Das M, Ghosh SK. Data-driven approaches for spatio-temporal analysis: A survey of the state-of-the-arts. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 35(3): 665–696 May 2020. DOI 10.1007/s11390-020-9349-0

Data-Driven Approaches for Spatio-Temporal Analysis: A Survey of the State-of-the-Arts

Monidipa Das[1], *Member*, *IEEE*, and Soumya K. Ghosh[2], *Senior Member*, *IEEE*

[1] *School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore*
[2] *Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India*

E-mail: monidipadas@ntu.edu.sg; skg@cse.iitkgp.ac.in

Received January 15, 2019; revised November 19, 2019.

Abstract With the advancement of telecommunications, sensor networks, crowd sourcing, and remote sensing technology in present days, there has been a tremendous growth in the volume of data having both spatial and temporal references. This huge volume of available spatio-temporal (ST) data along with the recent development of machine learning and computational intelligence techniques has incited the current research concerns in developing various data-driven models for extracting useful and interesting patterns, relationships, and knowledge embedded in such large ST datasets. In this survey, we provide a structured and systematic overview of the research on data-driven approaches for spatio-temporal data analysis. The focus is on outlining various state-of-the-art spatio-temporal data mining techniques, and their applications in various domains. We start with a brief overview of spatio-temporal data and various challenges in analyzing such data, and conclude by listing the current trends and future scopes of research in this multi-disciplinary area. Compared with other relevant surveys, this paper provides a comprehensive coverage of the techniques from both computational/methodological and application perspectives. We anticipate that the present survey will help in better understanding various directions in which research has been conducted to explore data-driven modeling for analyzing spatio-temporal data.

Keywords data-driven modeling, spatio-temporal data, prediction, change pattern detection, outlier detection, hotspot detection, partitioning/summarization, (tele-)coupling, visual analytics

1 Introduction

Spatio-temporal data analysis is a recently emerging area of research, centered on the development of advanced computational techniques for analyzing enormous set of spatio-temporal data. It has a huge scope in various domains including environmental management, transportation, epidemiology, climatology and so on. Weather prediction, traffic management, urban growth modeling, disease sprawl pattern analysis, MRI (Magnetic Resonance Imaging) analysis, crime pattern detection, flood monitoring, crowd behavior analysis, etc. are some typical application areas of spatio-temporal analysis. Several models, taking care of the underlying physics in such domain specific problems, have been proposed till date. For example, various global circulation models (GCMs), like CCSM, HadCM3 [1, 2], are physics-driven models which are widely used for climatological prediction. However, the two major limitations in such physics-driven models are as follows: firstly, these assume that all the physical systems are well understood, which is not true in reality; secondly, these models are computationally inefficient, requiring lots of computational power. Therefore, recently, the data-driven approaches have been emerged as a new paradigm in this regard, with an aim to extensively analyze historical data for generating insights, and utilize those in further studies.

The data-driven models involve mathematical equations that are derived not from the explicit knowledge of physical processes but by using empirical analysis [3]. The generic concepts and features of any data-driven approach have been summarized in Table 1. Recent developments in computational intelligence techniques, such as artificial neural network, probabilistic reasoning, fuzzy logic, genetic algorithms, chaos theory,

Survey

have greatly expanded the capabilities of data-driven modeling [4]. However, analyzing spatio-temporal data using data-driven approaches is still a challenging task. The major challenges here stem from the fact that unlike the classical datasets, these kinds of data are embedded in continuous space and also tend to show high auto-correlation. Moreover, such data are also not independent, and in most of the cases, these are influenced by various co-located variables in the spatial region of study. Further, the recent advancement in satellite and remote sensing technology has led to a huge amount of data availability, which is an added challenge in this regard. Therefore, the last few decades have encountered immense effort in exploring data-driven learning algorithms and their applications, with an objective to tackle the above-mentioned challenges.

Table 1. Physics-Driven vs Data-Driven Approaches

Categories of Approaches	Key Concepts and Features
Physics-driven approach [1, 2]	• Based on underlying physical properties of the system
	• Relying on existing model of the system dynamics
	• Prone to more uncertainty in modeling the system
Data-driven approach [3, 4]	• Based on various statistical or machine intelligence techniques
	• Relying on the available/historical data
	• Prone to less uncertainty in modeling the system

1.1 Spatio-Temporal Data

The spatio-temporal (ST) data involve variations across the space as well as the time. These can be defined as data to which labels have been assigned to indicate where and when these were collected. On the basis of extent of information available for spatial and temporal aspects of the data and how these aspects over both the dimensions are related to each other, ST data can be classified into three major categories (refer to Fig.1).

• *Spatio-Temporal (ST) Events*. These are the events associated with a location and a corresponding timestamp. However, these are static with respect to both space and time. In other words, these involve neither movement with respect to space nor any kind of evolution with respect to time. The geo-referenced record of an epidemic is a kind of ST event (refer to Fig.1(b)).

• *Spatial Time Series Data*. In this case, the space is fixed; however, the measurement value changes over a series of time (refer to Fig.1(a)). These are also termed as geo-referenced time series. Time series of precipitation data collected over various locations in a space, earth surface temperature data, etc. are some examples in this regard. The recent advancements in satellite remote sensing and spatially enabled sensors technology are the primary source of this kind of data.

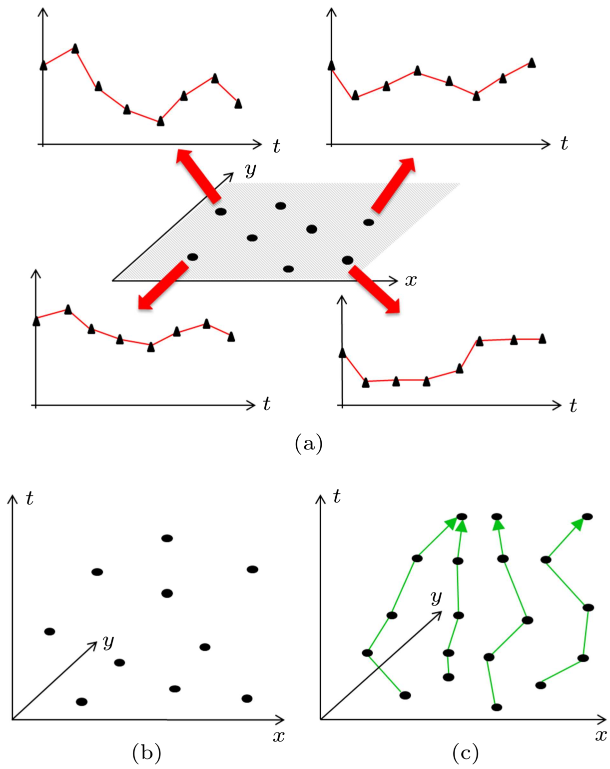

Fig.1. Three major categories of spatio-temporal data. (a) Spatial time series data. (b) ST events. (c) Moving object data.

