and spreading trends.

Simulation The third category includes studies that use simulation models or mechanistic models to describe the spread of infectious diseases. Mechanistic models are built with structures that make explicit hypotheses about the biological mechanisms that drive infection dynamics. Such hypotheses include the dynamics of disease process among individuals (e.g., susceptible, infected, immune) and social interactions of people in an entire country or even the world [59]. As early as 1930s, mechanical epidemic simulators are used as research tools as well as teaching tools for epidemic theory [58]. Since that time, modeling has become an integral part of epidemiology and public health. The history and typical methods of mechanistic models of infectious disease are reviewed in [58, 59]. The most commonly used mechanistic approach is the Susceptible-Exposed-Infectious-Recovered (SEIR) compartmental model [49]. Smieszek [95] presented a mechanistic model that considers the different duration and intensity of contacts. Balcan et al. [7] presented the Global Epidemic and Mobility (GLEaM) model to simulate the spread of epidemics at the worldwide scale. The model used in a spatially structured stochastic disease approach to integrate sociodemographic and population mobility data [7]. Andradóttir et al. [4] developed a stochastic, individual-level simulation model of influenza spread within a structured population to investigate reactive strategies for containing developing outbreaks of pandemic influenza [4].

The three categories of research, forecasting, surveillance and simulation tackle different aspects of the spatiotemporal problems in public health, therefore using different underlying methods. In the group of public health forecasting, the key question to answer is how would the trend develop in time and space for a particular kind of infectious diseases? Statistical or machine learning-based spatiotemporal methods are used in these studies. For the public heath surveillance problems, the focus is on detection, detecting the time and (or) location(s) of significant changes in disease incidence and death rates. Therefore, statistical process control (SPC) and related process monitoring methods are often used. Regarding the simulation models, the goal is to understand the dynamic nature of the process, to evaluate the impact of policy change over time, to develop better intervention strategies for epidemics. Therefore, the development of models that reveal the nature of infection dynamics is more important.

2.5 Challenges and Opportunities

To summarize, many methods for spatiotemporal forecasting have been applied in various contexts, and they generally adopt three approaches: (1) statistical models that consider mathematical description of the spatiotemporal process of related variables, (2) machine learning/deep learning approaches that use the complex model structures to learn spatiotemporal patterns and forecast future event occurrences, and (3) simulation models or mechanistic models built with structures describing the physical/biological mechanisms that drive dynamics. In the two specific applications, transportation and public health, classic statistical methods and simulation models have been used for traffic prediction, infectious disease forecasting, and other spatiotemporal modeling. Recent efforts have been made to utilize the power of artificial intelligence for spatiotemporal modeling and forecasting. Most papers focus on either data-driven approaches or mechanistic approaches. Few papers address both mechanistic and data-driven approaches as well as their interaction in feature engineering.

Each of the three approaches has its own limitations. For the statistical approaches, many models are limited to strong probability and distribution assumptions, which are not always valid for data from wireless sensors, text data, or data generated from human behaviors. They also have limited capability to model the spatiotemporal structure of multivariate data and are lack of flexibility to incorporate external factors that influence the spatiotemporal patterns. These limitations have prohibited broad applications of statistical methods involving high-dimensional variables, especially in complex socio-technical problems. For machine learning and deep learning methods, different methods/models work for different situations. Experience-based model construction, feature engineering, and parameter tuning have made it difficult for researchers and practitioners to generalize and implement for a wide range of applications. In addition, model interpretability is another major shortcoming of deep learning methods. The lack of a clear understanding of the model and the meaning of its results limits its deployment and impact in real world. The simulation models or mechanistic models require specific knowledge of particular problems, which is difficult to be scaled up to deal with different applications. To fill this gap, there is an urgent need to examine existing spatiotemporal monitoring and forecasting methods and develop improved solutions for practice and complex problems.

3 A General Framework for Spatiotemporal Modeling

With existing methods and algorithms, we propose a systematic framework for modeling spatiotemporal problems focusing on applications in transportation and public health. The framework covers statistical models, machine learning methods, and mechanistic simulations for spatiotemporal modeling. It introduces a unified way for effective and efficient mining and modeling of spatial and temporal dependences among diversified data sources while integrating domain knowledge and various forecasting methods. The proposed framework can serve as a guideline for researchers and practitioners to understand and structure the spatiotemporal problems they are facing and configure the modeling steps, i.e., feature engineering, model selection and fusion, parameter tuning, performance evaluation, and results interpretation. The proposed systematic framework for modeling spatiotemporal problems is shown in Fig. 1. The details of individual modules are explained in the following subsections.

3.1 Mechanistic/Simulation Approach

The mechanistic/simulation approach models the spatiotemporal dynamics based on explicitly theories or hypotheses about the traffic flow or the infectious diseases transmission mechanisms. Depending on whether the fundamental unit models the dynamics among individuals or the dynamics at a system level, agent-based simulations or system-level simulations are used. Many classic simulation models have been developed in the field of transportation as well as the field of public health, and new developments of mechanistic models and simulation are still evolving.

For the system (compartment)-level simulations, there are SEIR/SIR models [32] and stochastic Markov chain models [30, 77] in public health. Ghaffarzadegan [32] is an example of the system-level simulation model. The model analyzes the spread of COVID-19 in universities that can be used to conduct a what-if analysis and estimate infection cases under different policies. In transportation, the classic system-level simulations are LWR models [62, 90, 119] and traffic flow simulations [13, 82].

For agent-based (component-level) simulations, examples in public health include the FluTE model [15, 65], the EpiSimdemics simulations [8], GSAM [81], and Andradáttir's model [4]. Andradáttir et al. [4] is an example of the agent-based simulation model. It simulates the transmission of pandemic influenza with

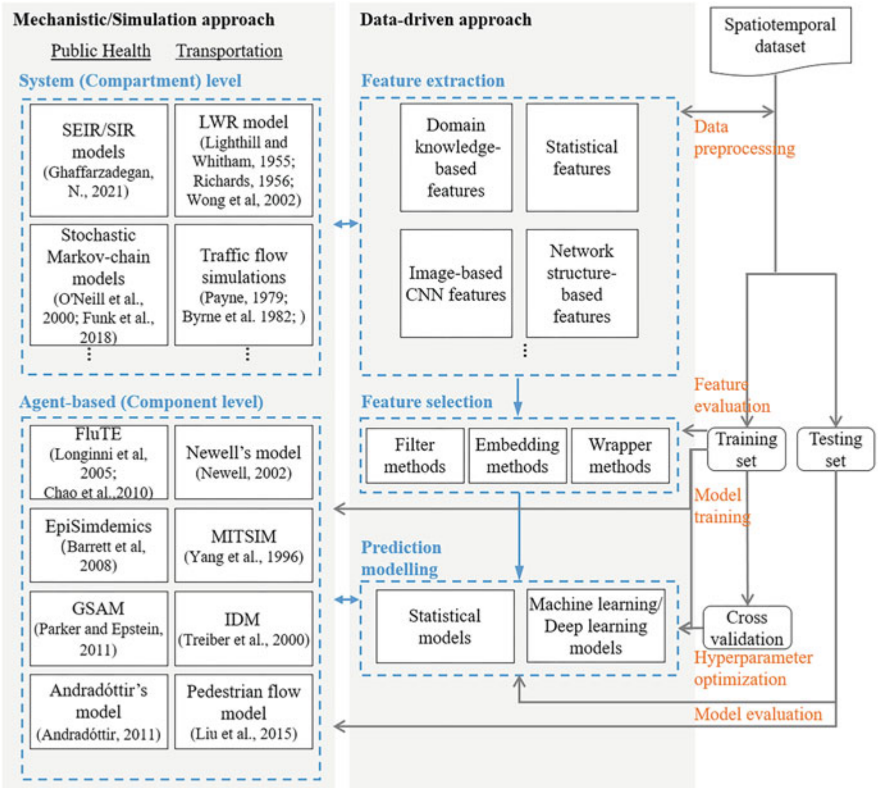

Fig. 1 A general framework for spatiotemporal modeling

the purpose of examining reactive strategies and concluded that in reaction to developing outbreaks combination strategies of reactive vaccination and limited antiviral use can be substantially more effective than vaccination alone in terms of controlling outbreaks and economic cost. In transportation, the well-known agent-based simulations include Newell's model [72], MITSIM [126], and IDM [106]. Liu et al. [63] is a new development that takes the mechanistic/simulation approach. The authors developed an agent-based simulation to model movement direction choice and collision avoidance for pedestrian flow. The results reveal the joint effect of several physical, psychological, and sociological factors dominating the real-world pedestrian walking behaviors.