• *Moving Objects Data*. In this case, the space is not fixed. Area of interest changes with the moving object. GPS track of a vehicle (trajectory data) is an appropriate example of such data. The development of wireless communication, mobile computing, and location sensing technologies has been the primary reason of a huge availability of these data in recent days.

It is to be noted that spatio-temporal data analysis is not a simple task because of the complex structure of the ST data which first needs appropriate representation, and indexing in the database. Depending on the characteristics of ST data, spatio-temporal database models can be classified into different categories among which event-oriented database (EOD) model, object-oriented database (OOD) model, and moving object database (MOD) model are most widely used in the realm of ST data analysis. In case of EOD models, each event, associated with changes that occurred since the last update of the event vector, is recorded sequentially in an increasing order of time. An EOD model supports neither the ST range queries nor the ST behavior-related queries. In the other case, the OOD models are built upon the concept of ST object, where the object has both spatial and temporal extents and represents the whole history of an entity. However, the OOD model does not support queries regarding the spatio-temporal behavior of the object. Contrarily, an MOD model supports both range and behavior-oriented queries. These are primarily designed to efficiently manage large volumes of trajectories and real-time GPS streams [5]. MODs can be either Euclidean-based or network-based, where the later offers more precision and efficiency in terms of location representation, storage, update, and indexing [6]. The existing MODs are mostly centralized, where the location update and the query processing are done at a single database server, which significantly affect the performance when the number of moving objects increases. Therefore, the recent research trend shows interest in designing parallel distributed network-constrained MOD [7], which can support location tracking as well as query processing in distributed and parallel fashion to achieve performance improvements over the centralized counterparts.

1.2 Spatio-Temporal Data Analysis

From the perspective of data mining, the spatio-temporal data analysis can be described as a process of identifying non-trivial, interesting and meaningful patterns from massive spatial/spatio-temporal databases. Based on the underlying objective and the output pattern of the process, spatio-temporal data analysis can be classified into six broad families [8] (refer to Fig.2 and Fig.3), as briefly described below.

1) *Spatio-Temporal* (*ST*) *Prediction*. Given spatio-temporal data items along with explanatory variables and a dependent (target) variable, the ST prediction is the process of learning a model, capable of predicting the dependent variable based on the explanatory variables [8].

2) *Spatio-Temporal* (*ST*) *Change Detection*. It basically refers to the ST change footprint discovery. Given a change definition and a dataset over a ST phenomenon, ST change detection aims to identify the location and/or time of such changes from the ST dataset [9].

3) *Spatio-Temporal* (*ST*) *Outlier Detection*. ST outlier detection is the process of identifying anomalous patterns from a given set of ST observations.

4) *Spatio-Temporal* (*ST*) *Hotspots Detection*. ST hotspots are the regions together with time intervals where object count is unexpectedly high. ST hotspot detection has huge application in epidemiology (finding disease hotspots) and criminology (finding crime hotspot).

5) *Spatio-Temporal Partitioning and Summarization*. ST partitioning is the process of grouping similar ST data items so as to partition the underlying space and time; whereas ST summarization is the process of generating a compact representation of ST data [10].

6) *Spatio-Temporal* (*ST*) *Coupling and Tele-Coupling*. ST coupling pattern represents the classes/types of spatio-temporal objects which often occur in close spatial and temporal proximity, whereas ST tele-coupling pattern indicates considerably high temporal correlation between time series data at long spatial distance [11].

Fig.3 exemplifies these major families of ST data analysis in terms of respective objectives and output patterns. The detailed statistical foundations for each of these ST data analysis families can be found in [8].

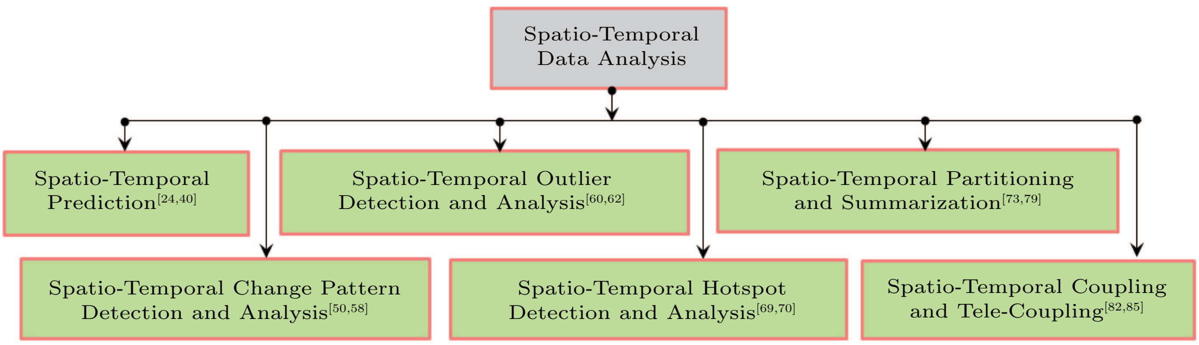

Fig.2. Spatio-temporal data analysis families.

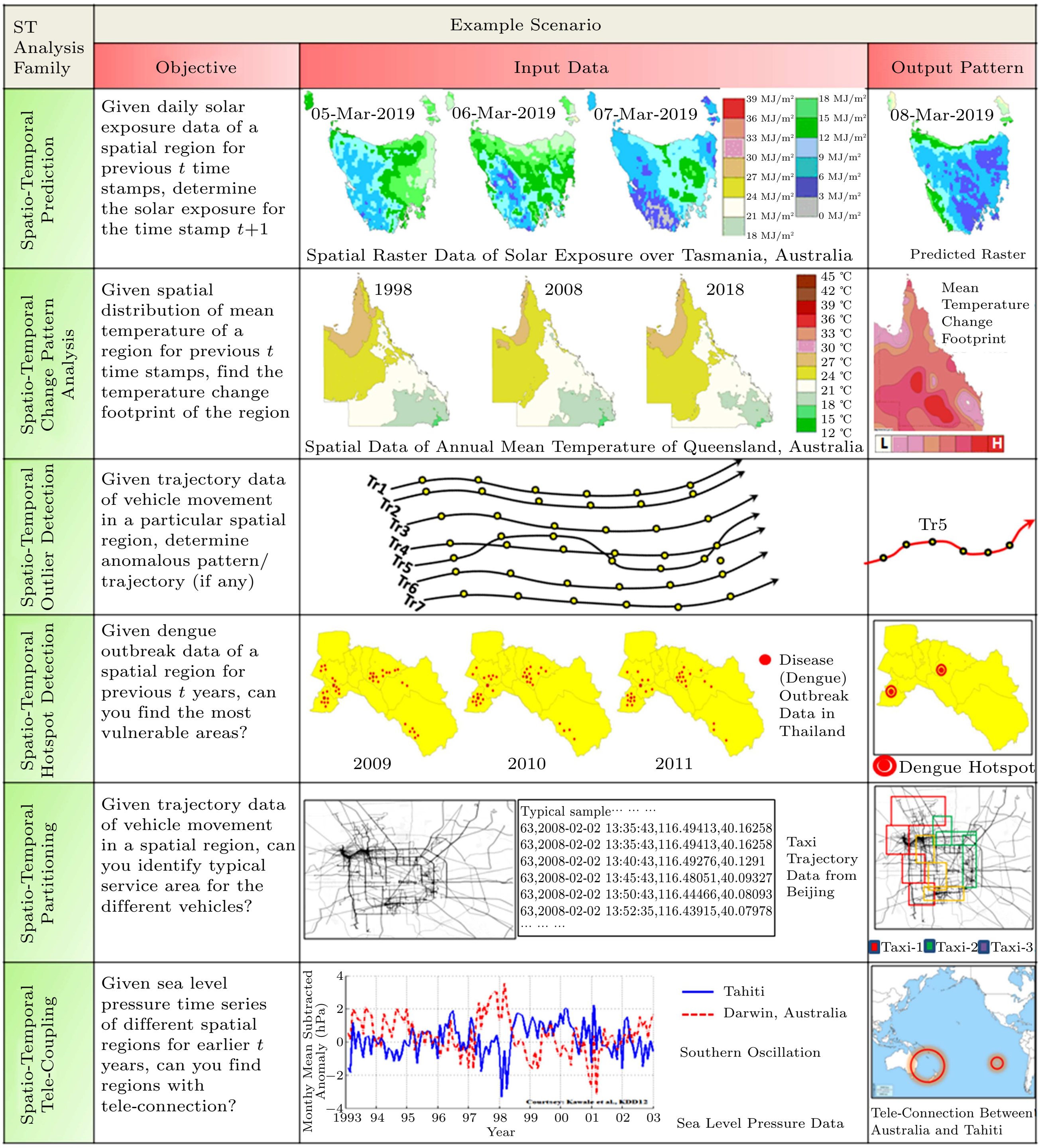

Fig.3. Example scenarios over major families of spatio-temporal data analysis.

1.3 Recent Surveys on Spatio-Temporal Data Analysis

Spatio-temporal data analysis is a recently emerging research field in computer science. However, within this short period of time, it has become a hot topic for a number of surveys, review articles, and books. Shekhar *et al.* provided an extensive survey of spatio-temporal data mining techniques, especially those with pure statistical backgrounds [8]. The authors of [8] mainly concentrated on the statistical foundations of these techniques with consideration to six major output pattern families. A broad review of spatio-temporal clustering, prediction, and visualization techniques, considering statistical as well as machine learning approaches has been presented by Cheng *et al.* in [12]. In [13], the author presented an extensive review of spatio-temporal outlier detection especially considering the trajectory data, weather data, and PET/MRI scans data. A comprehensive study of spatio-temporal clustering has been made by Kisilevich *et al.* in [10]. Substantial amount of research on spatio-temporal clustering over several application areas has been reported in their work. An extensive survey of spatio-temporal change pattern mining techniques from multi-disciplinary perspective has been conducted in [9]. Various ST data mining families and application areas, covered by our present survey and the other relevant survey articles, are summarized in Table 2.

1.4 Our Contributions

In this survey, we intend to provide a structured and wide-ranging overview of extensive research on data-driven approaches for analyzing spatio-temporal (ST) data, spanning multiple disciplines and application areas. Most of the existing surveys either focus on a single ST data analysis family along with extensive exploration of its domains of applications, or focus on statistical foundations of techniques under multiple ST data analysis family without much exploration of relevant application domains. The articles by Aggarwal [13], Kisilevich *et al.* [10], Zhou *et al.* [9], etc. are of the first category, whereas the survey work by Shekher *et al.* [8] is of the second category. On the contrary, our survey focuses on the state-of-the-art techniques from the perspectives of both ST data analysis family and application domain. Moreover, the present article covers traditional statistical, geo-statistical, computational intelligence based as well as deep learning based techniques, which is missing in a majority of the existing survey papers, even in those published recently [14].

Thus, the key contributions in our survey article can be summarized as follows.

• The present survey of data-driven approaches for ST data analysis focuses on the state-of-the-art tech-

Table 2. Comparison of Present Survey with Other Relevant Survey Articles

Survey Perspective	Covered Area	1	2	3	4	5	6	7
Spatio-temporal (ST)	ST prediction	√	√	√		√		
data mining families	ST change pattern mining	√	√	√			√	
	ST outlier detection	√	√	√	√			
	ST hotspot detection/clustering	√	√	√		√		√
	ST partitioning and summarization	√		√				
	ST coupling and tele-coupling	√		√				
Methodology	Pure/geo-statistical techniques	√	√	√	√	√	√	√
	Computational intelligence	√	√			√		
	Deep learning	√						
Application area	Climatology/meteorology	√	√	√	√	√	√	
	Hydrology	√						
	Environment and ecology	√	√	√	√		√	√
	Medical science and public health	√	√	√	√		√	
	Transport system	√	√	√	√	√		√
	Urban planning and development	√					√	
	Finance and economy	√		√	√	√		
	Bio-informatics	√	√					
	Molecular biology	√						
	Location-based services	√		√				√
	Mobility analysis	√	√		√			√
	Online and social network	√	√		√			
	Homeland security	√		√			√	√

Note: 1: our survey; 2: Atluri *et al.* (2018) [14]; 3: Shekhar *et al.* (2015) [8]; 4: Aggarwal (2015) [13]; 5: Cheng *et al.* (2014) [12]; 6: Zhou *et al.* (2014) [9]; 7: Kisilevich *et al.* (2009) [10].

niques from both analysis family and application domain perspectives. The survey covers six generic ST data analysis families and 13 application areas involving spatio-temporal aspects.