Empirical data are important for the development of the mechanistic/simulation approach. The data are used to estimate and calibrate the chosen parameters in the model, as well as to validate the output accuracy. In the process of calibration and validation of mechanistic/simulation models, the focus is on the mechanism of interest to investigate and the intermediate steps, while the data-driven approach strives for the final step accuracy.

3.2 Data-Driven Approach

The data-driven approach focuses on the performance of final predictions, rather than the intermediate steps/processes involved in the spatiotemporal modeling. It does not require domain knowledge, yet its performance can be enhanced by domain knowledge. The typical steps involved in a data-driven approach include feature engineering, feature selection, and prediction model.

3.2.1 Feature Engineering

Feature engineering and feature extraction plays a key role in spatiotemporal modeling, directly affecting modeling accuracy, reliability, and generality. It is the process to generate informative features from the existing raw data by discovering hidden patterns inside them. For real-life problems, the types and characteristics of raw data for spatiotemporal modeling can be extremely diverse as they come from multiple sources with complex correlations. Most of existing spatiotemporal models in transportation and public heath still rely on a trial-and-error approach to choose and construct features.

Exogeneous Versus Endogenous Features The spatiotemporal features can be divided into two basic groups: exogenous features and endogenous features. An exogenous feature is one whose value is determined outside the model and is imposed on the model, in other words, features that affect a model without being affected by it. An endogenous feature correlates with other factors within the system being studied. It is changed or determined by its relationship with other features within the model.

For transportation applications, the information of weather, holiday, and calendar, as well as social media information containing the big events, etc. is the commonly used exogenous features for assisting the spatiotemporal modeling in transportation [48, 104, 137]. Sun et al. [102] is an example of using exogenous features, Baidu Index and Google Index, in the forecasting model. The Internet search data were integrated into extreme learning machine (KELM) models, and the forecasting performance was significantly improved in terms of both forecasting accuracy and robustness analysis.

For public health applications, the feature engineering part focuses on identifying exogenous features for improving the prediction accuracy, rather than incorporating complex network structures as in transportation models. Taking the forecasts of infectious disease as an example, Internet search index is the exogenous feature for improving the prediction accuracy of ILI rate [74, 128, 139].

Temporal Versus Spatial Within endogenous features, they can be further classified into (1) temporal features, which refers to any feature that is associated with or changes over time, and (2) spatial correlation features, which refers to any feature that is associated with or changes over space. Temporal features are typically classified as three different components: correlation with previous temporal measurements, upward or downward trends, and cyclical or seasonal pattern. Some spatial features (Euclidean-based) are better represented by Euclidean-based distance measures or an image-like matrix. Other spatial features (network-based) are better captured by the topological structure of a network or a graph. In addition, the spatial structure may change over time. Thus, special feature extraction and modeling techniques are needed to deal with the dynamic spatial structure.

Spatial Versus Network Features When choosing between image-based CNN features and network structure-based features, it mainly depends on how to characterize different types of spatial dependency in the forecasting problem. There are mainly two classes, regular or network spatiotemporal problems, which require different sets of methods. Most existing spatiotemporal models for disease transmission forecasting do not consider the network features, while for transportation systems, the network effect is naturally embedded, e.g., road network, rail network, air traffic network. How to model the network structure also differs depending on the mode of transportation. Road traffic prediction focuses on traffic condition on links, while forecasting problems in rail operations and air transport pay more attention to the delays at stations and airports.

Casual Versus Correlated Features The input features of a data-driven approach model could be correlated with or (and) casual factors to the output variable(s). Correlation and causation can both exist at the same time. However, correlation does not imply causation. Causation means that one event or action causes another event to happen. Correlation simply means there is a relationship between two events or two variables. Most models in the data-driven approach are only assuming and checking correlations between input features and the output variable(s). In contrast, the mechanistic/simulation approach has a better chance to identify casual features because a physical or mechanistic process is explicitly modeled and causal relationships are normally used in these models.

Typical Ways to Extract Features from Raw Data There are four typical ways to construct these endogenous features from raw data: domain knowledge-based features, statistical features, image-based CNN features, and network structure-based features. Domain knowledge-based features are constructed using a set of variables (usually have physical meanings), rules, or mechanism based on the knowledge and experience of the specific system or problem. For example, [28] extracted 42 parameters as features for battery lifetime prediction based on domain knowledge. These parameters can effectively reflect the aging dynamics of lithium-ion batteries, such as the voltage, capacity, temperature, etc. This is an effective way to construct features as it utilizes the domain knowledge accumulated for many years. However, it requires deep understanding of the domain, which could be expensive and time consuming for less popular applications. Statistical features are commonly used in various applications, including mean, standard deviation, variance, skewness, and correlation coefficients, autoregressive coefficients. The advantage of using these statistical features is that they are not limited to a particular domain or system. However, the downside is that these statistical features may not be able to capture the hidden patterns and the fundamental structure of the system. For examples, in traffic forecasting, the raw data collected are normally flow-rates at major road segments. Statistical features of flow-rates are not enough to build a traffic forecasting model because flow-rate is not sufficient to determine the traffic condition—a small flow-rate value may correspond to either a very light traffic or a congested traffic [71].

3.2.2 Feature Selection

Feature selection is the method to reduce a large set of features to a small number of features. The reduced feature set size makes it computationally feasible and easier to interpret when using certain algorithms. It may also lead to better results by reducing overfitting.

Three typical methods are used for feature selection: (1) filter methods, (2) wrapper methods, and (3) embedded methods, as shown in Fig. 2. Filter methods assess feature importance based on some ranking criteria. Typical filter methods are ANOVA, Pearson correlation, variance threshold, and information gain. Wrapper methods evaluate and select feature subsets based on model performance. The most commonly used wrapper methods include forward selection, backward elimination, and bidirectional elimination (Stepwise Selection).

Embedding methods take all available features as input and perform feature selection in model training as part of the model construction process, e.g., LASSO, elastic net, decision tree, deep learning methods, etc.

In addition, dimension reduction techniques can also be broadly categorized under feature selection methods, such as PCA, SVD, autoencoders, etc. These methods transform the original features into other variables via parametric or nonparametric projection.

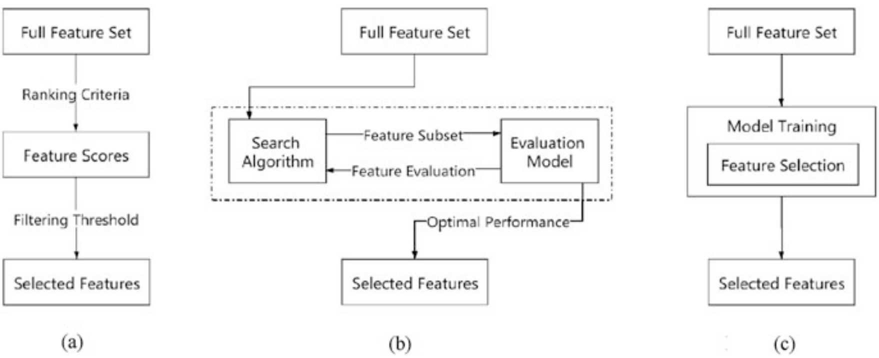

Fig. 2 Three categories of feature selection methods [28]. (**a**) Filter method. (**b**) Wrapper method. (**c**) Embedded method

3.2.3 Prediction Modeling

Prediction models take the input of selected features or transformed features, model the patterns and relationships among the features and their influence on the output variables (sometimes they are the input features with a different time window) in the training data, and predict the output variables. Both statistical methods and machine learning methods have been developed for prediction models.

Most statistical forecasting methods are developed based on autoregressive moving average models. However, these statistical forecasting methods are limited by the assumptions that they rely on, such as stationary of time series, known statistical distributions of features, etc. Such kind of methods have difficulty to incorporate complex spatiotemporal correlation into modeling.