• For each of the six data analysis families, we have categorized the existing techniques into variants of basic techniques. This hierarchical structure provides a more crisp and comprehensible representation of various state-of-the-art techniques under each data analysis family.

• While the majority of the relevant surveys only mention the application areas of various data analysis techniques, we attempt to provide an exhaustive discussion of the domains where these approaches are applied. For each domain we have illustrated the nature of spatio-temporal data, several challenges faced in data analysis, and the set of techniques that have been employed.

• The existing surveys mostly discuss on the pure statistical/geo-statistical techniques for ST data analysis. However, our article provides an extensive coverage of computational intelligence based and deep learning based techniques as well.

• Further, it is evident from the survey statistics as depicted in Fig.4 that, unlike the considered survey papers ([8–10, 12–14]), we concentrate more on the research made in the current years than the work done in the last decades.

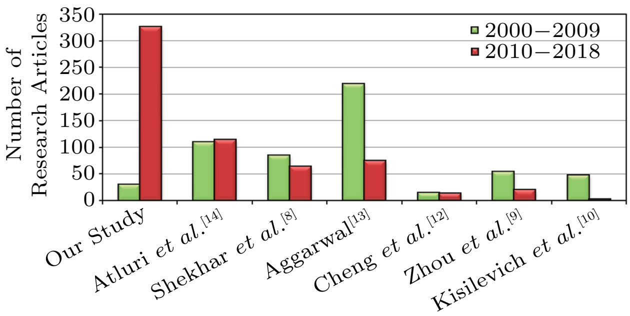

Fig.4. Survey statistics in terms of research articles considered from last two decades.

1.5 Organization

The remainder of this survey is structured as follows. Section 2 provides a discussion on the promises as well as major challenges in spatio-temporal data analysis. Section 3 surveys the state-of-the-arts in conventional data-driven approaches for spatio-temporal data analysis. Section 4 summarizes the state-of-the-art deep learning techniques applied for analyzing spatio-temporal data. Section 5 presents the ST data analysis approaches from the perspectives of their application areas. Section 6 identifies the recent trends and future scopes of research. Finally we conclude in Section 7. A schematic representation for the overall organization of the present survey paper is depicted in Fig.5.

2 Promises and Challenges of Spatio-Temporal Data Analysis

The spatio-temporal data are rich sources of knowledge and information, waiting to be discovered or extracted. An appropriate analysis of spatio-temporal data may provide several useful application-specific insights. For example, an accurate spatio-temporal prediction of climatological/meteorological data can detect extreme events like flood, drought, hurricane, and thereby can help in disaster planning; an unusual spatio-temporal pattern in medical data (ECG time series, MRI scans, PET scans, etc.) can reflect the disease conditions and thus can aid in diagnosis; the analysis of trajectory data collected from mobile devices or other sensors can reveal the category or motive of a person; the spatio-temporal analysis of phone calls made by people in a city can provide insights on urban activities; spatio-temporal data on crime events can be analyzed to locate crime generators in a city and thereby can help police department to take effective measures; spatio-temporal summarization process on traffic/road-network data can identify the accident prone routes and thus ensures the public safety. Many such examples with respect to the real-life events can be listed in this regard.

However, both the temporal and the spatial aspects add significantly high complexity to spatio-temporal data analysis/mining techniques. Various challenges in ST data mining are discussed in the subsequent part of this section.

2.1 Challenges

The major challenges in spatio-temporal data mining arise mainly because of the nature/characteristics of the ST data itself.

• First of all, unlike the traditional data, the spatio-temporal data follow the first law of geography, i.e., the data that are in more spatio-temporal proximity are more likely to be similar. For example, the weather of a day is more similar to that of the previous day. Likewise, the land surface temperature of one location is more likely to be the same as that of its nearby locations. This is primarily termed as autocorrelation

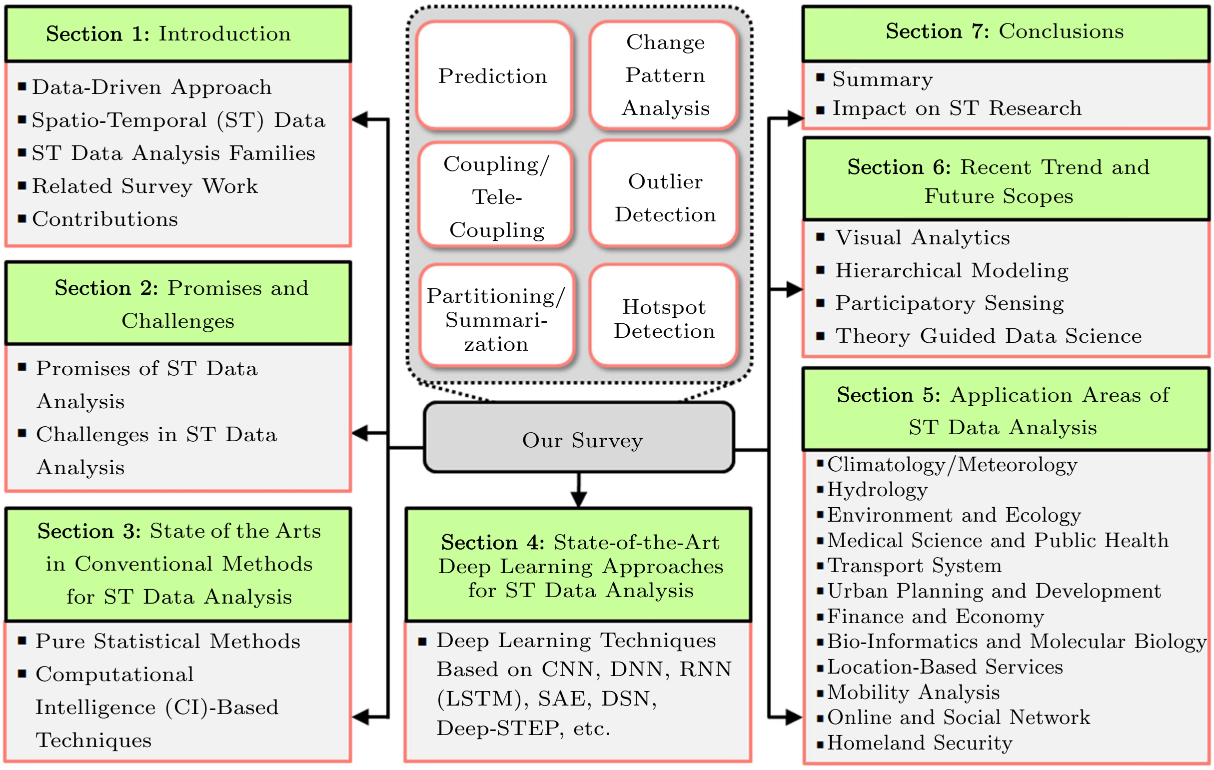

Fig.5. Overall organization of our survey paper.

property [15], and it indicates that spatio-temporal data should not be assumed as statistically independent data during the analysis or modeling process.

• Secondly, the spatio-temporal phenomena are not "concrete objects" [16]. These are continuous patterns that evolve over space and time, and can be well captured by existing physics-driven approaches using differential equations. However, solving differential equation is expensive and also suffers from several well-known limitations. Hence, providing an alternative means of modeling ST phenomena becomes a challenging task.

• Thirdly, the spatial/spatio-temporal data sometimes show inter-dependency with the co-located variables. Therefore, instead of only dealing with the target variable, considering the effects of other influencing variables may improve the results of spatio-temporal data mining. A proper modeling of such spatio-temporal interrelationships among the variables is also a critical issue.

• Apart from the above-mentioned common properties, a particular kind of spatio-temporal data may also have its own special properties. For example, in the case of climatological data, there are certain central processes that affect the climate system in such a complicated manner that the data become inherently chaotic in nature. Besides, it may happen that a similar pattern, that used to take place in distant past, may again be repeated in recent days. Therefore, the climatological data may sometimes show long memory time series effect. All these inherent as well as special properties of spatio-temporal data make the analysis process a complicated task.

• Besides, in most of the cases, the spatio-temporal data are relatively abundant in either space, or time, but not in both [17]. For example, the satellite remote sensing imagery is significantly profuse in space, providing a detailed view of large areas. However, these are relatively scarce with respect to time. On the other hand, the data from fixed sensors are plentifully available over time. However, these provide relatively little detail in space, because of limitation in the number of spatially distributed sensors.

• Further, the recent advancement in satellite and remote sensing technology has led to explosive growth in spatial and spatio-temporal data. This avalanche of data is also an added challenge in the present context.

In the next two sections we discuss the state-of-the-art data-driven approaches, extensively applied for analyzing spatio-temporal (ST) data. Various techniques in this regard are broadly classified into conventional and deep learning based techniques, and these are presented in Section 3 and Section 4, respectively.

3 State of the Arts in Conventional Techniques for Data-Driven Analysis with Spatio-Temporal Data

In this section we report the state-of-the-arts in various conventional techniques (including pure statistical and AI/CI-based approaches) which have been widely used to analyze spatio-temporal data. The techniques have been discussed with respect to each data analysis family separately. Further, as per the underlying objective and/or the base approach used, we have hierarchically classified the existing techniques into variants of basic techniques. This hierarchical structure offers an easier and more laconic understanding of the state-of-the-arts in spatio-temporal data mining/analysis.

3.1 Spatio-Temporal Prediction Techniques

In spite of the fact that the spatial relationships are powerful and informative, while predicting the ST data, most of the earlier researches focused only on the temporal aspects without taking into account the spatial dependencies. Various traditional statistical time series prediction models formed the base structure of these techniques. On the contrary, the recent research has focused more on utilizing the rich set of spatial information to improve the prediction accuracy. Therefore, a number of spatially-enhanced prediction techniques have been proposed in recent days. In this subsection, we have discussed both conventional statistical and spatially extended techniques for the prediction of ST data. Moreover, as the significant driver of data-driven approaches, we have also explored the state-of-the-art artificial intelligence (AI) and computational intelligence (CI) techniques. A hierarchical representation of these techniques is depicted in Fig.6.

3.1.1 Traditional Statistical Techniques

Among various conventional statistical techniques, the Exponentially Weighted Moving Average (EWMA) model, the Autoregressive Moving Average (ARMA) model, the Autoregressive Integrated Moving Average (ARIMA) model, and the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model have been extensively used especially for time series prediction of spatio-temporal data.

In earlier days, the EWMA model has been widely used for financial and economic time series prediction, whereas the ARMA and ARIMA models are generally applied in meteorological and other atmospheric prediction work. Further, the recent research shows a tendency of using hybrid ARIMA models along with artificial neural network (ANN) [18, 19]. On the other side, the GARCH model is mostly used for analyzing time series data in financial application [20]. A number of variations for GARCH models have been proposed till date. NGARCH, IGARCH, fGARCH, etc. are a few examples in this regard. However, the major limitations of applying these traditional statistical techniques for ST prediction are: 1) most of these techniques suffer from linear and/or univariate nature and the backward looking problem; 2) none of these take the spatial properties of the data into account.

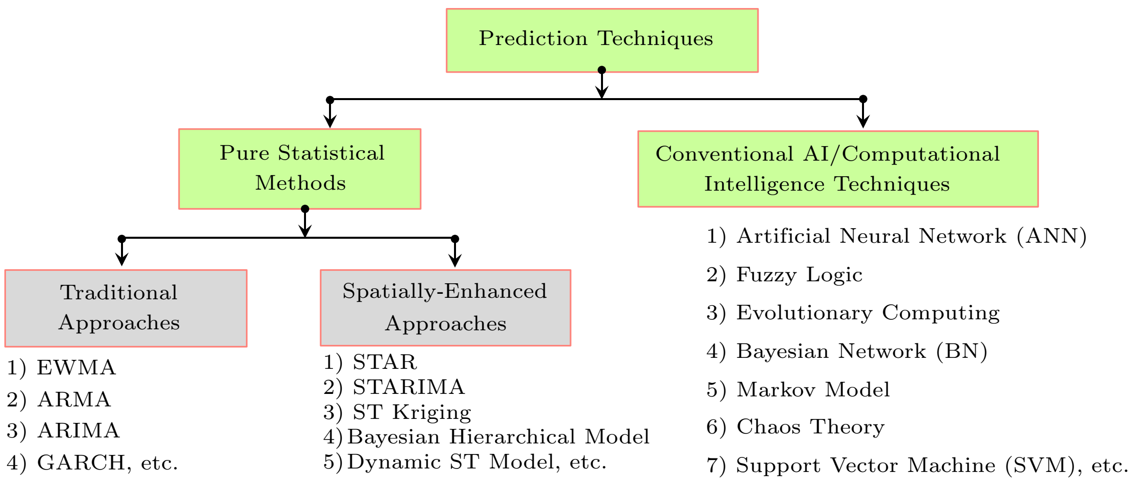

Fig.6. State of the arts in conventional techniques for prediction of ST data.