Deep learning methods have advanced in many application fields over the past years, especially under the big data environment. Compared with conventional statistical methods, deep learning methods have more flexibility in handling data with complex structure, such as spatial data like maps, rasters, graphs and spatiotemporal data like sequence of spatial data, and 3D tensors. However, the training procedure may become time consuming as the size of dataset becomes very large. Besides, existing methods may have difficulty in state forecasting in irregular regions, directed network flow forecasting, or graph-structured data forecasting in non-Euclidian spaces. Many ongoing research efforts are made to develop better learning-based prediction models to overcome these difficulties.

3.3 Combining the Mechanistic and the Data-Driven Approach

A hybrid approach that involves both the mechanistic/simulation approach and the data-driven approach is expected to combine strengthens of both. For example, the mechanism and knowledge used in the mechanistic/simulation approach can be used to generate better domain knowledge-based features, as well as better structured features, such as network structure-based features. The formulated dynamics used in the mechanistic/simulation approach can also be used to design better prediction models in the data-driven approach, such as the structure of the statistical models, the architecture of the deep learning models, etc.

3.4 Evaluation Metrics and Methods

The standard way of evaluating the spatiotemporal forecasting models is to test the prediction accuracy on testing dataset. In the data preprocessing part, the original spatiotemporal dataset is divided into training set and testing set. Training set is used for feature evaluation, model training, and hyper-parameter optimization, while testing set is used for the model evaluation.

The model performance is typically evaluated based on three evaluation metrics: RMSE (root-mean-squared error), MAPE (mean absolute percentage error), and R2 (coefficient of determination). The RMSE describes the absolute error measured by variations in data errors, and the MAPE indicates the relative error, in terms of percentages of predicted values. Smaller values of RMSE and MAPE indicate better prediction accuracy. R2 measures how much of the variance in the predicted variables can be explained by the model. A larger value of R2 represents better performance. To evaluate the spatiotemporal forecasting models comprehensively, the models can also be assessed from other several aspects in addition to the prediction accuracy and prediction robustness, including computational efficiency, model interpretability, etc.

One important aspect of spatiotemporal forecasting models is about the prediction performance across different forecast horizons. The comparison of short-term and long-term predictions can help provide a better understanding of the robustness and adaptability of forecasting methods. For instance, an influenza forecast model might have different performance when forecast influenza activity 1 week ahead, 1 month ahead, and 1 year ahead. However, the natural difference in the structures of statistical models, deep learning models, and the mechanistic simulation models may cause unfair comparisons in sliding window evaluation. Most statistical models are naturally structured for sliding window evaluation. Many of these models are also capable of doing one-step or multi-step ahead forecasting with parameter re-estimation. Model parameters can be adjusted based on newly available data after each window sliding is performed. However, conventional machine learning/deep learning-based models reuse data at time t-1, t-2, . . . to predict output values at time t. The hyper-parameters of the model are not refitted as the prediction time shifts in sliding window evaluations unless the training process is explicitly re-performed on the new training set. Moreover, the comparison of mechanistic forecasting model may be different from the other types of forecasting models. In the simulation-based forecasting, the changes in policy and behavior could be factored in the model, while the data-driven approach does not have the natural structure to do this. Therefore, the comparison among different methods needs to be carefully designed to make it fair.

4 Examples of the General Framework for Spatiotemporal Modeling

Many real-world data sets are available for research in transportation and public health applications in recent years, including ILI data at different region levels, inter-city passenger travel data, the smart card records data collected from metro Automatic Fare Collection (AFC) system, and datasets for external factors, such as Google search and meteorological data. We performed a number of studies following proposed systematic framework for spatiotemporal modeling to achieve better prediction results, understand the spatiotemporal patterns better, and generate application insights. In this section, we illustrate the systematic framework for spatiotemporal modeling via several examples. The methods or models developed in these examples are different and they have unique characteristics best suited for a particular kind of application problem, since the framework is not meant to solve any specific spatiotemporal problems, but rather to structure the problems and construct appropriate spatiotemporal models.

4.1 Spatiotemporal Modeling for Road Traffic

This is a fast evolving field. Many papers have been published using machine learning or deep learning-based approaches to forecast traffic conditions on road network in recent years. A number of open datasets provide traffic speed and traffic flow over major road segments or intersections measured by sensors installed on the road or real-time information provided by the vehicles on the road. [43] provides a summary of open data and big data tools used for traffic estimation and prediction.

How to structure the problem, including the selection of the output variables, feature extraction, and the design of model structures, is critical for spatiotemporal modeling in the field of road transportation. Regarding the design and selection of the output variables, the output variable is normally flow, speed, congestion level, relative velocity, and other traffic condition measures. Yet it can be categorical or continuous. The selection of output variables depends on data availability and model specification.

In terms of feature extraction, most of the recent deep learning-based approaches to forecast traffic conditions on road network have been focused on how to extract and model spatial features in the road network. In addition to these endogenous features, [102] is an example of using exogenous features, Baidu Index and Google Index, in the forecasting model. The Internet search data were integrated into extreme learning machine (KELM) models, and the forecasting performance was significantly improved in terms of both forecasting accuracy and robustness analysis.

One of our ongoing work is the development of data-driven approaches to predict traffic condition on a city road network [115]. In this study, we are using a dataset provided by Baidu, named MapBJ, which provides the traffic conditions categorized into four levels (unblocked, slow, congested, extreme congested) over major roads in Beijing [18], and another dataset provided by DiDi Chuxing for a similar set of traffic condition measures in Xi'an. Following the proposed general framework, the challenges of developing a data-driven approach to predict traffic condition on a city road network come from how to structure the model so that it can capture the temporal dependency, the spatial dependency, and the changes of spatial dependency overt time over a road network. A number of deep learning model architectures are being examined, and below is an architecture that we proposed. The proposed architecture is named, periodic spatial–temporal deep neural network (PSTN) as

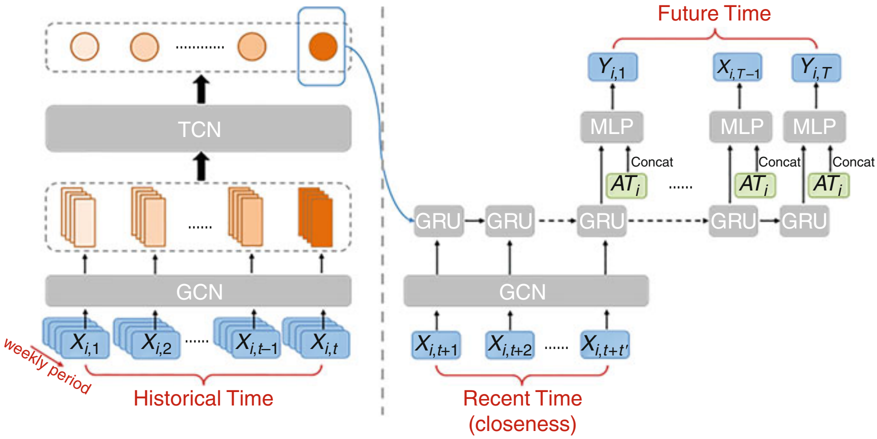

Fig. 3 Illustration of PSTN architecture of spatiotemporal model for road traffic. From [115], ©Tiange Wang, Zijun Zhang, Kwok-Leung Tsui, 2021, used under the Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/

shown in Fig. 3. The basic idea is to have three sequentially parts: (1) graph convolutional networks (GCNs) to capture the topological structure of road network, (2) the temporal convolutional networks (TCNs) and the gated recurrent units (GRUs) to capture periodic temporal dependency and local temporal dependency, respectively, and (3) the multi-layer perceptron (MLP) to combine road attributes and make the final prediction.

4.2 *Spatiotemporal Modeling for Transit Passenger Flow*

Forecasting short-term passenger flow on urban metro networks is an essential task for proactive traffic management in cities, which monitors real-time traffic conditions and forecast the condition in the immediate future. The challenges are mainly driven by complex spatiotemporal characteristics in metro passenger flow data and other external influence factors such as weather effect. We performed a number of studies following proposed systematic framework for spatiotemporal modeling to achieve better prediction results, understand the spatiotemporal patterns better, and generate application insights [36, 63, 104, 105].

Liu et al. [63] is an example of taking the mechanistic/simulation approach. The authors developed an agent-based simulation to model movement direction choice and collision avoidance for pedestrian flow. This developed microscopic simulation of pedestrian flow could be used for studying problems related to pedestrian traffic and evacuation dynamics.