3.1.2 Spatially-Enhanced Statistical Techniques

In order to overcome the limits of traditional techniques, the spatial statisticians have proposed a number of spatially extended prediction techniques to apply on spatio-temporal data. The Space-Time Autoregressive Moving Average (STARMA) model, the Space-Time ARIMA (STARIMA) model, Spatio-Temporal Kriging (ST Kriging), the Bayesian Hierarchical model, Dynamic Spatio-Temporal Models (DSMs), etc. are most commonly used statistical ST prediction techniques. A summary of these techniques are provided in Table 3.

3.1.3 AI/Computational Intelligence Techniques

The recent advancement in data-driven modeling for spatio-temporal prediction is mainly due to the progress in various computational intelligence (CI) techniques. In this subsection, we have discussed both the conventional and spatially enhanced CI techniques that have been widely used for predicting ST data.

Artificial Neural Network (*ANN*). Several research papers have been put forward employing ANN-based approach for predicting spatio-temporal data from various domains, including climatology, hydrology [25], transport system, biology [26] and so on. Variants of ANN models, like feed forward neural network trained with Levenberg-Marquardt algorithm [25], echo state network or ESN [27] have been employed for these purposes. In [28], Daliakopoulos and Tsanis claimed that, compared with the traditional models, ANN can show superior performance in modeling complex hydrological processes. ANNs have widely been used for traffic prediction also. The work in [29] is an example in this respect. However, most of these studies are developed with less exploration of uncertainty management issues, and it is also necessary to incorporate robustness in those approaches.

Fuzzy Logic. This is one of the popular CI techniques, in which the computation is performed based on the degree of truth, rather than the classical Boolean logic. Recently, fuzzy logic has featured in many successful applications including the prediction of spatio-temporal data. For example, Bazartseren *et al.* [30] proposed a system based on neuro-fuzzy technique which is proved to be effective in short-term forecast of water level. Two different adaptive neuro-fuzzy approaches, namely ANFIS-GP and ANFIS-SC, have been proposed in [31] for estimating the house selling price. More research studies on hybrid fuzzy logic systems are discussed along with the Bayesian network based prediction approaches.

Evolutionary Computing. Evolutionary algorithms are biologically inspired algorithms which are based on the natural principle of survival of the fittest. The genetic programming (GP), genetic algorithm (GA), ant coloney optimization (ACO), particle swarm optimization (PSO), etc. are some variants of the evolutionary computing algorithms. A number of research literatures employing hybrid evolutionary computing to predict spatio-temporal data can be found in the literature. For example, Semero *et al.* [32] developed a GA-based ANN model that can predict wind speed on short-term basis; hybrid SVM-PSO and SVM-GA based approaches have been used for real estate price forecasting in [33] and [34], respectively, and so on.

Bayesian Network (*BN*). The Bayesian network is a powerful tool for representing and reasoning with uncertain knowledge. It has the capability of intuitively representing relevant dependencies and automatically capturing probabilistic information from the data.

Being able to efficiently model complex systems with numerous variables, BNs are extremely suitable for various applications in environmental modeling, especially in meteorology [3, 35]. Recently, Das and

Table 3. Summary of the State-of-the-Arts in Spatially-Enhanced Statistical Techniques for ST Prediction

Technique	Key Feature	Primary Application Area
STARMA & STARIMA	Explicitly models both the temporal and the spatial dependency among variables from various locations; suitable for spatial time series data and moving object data; however, the number of parameters increases geometrically with the increasing number of locations	Transport system, hydrology, remote sensing, climatology/ meteorology [21]
ST Kriging	Generalizes the spatial Kriging process [22] with a ST covariance matrix and variograms; widely used for predictions from incomplete and noisy ST events and spatial time series of raster data; however, the performance decreases with increasing forecast horizon	Climate/environmental science, ecology, and meteorology [23]
DSTM	Focuses on modeling the latent ST processes; assumes that the current values of a process at any location evolve from past values from neighboring locations; based on three-stage factorization of data, process, and parameters; suitable for spatial time series data	Climate/environmental science [24]

Ghosh [36] proposed a semantically enhanced Bayesian network (semBnet) which is able to improve performance by incorporating domain knowledge during the BN-based spatio-temporal prediction of meteorological time series data. Apart from the meteorology, the BNs have widely been applied in hydrological prediction as well [37, 38]. The research work in [39] is another example where spatially enhanced BN or SpaBN has been used to model the spatial influence of the meteorological and topographical factors while predicting reservoir water dynamics. In [40] and [41], the BN-based approaches are found to show considerable improvement when upgraded with added residual correction mechanism during inference generation process under scarcity of influencing factors. However, sometimes due to the lack of appropriate data/information, it becomes difficult to express knowledge in standard BNs. In such cases, a BN with incorporated fuzziness, can be used to resolve the issue [42].

Markov Models. Markov chain, Hidden Markov model, etc. are probabilistic models commonly used for simulating and exploring the process of dynamic systems. These directed graph-theoretic approaches provide an easy mechanism to represent the state transitions and thereby become effective for predictive analytics. The application of Markov models is well-observed in predicting spatial time series and moving object data. For example, Yang *et al.* [43] used spatio-temporal hidden Markov model (STHMM) for predicting travel cost in transportation network, and Yuan *et al.* [44] utilized Markov chain for predicting the traffic condition. Use of Markov models is also popular in urban growth modeling [45]. To be noted, the Markov models are effective in a predictive scenario which exhibits the Markov property.

Chaos Theory. This is another new area for dynamical system analysis. It has shown tremendous growth in analyzing nonlinear dynamics of time series data, obtained from the real observations of natural phenomenon [46]. One of the commonly used chaos analysis tool is the fractal/multi-fractal analysis, which is applied often, for analyzing ST data, especially in the domains of finance [47], biology, and climatology [46].

Support Vector Machines (*SVMs*). SVMs offer a set of supervised learning algorithms that can be used in classification and regression problems. The use of SVMs for the ST prediction is often observed in the field of finance, meteorology and hydrology. For example, hybrid SVM-based approaches, along with PSO and genetic algorithm, are used in [33] and [34], respectively, for real estate price forecasting. The authors in [48] used SVM in combination with discrete wavelet transform to forecast streamflow on monthly basis.

Overall, all the above mentioned conventional CI techniques have their own advantages and disadvantages over the others. None of these can be treated as the best technique with respect to all prediction scenarios. Although the ANN-based approaches are efficient alternatives of the traditional statistical approaches, sometimes these require quite complex models to become good predictors. Further, these are sensitive to the initial weight assignments and hyper-parameters setup. In case of SVMs, finding heuristic to determine the free parameters is also challenging. The approaches based on chaos theory are also highly sensitive to the initial condition. Though the issue of initialization is quite resolved when applying evolutionary computing and probabilistic reasoning, both of these suffer from the curse of dimensionality, and also, the evolutionary computing models may result in premature convergence to the local maxima. A detailed comparative study of all these CI techniques can be found in [3].