Tang et al. [104] is an example of taking the statistical approach to forecast the short-term passenger flow on Shenzhen metro. In the proposed framework, there are three modules: traffic data profiling (feature engineering), feature extraction, and predictive modeling. In the feature engineering and feature extraction part, three types of features were comprehensively investigated in this study. They are (1) temporal features from passenger flow time series data, (2) spatial features based on origin-destination (OD) patterns, and (3) external weather factors. In the prediction model part, this study employed a number of forecasting models to evaluate the performance of the proposed framework, i.e., the time series model autoregressive integrated moving average, linear regression, and support vector regression. Moreover, the evaluation of this framework pays special attention to forecasting steps and horizons. The results suggest that smaller forecasting step predicts better for longer forecasting horizon, while larger forecasting step performs well for t + 1 prediction yet the prediction performance degrades when forecasting horizon grows.

Focusing on how to construct the features and extract the complex spatiotemporal relationships, we have studied both the statistical approach and the deep learning approach. He et al. [36] used a statistical approach, focusing on the travel demand forecasting and exploring the influencing factors on urban rail transit (URT) ridership. In this paper, the authors proposed an approach based on spatial models considering spatial autocorrelation of variables, which outperform the traditional global regression model, OLS, in terms of model fitting and spatial explanatory power. A following study investigated local model selection in ridership prediction [38]. In this study, an adapted geographically weighted LASSO (Ada-GWL) framework was proposed for modeling subway ridership, which involves regression coefficient shrinkage and local model selection. It takes subway network layout into account and adopts network-based distance instead of Euclidean-based distance. In addition, [37] made an effort to incorporate multiple factors, including spatial factors (distance and network topology), temporal factors (e.g., period and trend), and external factors (e.g., land use and socioeconomics) to estimate metro ridership based on general estimating equation (GEE) models.

He et al. [39] is an example of taking the deep learning approach for short-term passenger flow forecasting on Shenzhen metro. This work focused on investigating how to encode the network-based spatial features and other heterogeneous inter-station correlations in the model. The solution proposed in this work is a multi-graph convolutional recurrent neural network (MGC-RNN) (shown in Fig. 4) to generate multiple graphs that each represents a type of network structure and then to employ multiple parallel graph convolutional operators on multigraphs in the prediction model. This work illustrates that feature engineering and feature selection are both embedded in the deep learning-based prediction model. Specifically, by incorporating various types of inter-station correlations, temporal dependencies, and exogenous factors, the framework exhibits a possibility for multi-source heterogeneous data fusion in a big data environment.

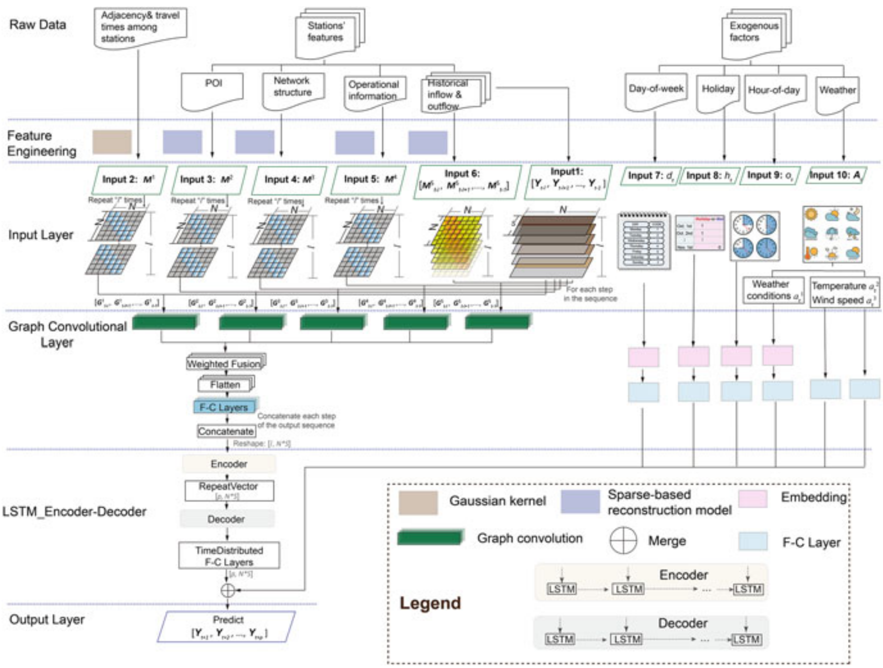

Fig. 4 The architecture of MGC-RNN [39]

4.3 *Spatiotemporal Modeling for Air Traffic*

Flight time prediction and, a related topic, flight delay prediction have been studied for years in the field of aviation. Many prior works employ statistical methods or probabilistic approaches. However, the accuracy of these models is not sufficient for the individual flight predictions. With the increasing amount of aviation system data being collected and available, such as Automatic Dependent Surveillance—Broadcast (ADS-B) data, aviation meteorological data, it is possible to utilize machine learning methods to learn the patterns of aircraft movement on a national air traffic network and predict individual flight time.

Sun et al. [101] is an example of combing statistical approach and deep learning approach to forecast air passenger flows. The proposed model incudes nonlinear vector autoregression and neural network. The results show that it outperforms single models and other hybrid approaches in terms of level forecasting accuracy, directional forecasting accuracy, and robustness analysis.

Zhu and Li [141] developed a novel spatial weighted recurrent neural network (SWRNN) model to provide flight time predictions for individual flights at a scale of national air traffic network, as shown in Fig. 5. Following the systematic framework for spatiotemporal modeling, the feature engineering part is a combination of domain knowledge-based and imagine-based CNN features. Based on domain

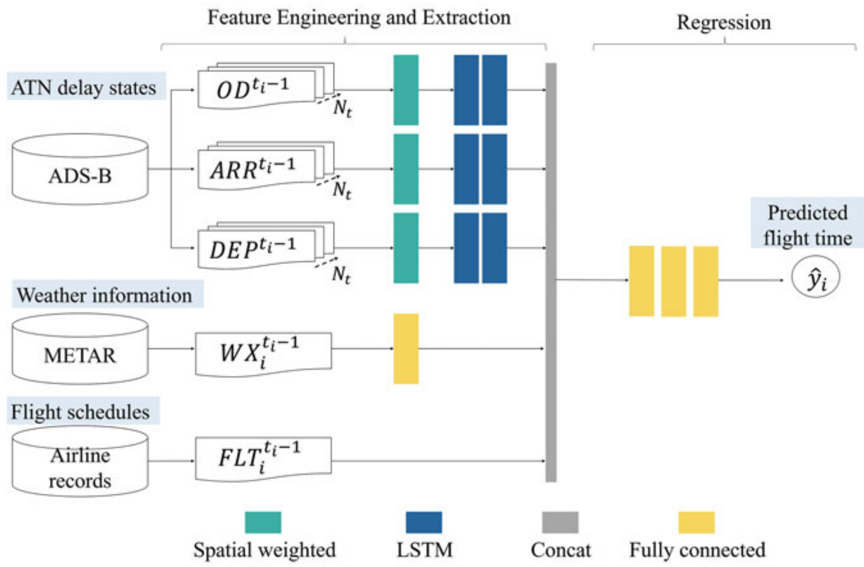

Fig. 5 Framework of SWRNN model [141]

knowledge, the network delay state features are extracted from the aircraft position tracking data, ADS-B, manually, including the average flight delay of each origin–destination (OD) pair, the average flight delay at each arrival airport, and the average flight delay at each departure airport for a specific time interval. Then, these network delay state matrices are sequenced based on time and fed into the spatial weighed layer to extract the spatial dependency and reduce the dimensionality for the network delay state features. The learnable weights of the spatial weighted layer show the importance of different OD pair/airports to the sample flight. Then, long short-term memory (LSTM) networks are used after the spatial weighted layer to extract the temporal dependency of network delay states. Therefore, the feature selection is an embedded method in this work. Finally, features from delays, weather, and flight schedules are fed into a fully connected neural network to predict the flight time of a particular flight. Evaluation of the SWRNN model was conducted using 1 year of historical data from an airline's real operations. Results show that the SWRNN model can provide a more accurate flight time predictions than baseline methods, especially for flights with extreme delays. In the paper, the authors also demonstrated that fuel loading can be optimized with the improved flight time prediction and resulting reduced fuel consumption by 0.016%–1.915% without increasing the fuel depletion risk for airlines.

4.4 Spatiotemporal Modeling for Infectious Disease Transmission

Taking the forecast of infectious disease transmission as an example, the systematic framework for spatiotemporal modeling can also be followed. Tsui et al. [108] provided a comprehensive review of research and developments in temporal and spatiotemporal surveillance for public health. Compared with the transportation problems, the feature engineering part focuses on identifying exogenous features for improving the prediction accuracy, rather than incorporating complex network structures. Several studies have shown that Google search data are effective exogenous features for improving the prediction accuracy of ILI rate [74, 128, 139].

Following the statistical approach, [125] studied the value of using online social media and web search queries to forecast new cases of influenza-like illness (ILI) in general outpatient clinics (GOPC) in Hong Kong. The study tested four individual models to forecast ILI-GOPC both 1 week and 2 weeks in advance, which are generalized linear model (GLM), least absolute shrinkage and selection operator (LASSO), autoregressive integrated moving average (ARIMA), and deep learning (DL) with feedforward neural networks (FNNs). Furthermore, the authors also proposed a statistical fusion model using Bayesian model averaging (BMA) to integrate multiple forecast scenarios.

Regarding the machine learning approach, [64] used a deep learning method to forecast influenza epidemics in Hong Kong, which also uses Google search queries. In this method, the innovative parts are mainly feature engineering on the output data. Variational mode decomposition (VMD), a signal decomposition method, is used to decompose the influenza data (the output data) into modes with different frequencies. Then, each mode extracted by VMD is forecasted by artificial neural networks (ANNs) and then these forecasts of each mode are added to generate the final forecasting results.

Zhao et al. [139] is an example of combining both the traditional statistical approach and the machine learning approach for spatiotemporal modeling of infectious disease transmissions. A meta learning framework (shown in Fig. 6) is proposed to select appropriate predictive model based on the statistical and time series meta features to nowcast the monthly hand, foot, and mouth disease (HFMD). In addition, the feature engineering part incorporated search engine index. The proposed meta learning method significantly improves the HFMD prediction accuracy, demonstrating that (1) the Internet-based information offers the possibility for effective HFMD nowcasts and (2) the meta learning approach is capable of adapting to a wide variety of data and enables selecting appropriate method for improving the nowcasting accuracy.

More recently, a study evaluated thirteen different methods for short-term forecasting of COVID-19 in Germany and Poland for 10 weeks, 12 October–19 December 2020, [12]. The study found that these forecasts from thirteen different teams are heterogeneous in terms of both point predictions and forecast spread. The performance of ensemble forecasts was relatively better on coverage, but ensemble

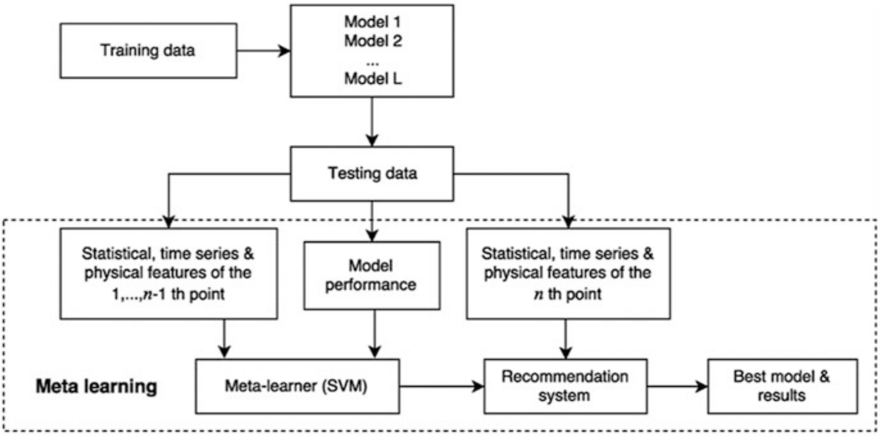

Fig. 6 Meta learning framework. From [139], ©The Author(s), 2018, used under the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

forecasts did not clearly dominate single-model predictions. The findings are consistent with the increasingly trend of acknowledgment that combining multiple models can improve the reliability of outputs.

5 Conclusion

There is an increasing demand for spatiotemporal monitoring and forecasting under various applications. This work reviews recent research and developments of spatiotemporal modeling in transportation and public health. Current spatiotemporal modeling methods are designed for specific applications, and various techniques and algorithms are proposed at different stages involved in the spatiotemporal modeling. This chapter proposes a systematic framework for developing spatiotemporal modeling, covering mechanistic/simulation approaches, statistical methods, and machine learning/deep learning methods. The proposed framework is illustrated via a few examples of spatiotemporal modeling in transportation and public health. The proposed framework will be useful to help researchers and practitioners formulate and structure the spatiotemporal modeling and forecasting problems, develop effective and accurate models, and improve the effectiveness of spatiotemporal modeling in solving real-life problems.

References

1. Abuella, M., Chowdhury, B.: Solar power probabilistic forecasting by using multiple linear regression analysis. In: SoutheastCon 2015, pp. 1–5. IEEE, New York (2015)
2. Agouris, P., Mountrakis, G., Stefanidis, A.: Automated spatiotemporal change detection in digital aerial imagery. In: Roper, W.E., Hamilton, M.K. (eds.) Automated Geo-Spatial Image and Data Exploitation, vol. 4054, pp. 2–12. International Society for Optics and Photonics, SPIE, New York (2000)
3. Ai, Y., Li, Z., Gan, M., Zhang, Y., Yu, D., Chen, W., Ju, Y.: A deep learning approach on short-term spatiotemporal distribution forecasting of dockless bike-sharing system. Neural Comput. Applic. **31**(5), 1665–1677 (2019)
4. Andradáttir, S., Chiu, W., Goldsman, D., Lee, M.L., Tsui, K.L., Sander, B., Fisman, D.N., Nizam, A.: Reactive strategies for containing developing outbreaks of pandemic influenza. BMC Public Health **11**(SUPPL. 1), 1–15 (2011)
5. Atwood, J., Towsley, D.: Diffusion-Convolutional neural networks. In: Advances in Neural Information Processing Systems (2016)
6. Austin, M.P., Nicholls, A.O., Margules, C.R.: Measurement of the realized qualitative niche: environmental niches of five Eucalyptus species. Ecol. Monogr. **60**(2), 161–177 (1990)
7. Balcan, D., Gonçalves, B., Hu, H., Ramasco, J.J., Colizza, V., Vespignani, A.: Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model. J. Comput. Sci. **1**(3), 132–145 (2010)
8. Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: EpiSimdemics: An efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In 2008 SC—International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2008 (2008)
9. Boulila, W., Farah, I.R., Ettabaa, K.S., Solaiman, B., Ghézala, H.B.: Improving spatiotemporal change detection: A high level fusion approach for discovering uncertain knowledge from satellite image databases. In: Icdm, vol. 9, pp. 222–227. Citeseer, New York (2009)
10. Bowman, J.L., Ben-Akiva, M.E.: Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. A Policy Pract. **35**(1), 1–28 (2001)
11. Box, G., Jenkins, G.M., Reinsel, G.C.: Time Series Analysis: Forecasting and Control, 3rd ed.. Prentice-Hall, Englewood Cliffs (1994)
12. Bracher, J., Wolffram, D., Deuschel, J., Görgen, K., Ketterer, J., Ullrich, A., Abbott, S., Barbarossa, M., Bertsimas, D., Bhatia, S., et al.: A pre-registered short-term forecasting study of COVID-19 in Germany and Poland during the second wave. Nat. Commun. **12**(1), 1–16 (2021)
13. Byrne, A.S., Courage, K.G., Wallace, C.E.: Handbook of computer models for traffic operations analysis. Technical report, Technology Sharing Report FHWA-TS-82-213, Washington, D.C. (1982)
14. Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., Sun, J.: A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. Transportation Research Part C: Emerging Technologies **62**, 21–34 (2016)
15. Chao, D.L., Halloran, M.E., Obenchain, V.J., Longini, I.M.: FluTE, a publicly available stochastic influenza epidemic simulation model. PLoS Comput. Biol. **6**(1), e1000656 (2010)
16. Chen, C., Li, K., Teo, S.G., Chen, G., Zou, X., Yang, X., Vijay, R.C., Feng, J., Zeng, Z.: Exploiting spatio-temporal correlations with multiple 3D convolutional neural networks for citywide vehicle flow prediction. Proceedings—IEEE International Conference on Data Mining, ICDM **2018-Novem**(61661146006), 893–898 (2018)
17. Chen, M., Yu, X., Liu, Y.: PCNN: Deep Convolutional Networks for Short-term Traffic Congestion Prediction. IEEE Trans. Intell. Transp. Syst. **19**(11), 3550–3559 (2020)
18. Cheng, X., Zhang, R., Zhou, J., Xu, W.: DeepTransport: learning spatial-temporal dependency for traffic condition forecasting. In: Proceedings of the International Joint Conference on Neural Networks, vol. 2018-July. Institute of Electrical and Electronics Engineers Inc. (2018)