3.2 Spatio-Temporal Change Detection and Analysis Techniques

Over the past few decades, there has been extensive research on analyzing change pattern in spatio-temporal data. The existing approaches deal with either the thematic or the geometric change pattern. This subsection discusses on various change detection and change pattern analysis techniques used for spatio-temporal data. A hierarchical representation of various change detection techniques is depicted in Fig.7.

3.2.1 Thematic Change Detection Approach

The thematic change refers to how the thematic information (e.g., vegetation cover, surface temperature), relevant to a spatio-temporal dataset, changes with space and time. Most of the conventional methods focus on the spatio-temporal analysis of such thematic change. The cumulative sum (CUSUM), subpath enumeration and pruning (SEP), statistical spatial wombling, S-outlier detection, ST scan statistics, ST cluster detection, and image processing are some well-used techniques in this regard. The key features of each of these techniques are summarized in Table 4. As mentioned in the table, the image processing techniques are most applicable for the sparial raster time series (e.g., remote sensing imagery), whereas the others can be ap-

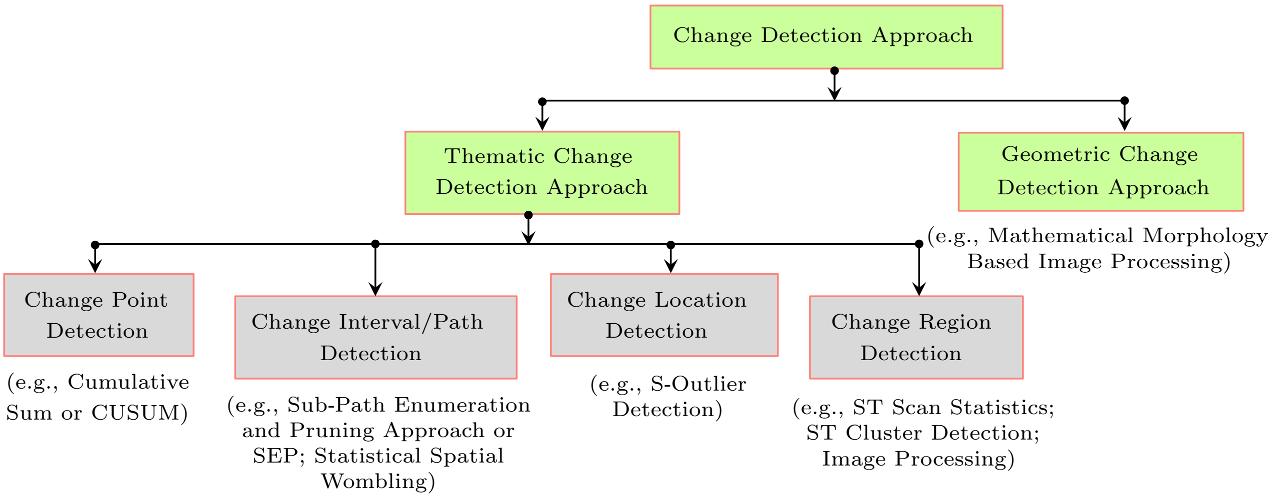

Fig.7. State-of-the-arts in conventional techniques for change detection in ST data.

Table 4. Summary of the State-of-the-Arts in Conventional Techniques for Thematic Change Detection

Technique	Key Feature	Primary Application Area
CUSUM	Mainly used for change point detection purpose; suitable for spatial time series data	Social network, system control [49]
SEP	Finds collections of long interesting sub-paths defined by some interest measure; commonly used for change interval detection; useful to understand abrupt changes; suitable for spatial time series data	Climatology [50]
Statistical spatial wombling	Based on the idea of identifying spatial boundaries separating regions with significantly different observed values of the spatial variable; suitable for ST events and spatial time series data	Ecology, meteorology, public health [51]
S-outlier detection	Used for detecting change locations; suitable for ST events and spatial time series data	Transport system, meteorology/climatology [52]
ST scan statistics	Detects change regions by means of identifying spatial clusters; based on the hypothesis testing and maximum likelihood ratio score; suitable for ST events and spatial time series data	Public health [53], ecology
ST cluster detection	Determines spatial region, having higher risk or intensity of spatial event during a certain period of time; suitable for ST events and spatial time series data	Public health, homeland security, climatology, ecology [54, 55]
Image processing	Mostly used for detecting changes in remote sensing imagery; can be either pixel-based or object-based; focuses either on detailed change trajectories or on detecting binary change; suitable for ST events and spatial time series data	Public health, ecology, climatology, urban development and planning [56, 57]

plied on non-image spatial time series data and spatio-temporal events. The CUSUM technique is simple and easy to implement. However, this is only applicable for detecting the change point. Further, with the decrement of the number of samples, its power of detecting small changes reduces considerably. The SEP technique can detect the interval of changes, rather than simply identify change points. Yet, this method needs predefined interest measures for change and stability, and the model performance substantially depends on the same. The statistical spatial wombling techniques are ideal for identifying spatial zones of temporal changes. However, with the increase in spatially referenced data, this technique requires a more precise measure for quantifying the significance of the change boundaries. The S-outlier based techniques [52] primarily aim at detecting the ST changes by determining the spatial outliers at different time stamps. However, these may not be effective in several cases as these do not take into account the temporal aspects of the data while detecting the S-outliers. ST scan statistics, which typically scan the space while looking for space-time clusters, are quite popular for identifying ST changes [53, 54]. Nevertheless, these are based on several assumptions on data distribution and the shape of cluster, which restrict them for being applied in many scenarios. Incidentally, there exist different other techniques [55] employing ST clustering for ST changes detection, which are free from so many assumptions over data. A more extensive discussion over these techniques can be found in [50].

3.2.2 *Geometric Change Detection Approach*

Analyzing the geometric characteristics of the change in spatio-temporal data is not a much explored area. However, the analysis of geometric change pattern is also essential, especially to provide insights into how the spatial distribution of the ST phenomena changes with time. This needs some special methods, different from the conventional statistical exploratory data analysis techniques. In this regard, variants of set-theoretic approaches can be found in the literature. For example, in [58], the authors proposed a modeling approach based on the basic operators from mathematical morphology to analyze the changes in spatially distributed events/objects or phenomena. Morphological study has also been employed in [59] to categorize the spatio-temporal change in objects.