19. Choi, S., Kim, Y.J., Briceno, S., Mavris, D.: Prediction of weather-induced airline delays based on machine learning algorithms. In: 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), pp. 1–6. IEEE, New York (2016)
20. Clark Dougherty, S.D., M.S., Kirby, H.R.: The Use of Neural Network and Time Series Modes for Short Term Forecasting: A Comparative Study (1993)
21. Cong, Y., Wang, J., Li, X.: Traffic flow forecasting by a least squares support vector machine with a fruit fly optimization algorithm. Procedia Engineering **137**, 59–68 (2016)
22. Cui, Z., Henrickson, K., Ke, R., Wang, Y.: Traffic Graph Convolutional Recurrent Neural Network: A Deep Learning Framework for Network-Scale Traffic Learning and Forecasting. IEEE Trans. Intell. Transp. Syst. **21**(11), 4883–4894 (2020)
23. Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J., Picado, R.: Synthesis of first practices and operational research approaches in activity-based travel demand modeling. Transp. Res. A Policy Pract. **41**(5), 464–488 (2007)
24. Duan, L., Hu, T., Cheng, E., Zhu, J., Gao, C.: Deep convolutional neural networks for spatiotemporal crime prediction. In: Proceedings of the International Conference on Information and Knowledge Engineering (IKE), pp. 61–67 (2017)
25. Ermagun, A., Levinson, D.: Spatiotemporal traffic forecasting: review and proposed directions. Transp. Rev. **38**(6), 786–814 (2018)
26. Fan, K., Li, C., Heller, K.: A unifying variational inference framework for Hierarchical Graph-Coupled HMM with an application to influenza infection. In: Thirtieth AAAI Conference on Artificial Intelligence (2016)
27. Federal Aviation Administration: Cost of Delay Estimates. Technical report, Federal Aviation Administration, Washington (2020)
28. Fei, Z., Yang, F., Tsui, K.L., Li, L., Zhang, Z.: Early prediction of battery lifetime via a machine learning based framework. Energy **225**, 120205 (2021)
29. Fleming, S. Traffic Congestion Cost the US Economy Nearly $87 Billion in 2018. Future of the environment, World Economic Forum (2019)
30. Funk, S., Camacho, A., Kucharski, A.J., Eggo, R.M., Edmunds, W.J.: Real-time forecasting of infectious disease dynamics with a stochastic semi-mechanistic model. Epidemics **22**, 56–61 (2018)
31. Fusco, G., Colombaroni, C., Isaenko, N.: Comparative analysis of implicit models for real-time short-term traffic predictions. IET Intell. Transp. Syst. **10**(4), 270–278 (2016)
32. Ghaffarzadegan, N.: Simulation-based what-if analysis for controlling the spread of Covid-19 in universities. PLoS One **16**(2 February), 1–24 (2021)
33. Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L.: Detecting influenza epidemics using search engine query data. Nature **457**(7232), 1012–1014 (2009)
34. Guo, J., Huang, W., Williams, B.M.: Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. Transportation Research Part C: Emerging Technologies **43**, 50–64 (2014)
35. Hamed, M.M., Al-Masaeid, H.R., Said, Z.M.B.: Short-term prediction of traffic volume in urban arterials. J. Transp. Eng. **121**(3), 249–254 (1995)
36. He, Y., Zhao, Y., Tsui, K.L.: Exploring influencing factors on transit ridership from a local perspective. Smart and Resilient Transport **1**(1), 2–16 (2019)
37. He, Y., Zhao, Y., Tsui, K.L.: Modeling and analyzing modeling and analyzing impact factors of metro station ridership: an approach based on a general estimating equation factors influencing metro station ridership: an approach based on general estimating equation. IEEE Intell. Transp. Syst. Mag. **12**(4), 195–207 (2020)
38. He, Y., Zhao, Y., Tsui, K.L.: An adapted geographically weighted LASSO (Ada-GWL) model for predicting subway ridership. Transportation **48**(3), 1185–1216 (2021)
39. He, Y., Li, L., Zhu, X., Tsui, K.L.: Multi-Graph Convolutional-Recurrent Neural Network (MGC-RNN) for Short-Term Forecasting of Transit Passenger Flow, arXiv:2107.13226, (2021)

40. Hoang, M.X., Zheng, Y., Singh, A.K.: FCCF: forecasting citywide crowd flows based on big data. In: Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, p. 6. ACM, New York (2016)
41. Hong, W.-C.: Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. Neurocomputing **74**(12–13), 2096–2107 (2011)
42. Jacobson, O., Dalianis, H.: Applying deep learning on electronic health records in Swedish to predict healthcare-associated infections. In: Proceedings of the 15th workshop on biomedical natural language processing, pp. 191–195 (2016)
43. Jiang, W., Jiayun L.: Big data for traffic estimation and prediction: a survey of data and tools. Appl. Syst. Innov. **5**(1) 23, (2022)
44. Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic Eng. **82**(1), 35–45 (1960)
45. Kalman, R.E., Bucy, R.S.: New results in linear filtering and prediction theory. J. Basic Eng. **83**(1), 95–108 (1961)
46. Kamarianakis, Y., Prastacos, P.: Forecasting traffic flow conditions in an Urban network: comparison of multivariate and univariate approaches. Transp. Res. Rec. **1**(1857), 74–84 (2003)
47. Kaytez, F., Taplamacioglu, M.C., Cam, E., Hardalac, F.: Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. Int. J. Electr. Power Energy Syst. **67**, 431–438 (2015)
48. Ke, J., Zheng, H., Yang, H., Chen, X.M.: Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. Transportation Research Part C: Emerging Technologies **85**(October), 591–608 (2017)
49. Kermack, W., Mckendrick, A.G.: A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **115**(772), 700–721 (1927)
50. Kim, Y.J., Choi, S., Briceno, S., Mavris, D.: A deep learning approach to flight delay prediction. In: AIAA/IEEE Digital Avionics Systems Conference—Proceedings, 2016-Decem:1–6 (2016)
51. Kira, Z., Li, W., Allen, R., Wagner, A.R., Georgia, A.: Leveraging deep learning for spatio-temporal understanding of everyday environments. In: IJCAI Workshop on Deep Learning and Artificial Intelligence (2016)
52. Kitamura, R., Chen, C., Pendyala, R.M., Narayanan, R.: Micro-simulation of daily activity-travel patterns for travel demand forecasting. Transportation **27**(1), 25–51 (2000)
53. Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., Biller, A.: Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. NeuroImage **129**, 460–469 (2016)
54. Kulldorff, M.: A spatial scan statistic. Communications in Statistics—Theory and Methods **26**(6), 1481–1496 (1997)
55. Kulldorff, M.: Prospective time periodic geographical disease surveillance using a scan statistic. J. R. Stat. Soc. Ser. A Stat. Soc. **164**(1), 61–72 (2001)
56. Lamberti, A., Naccarato, A.: VAR models for spatio-temporal structures: An application to environmental data. In: Studies in Classification, Data Analysis, and Knowledge Organization (2005)
57. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to digit recognition. Neural Comput. **1**(4), 541–551 (1989)
58. Lessler, J., Cummings, D.A.: Mechanistic models of infectious disease and their impact on public health. Am. J. Epidemiol. **183**(5), 415–422 (2016)
59. Lessler, J., Azman, A.S., Grabowski, M.K., Salje, H., Rodriguez-Barraquer, I.: Trends in the mechanistic and dynamic modeling of infectious diseases. Current Epidemiology Reports **3**(3), 212–222 (2016)
60. Li, Y., Yu, R., Shahabi, C., Liu, Y.: Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. arXiv preprint arXiv:1707.01926, pp. 1–16 (2017)

61. Liebig, T., Piatkowski, N., Bockermann, C., Morik, K.: Predictive trip planning-smart routing in smart cities. In: EDBT/ICDT Workshops, pp. 331–338 (2014)
62. Lighthill, M., Whitham, G.: On kinematic waves II: A theory of traffic flow on long crowded roads. Proc. R. Soc. Lond. A Math. Phys. Sci. **229**(1178), 317–345 (1955)
63. Liu, S.B., Lo, S.M., Tsui, K.L., Wang, W.L.: Modeling movement direction choice and collision avoidance in agent-based model for pedestrian flow. J. Transp. Eng. **141**(6), 04015001 (2015)
64. Liu, Y., Feng, G., Tsui, K.-L., Sun, S.: Forecasting influenza epidemics in Hong Kong using Google search queries data: a new integrated approach. Expert Systems with Applications **185**, 115604 (2021)
65. Longini, I.M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D.A., Halloran, M.E.: Containing pandemic influenza at the source. Science **309**(5737), 1083–1087 (2005)
66. Lu, F.S., Hou, S., Baltrusaitis, K., Shah, M., Leskovec, J., Hawkins, J., Brownstein, J., Conidi, G., Gunn, J., Gray, J.: Accurate influenza monitoring and forecasting using novel internet data streams: a case study in the Boston Metropolis. JMIR Public Health Surveill. **4**(1), e4 (2018)
67. Luo, X., Li, D., Yang, Y., Zhang, S.: Spatiotemporal traffic flow prediction with KNN and LSTM. J. Adv. Transp. 0197–6729 (2019)
68. Lv, J., Li, Q., Sun, Q., Wang, X.: T-CONV: A convolutional neural network for multi-scale taxi trajectory prediction. In: *Proceedings—2018 IEEE International Conference on Big Data and Smart Computing, BigComp 2018*, pp. 82–89 (2018)
69. Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., Wang, Y.: Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. Sensors **17**(4), 818 (2017)
70. Mao, L.-f., Jiang, Y.-c.: Medium-and long-term load forecasting based on partial least squares regression analysis. Power System Technology **32**(19), 71–77 (2008)
71. McCrea, J., Moutari, S.: A hybrid macroscopic-based model for traffic flow in road networks. Eur. J. Oper. Res. **207**(2), 676–684 (2010)
72. Newell, G.F.: A simplified car-following theory: A lower order model. Transp. Res. B Methodol. **36**(3), 195–205 (2002)
73. Ng, S.T., Cheung, S.O., Skitmore, M., Wong, T.C.Y.: An integrated regression analysis and time series model for construction tender price index forecasting. Constr. Manag. Econ. **22**(5), 483–493 (2004)
74. Ning, S., Yang, S., Kou, S.C.: Accurate regional influenza epidemics tracking using Internet search data. Sci. Rep. **9**(1), 1–8 (2019)
75. Niu, X., Zhu, Y., Zhang, X.: DeepSense: A novel learning mechanism for traffic prediction with taxi GPS traces. In: 2014 IEEE Global Communications Conference, GLOBECOM 2014, pp. 2745–2750 (2014)
76. Okutani, I., Stephanedes, Y.J.: Dynamic prediction of traffic volume through Kalman filtering theory. Transp. Res. B Methodol. **18**(1), 1–11 (1984)
77. O'Neill, P.D., Balding, D.J., Becker, N.G., Eerola, M., Mollison, D.: Analyses of infectious disease data from household outbreaks by Markov chain Monte Carlo methods. J. R. Stat. Soc.: Ser. C: Appl. Stat. **49**(4), 517–542 (2000)
78. Oppenheim, N.: Urban Travel Demand Modeling: From Individual Choices to General Equilibrium. Wiley, New York (1995)
79. Pace, R.K., Barry, R., Gilley, O.W., Sirmans, C.F.: A method for spatial–temporal forecasting with an application to real estate prices. Int. J. Forecast. **16**(2), 229–246 (2000)
80. Park, C., Lee, C., Bahng, H., Tae, Y., Jin, S., Kim, K., Ko, S., Choo, J.: ST-GRAT: A Novel Spatio-temporal graph attention networks for accurately forecasting dynamically changing road speed. In: Proceedings of theInternational Conference on Information and Knowledge Management, pp. 1215–1224 (2020)
81. Parker, J., Epstein, J.M.: A distributed platform for global-scale agent-based models of disease transmission. ACM Transactions on Modeling and Computer Simulation: A Publication of the Association for Computing Machinery **22**(1), 2 (2011)

82. Payne, H.J.: FREFLO: A macroscopic simulation model of freeway traffic. Transp. Res. Rec. **722**, 68–77 (1979)
83. Pérez–Rodríguez, J.V., Pérez–Sánchez, J.M., Gómez–Déniz, E.: Modelling the asymmetric probabilistic delay of aircraft arrival. J. Air Transp. Manag. **62**, 90–98 (2017)
84. Pfeifer, P.E., Deutsch, S.J.: A STARIMA model-building procedure with application to description and regional forecasting. Trans. Inst. Brit. Geogr. **5**(3), 330–349 (1980)
85. Prause, A., Steland, A.: Detecting changes in spatial-temporal image data based on quadratic forms. In: Steland, A., Rafajłowicz, E., Szajowski, K. (eds.) Stochastic Models, Statistics and Their Applications, pp. 139–147. Springer International Publishing, Cham (2015)
86. Rafajłowicz, E.: Detection of essential changes in spatio-temporal processes with applications to camera based quality control. In: Steland, A., Rafajłowicz, E., Szajowski, K. (eds.) Stochastic Models, Statistics and Their Applications, pp. 433–440. Springer International Publishing, Cham (2015)
87. Rebollo, J.J., Balakrishnan, H.: Characterization and prediction of air traffic delays. Transportation Research Part C: Emerging Technologies **44**, 231–241 (2014)
88. Reis, B.Y., Mandl, K.D.: Time series modeling for syndromic surveillance. BMC Med. Inform. Decis. Mak. **3**(1), 2 (2003)
89. Reza, R.M.Z., Pulugurtha, S.S., Duddu, V.R.: ARIMA Model for Forecasting Short-Term Travel Time due to Incidents in Spatio-Temporal Context. In: 94th Annual Meeting of the Transportation Research Board, vol. 257 (2015)
90. Richards, P.I.: Shock waves on the highway. Oper. Res. **4**(1), 42–51 (1956)
91. Santillana, M., Nguyen, A.T., Dredze, M., Paul, M.J., Nsoesie, E.O., Brownstein, J.S.: Combining search, social media, traditional data sources to improve influenza surveillance. PLoS Comput. Biol. **11**(10), e1004513 (2015)
92. Shekhar, S., Jiang, Z., Ali, R.Y., Eftelioglu, E., Tang, X., Gunturi, V., Zhou, X.: Spatiotemporal data mining: A computational perspective. ISPRS Int. J. Geo Inf. **4**(4), 2306–2338 (2015)
93. Shi, X., Chen, Z., Wang, H., Yeung, D.-Y.: Convolutional LSTM network: a machine learning approach for precipitation nowcasting. Adv. Neural Inf. Proces. Syst. **28**, 802–810 (2015)
94. Sirvio, K., Hollmén, J.: Spatio-temporal road condition forecasting with Markov chains and artificial neural networks. In: International Workshop on Hybrid Artificial Intelligence Systems, pp. 204–211. Springer, Berlin (2008)
95. Smieszek, T.: A mechanistic model of infection: Why duration and intensity of contacts should be included in models of disease spread. Theor. Biol. Med. Model. **6**(1), 1–10 (2009)
96. Song, K.-B., Baek, Y.-S., Hong, D.H., Jang, G.: Short-term load forecasting for the holidays using fuzzy linear regression method. IEEE Trans. Power Syst. **20**(1), 96–101 (2005)
97. Sternberg, A., Soares, J., Carvalho, D., Ogasawara, E.: A review on flight delay prediction. arXiv preprint arXiv:1703.06118, pp. 1–21 (2017)
98. Stoffer, D.S.: Estimation and identification of space-time ARMAX models in the presence of missing data. J. Am. Stat. Assoc. **81**(395), 762–772 (1986)
99. Stroud, J.R., Müller, P., Sansö, B.: Dynamic models for spatiotemporal data. J. R. Stat. Soc. Ser. B Stat Methodol. **63**(4), 673–689 (2001)
100. Sun, J., Zhang, J., Li, Q., Yi, X., Zheng, Y.: Predicting Citywide Crowd Flows in Irregular Regions Using Multi-View Graph Convolutional Networks. arXiv preprint arXiv:1903.07789 (2019)
101. Sun, S., Lu, H., Tsui, K.L., Wang, S.: Nonlinear vector auto-regression neural network for forecasting air passenger flow. J. Air Transp. Manag. **78**(September 2018), 54–62 (2019)
102. Sun, S., Wei, Y., Tsui, K.L., Wang, S.: Forecasting tourist arrivals with machine learning and internet search index. Tour. Manag. **70**(July 2018), 1–10 (2019)
103. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. In: Advances in neural information processing systems, pp. 3104–3112 (2014)
104. Tang, L., Zhao, Y., Cabrera, J., Ma, J., Tsui, K.L.: Forecasting short-term passenger flow: an empirical study on Shenzhen Metro. IEEE Trans. Intell. Transp. Syst. **20**(10), 3613–3622 (2019)

105. Tang, L., Zhao, Y., Tsui, K.L., He, Y., Pan, L.: A clustering refinement approach for revealing urban spatial structure from smart card data. Applied Sciences (Switzerland) **10**(16), 5606 (2020)
106. Treiber, M., Hennecke, A., Helbing, D.: Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics **62**(2), 1805–1824 (2000)
107. Tsui, K.L., Chiu, W., Gierlich, P., Goldsman, D., Liu, X., Maschek, T.: A review of healthcare, public health, and syndromic surveillance. Qual. Eng. **20**(4), 435–450 (2008)
108. Tsui, K.L., Wong, S.Y., Jiang, W., Lin, C.J.: Recent research and developments in temporal and spatiotemporal surveillance for public health. IEEE Trans. Reliab. **60**(1), 49–58 (2011)
109. Tsui, K.L., Han, S.W., Jiang, W., Woodall, W.H.: A review and comparison of likelihood-based charting methods. IIE Transactions (Institute of Industrial Engineers) **44**(9), 724–743 (2012)
110. Tsui, K.L., Wong, Z.S.Y., Goldsman, D., Edesess, M.: Tracking infectious disease spread for global pandemic containment. IEEE Intell. Syst. **28**(6), 60–64 (2013)
111. Tu, Y., Ball, M.O., Jank, W.S.: Estimating flight departure delay distributions—A statistical approach with long-term trend and short-term pattern. J. Am. Stat. Assoc. **103**(481), 112–125 (2008)
112. Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C.: Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks. Comput. Aided Civ. Inf. Eng. **22**(5), 317–325 (2007)
113. Volkova, S., Ayton, E., Porterfield, K., Corley, C.D.: Forecasting influenza-like illness dynamics for military populations using neural networks and social media. PLoS One **12**(12), e0188941 (2017)
114. Wang, J., Shi, Q.: Short-term traffic speed forecasting hybrid model based on chaos–wavelet analysis-support vector machine theory. Transportation Research Part C: Emerging Technologies **27**, 219–232 (2013)
115. Wang, T., Zhang, Z., Tsui, K.-L.: PSTN: Periodic Spatial-Temporal Deep Neural Network for Traffic Condition Prediction, arXiv:2108.02424, (2021)
116. Wei, C., Sheng, J.: Spatial-temporal graph attention networks for traffic flow forecasting. IOP Conference Series: Earth and Environmental Science **587**(1), 1853–1862 (2020)
117. West, M., Harrison, J.: Bayesian Forecasting and Dynamic Models, 2nd edn. Springer, New York (1997)
118. Williams, B.M., Hoel, L.A.: Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. J. Transp. Eng. **129**(6), 664–672 (2003)
119. Wong, G.C., Wong, S.C.: A multi-class traffic flow model—An extension of LWR model with heterogeneous drivers. Transp. Res. A Policy Pract. **36**(9), 827–841 (2002)
120. Woodall, W.H., Tsui, K.L.: Comments on 'Some methodological issues in biosurveillance'. Stat. Med. **30**(5), 430–433 (2011)
121. Wu, Y., Tan, H.: Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework. arXiv preprint arXiv:1612.01022 (2016)
122. Xie, Y., Zhang, Y., Ye, Z.: Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. Comput. Aided Civ. Inf. Eng. **22**(5), 326–334 (2007)
123. Xu, N., Donohue, G., Laskey, K.B., Chen, C.H.: Estimation of delay propagation in the national aviation system using Bayesian networks. In: Proceedings of the 6th USA/Europe Air Traffic Management Research and Development Seminar, ATM 2005, pp. 353–363 (2005)
124. Xu, Q., Tsui, K.L., Jiang, W., Guo, H.: A Hybrid approach for forecasting patient visits in emergency department. Qual. Reliab. Eng. Int. **32**(8), 2751–2759 (2016)
125. Xu, Q., Gel, Y.R., Ramirez, L.L.R., Nezafati, K., Zhang, Q., Tsui, K.L.: Forecasting influenza in Hong Kong with Google search queries and statistical model fusion. PLoS One **12**(5), 1–17 (2017)

126. Yang, Q., Koutsopoulos, H.N.: A microscopic traffic simulator for evaluation of dynamic traffic management systems. Transportation Research Part C: Emerging Technologies **4**(3 PART C), 113–129 (1996)
127. Yang, Z., Mei, D., Yang, Q., Zhou, H., Li, X.: Traffic flow prediction model for large-scale road network based on cloud computing. Math. Probl. Eng. 926251 (2014)
128. Yang, S., Santillana, M., Kou, S.C.: Accurate estimation of influenza epidemics using Google search data via ARGO. Proc. Natl. Acad. Sci. U. S. A. **112**(47), 14473–14478 (2015)
129. Yin, W., Murray-Tuite, P., Ukkusuri, S.V., Gladwin, H.: An agent-based modeling system for travel demand simulation for hurricane evacuation. Transportation Research Part C: Emerging Technologies **42**, 44–59 (2014)
130. Yu, B., Song, X., Guan, F., Yang, Z., Yao, B.: k-nearest neighbor model for multiple-time-step prediction of short-term traffic condition. J. Transp. Eng. **142**(6), 04016018 (2016)
131. Yu, H., Wu, Z., Wang, S., Wang, Y., Ma, X.: Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. Sensors (Switzerland) **17**(7), 1501 (2017)
132. Yu, L., Chan, W.M., Zhao, Y., Tsui, K.L.: Personalized health monitoring system of elderly wellness at the community level in Hong Kong. IEEE Access **6**, 35558–35567 (2018)
133. Yu, B., Guo, Z., Asian, S., Wang, H., Chen, G.: Flight delay prediction for commercial air transport: a deep learning approach. Transportation Research Part E: Logistics and Transportation Review **125**(March), 203–221 (2019)
134. Zhang, L., Levinson, D.: Agent-based approach to travel demand modeling exploratory analysis. Transp. Res. Rec. **1898**, 28–36 (2004)
135. Zhang, Y., Xie, Y.: Forecasting of short-term freeway volume with v-support vector machines. Transp. Res. Rec. **2024**(1), 92–99 (2007)
136. Zhang, L., Liu, Q., Yang, W., Wei, N., Dong, D.: An improved k-nearest neighbor model for short-term traffic flow prediction. Procedia. Soc. Behav. Sci. **96**, 653–662 (2013)
137. Zhang, J., Zheng, Y., Qi, D., Li, R., Yi, X., Li, T.: Predicting citywide crowd flows using deep spatio-temporal residual networks. Artif. Intell. **259**, 147–166 (2018)
138. Zhao, L., Song, Y., Deng, M., Li, H.: Temporal graph convolutional network for urban traffic flow prediction method. arXiv preprint arXiv:1811.05320 (2018)
139. Zhao, Y., Xu, Q., Chen, Y., Tsui, K.L.: Using Baidu index to nowcast hand-foot-mouth disease in China: A meta learning approach. BMC Infect. Dis. **18**(1), 1–11 (2018)
140. Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H.: T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. IEEE Trans. Intell. Transp. Syst. **21**(9), 3848–3858 (2020)
141. Zhu, X., Li, L.: Flight time prediction for fuel loading decisions with a deep learning approach. Transportation Research Part C: Emerging Technologies **128**, 103179 (2021)
142. Zhu, Z., Peng, B., Xiong, C., Zhang, L.: Short-term traffic flow prediction with linear conditional Gaussian Bayesian network. J. Adv. Transp. **50**(6), 1111–1123 (2016)