3.3 Spatio-Temporal Outlier Detection Techniques

The ST outlier detection techniques are broadly classified into two major categories, based on whether these are meant for the trajectory data or for the spatial time series data. Fig.8 represents a compact hierarchical representation for the same.

3.3.1 *Outlier Detection in Spatial Times Series Data*

These techniques are mostly based on neighborhood-based approaches, auto-regressive models, visualization approaches, and shape analysis. The key features of these techniques along with the associated challenges are summarized in Table 5.

The neighborhood based approaches taking care of the contextual information are widely used in literature. However, deriving the spatial neighborhood, defining "considerable deviation", combining the contextual and spatiotemporal distances, etc. become major issues for

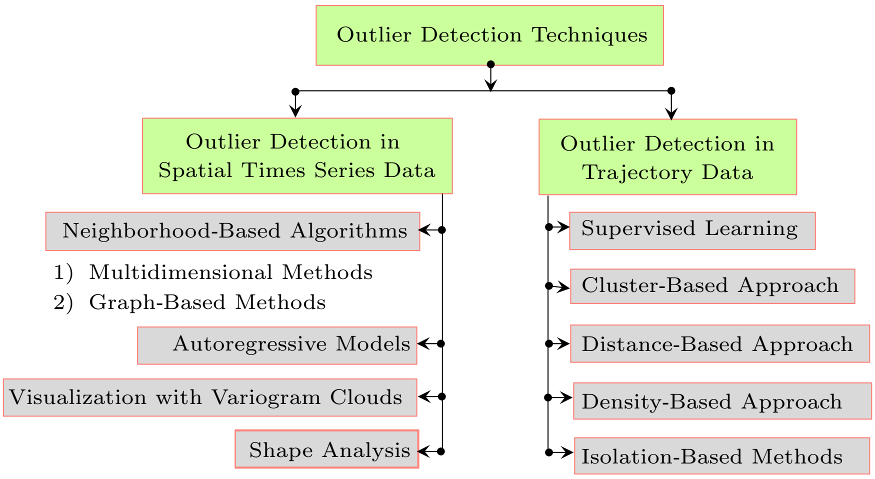

Fig.8. State-of-the-arts in conventional techniques for outlier detection from ST data.

Table 5. State-of-the-Arts in Conventional Techniques for Outlier Detection in Spatial Time Series

Technique	Key Feature	Primary Application Area
Neighborhood-based algorithm	Neighborhoods are defined based on expanded set of contextual attributes along with the spatio-temporal dimensions; however, combining the contextual and spatio-temporal distances is a major challenge	Census, video analysis, network analysis [60]
Autoregressive model	Extension from both temporal and spatial autoregressive models, shows higher degree of robustness than the neighborhood-based models; commonly used when large amounts of reasonably complete data are available; high computational complexity	Not so popular for outlier detection
Visualization	Provides insights into distribution of data; aids in selecting appropriate model for subsequent processing; variogram cloud, pocket plot, Moran scatter plot, etc. are some well-used tools in this regard	Urban planning and development [13, 61]
Shape analysis	Used for identifying ST outliers from images, like MRI scan, PET scan, weather datasets; contours of the shapes are constructed on the basis of the changes in the behavioral attributes between two snapshots	Medical science, meteorology/climatology, traffic [62]

these techniques[13]. Comparatively, the autoregressive models are found to be more robust in this context. Nevertheless, because of a large number of coefficient requirements and high computational complexity, these are rarely used in practice. Apart from the standard statistical outlier detection techniques, sometimes the visualization becomes useful by providing insights into data distribution and subsequently facilitating the outlier analysis with ST data. However, some of the visualization techniques, like variogram cloud, suffer from high computational complexity and can become intractable when the data contains even only hundreds/thousands of spatial data points. The shape analysis techniques, on the other hand, aim at identifying unusual shapes from the distribution of the spatial attributes. Nevertheless, the shape contour detection sometimes becomes an issue for these techniques.

3.3.2 Outlier Detection in Trajectory Data

Detection of spatio-temporal outliers from trajectory data is challenging because of the dynamic and high-dimensional nature of the trajectory data. The existing work in this regard can be roughly classified into five major categories, namely, supervised learning methods, clustering-based methods, distance-based approaches, density-based approaches, and isolation-based methods. Moreover, there exist stochastic model based[63] approaches for ST outlier detection. The key features of these techniques along with the associated challenges are summarized in Table 6.

Among various techniques for outlier detection from moving object data, the distance-based methods are the most widely used ones[64]. However, these generally apply nested loop for every anomaly candidate, and thus are not scalable to high-dimensional datasets. On the other side, the distance-based, density-based, and cluster-based techniques use only time and distance to directly judge whether a trajectory is anomalous or not, and so, these are able to provide insights into only spatially/temporally local distributions of data points. The temporal locality issue is handled by isolation-based techniques[67] that compare the test trajectory against a set of sampled historical trajectories to determine its normal/anomalous nature, whereas the spatial locality issue is resolved by the motif-driven supervised learning approaches[65] which analyze complex relationships between multiple features associated with different spatial granularities.

3.4 Spatio-Temporal Hotspot Detection Techniques

The state-of-the-art ST hotspot detection techniques (refer to Fig.9) can be broadly classified into: 1) clustering-based approach, 2) ST scan statistics based approach, and 3) eigenspace method. The major aspects of each of these techniques are briefly presented in Table 7.

Among various techniques for ST hotspot detection, the most popular and widely used approach is the Space-Time Scan statistics or ST Scan statistics[68]. Typically, the ST scan statistics approach employs a sliding window to search for significant spatiotemporal clusters in entire space. However, these approaches make several assumptions especially regarding the data distribution and the shape of the hotspots which often restrict them to be applied on real-world scenario. The recently proposed eigenspace method[69] overcomes all

Table 6. Summary of State-of-the-Arts in Conventional Techniques for Outlier Detection in Trajectory Data

Technique	Key Feature	Primary Application Area
Supervised learning	Views the movement/change paths as a sequence of object movement features or patterns, called *motif*; based on the motifs, a classifier is learnt to distinguish between normal and the outliers	Transportation, mobility analysis[65], online and social network, climatology/meteorology
Cluster-based	Outlier score is determined based on whether a data point is not within any cluster, its nearness to the other clusters, and the size of the cluster nearest to the data point; not meant for optimizing the outlier detection process	Mobility analysis, transport system[66]
Distance-based	Assumes that the distances between outlier data-points and their k-nearest ST neighbors are notably higher than that in normal case; analysis is performed at detailed granular level; however, the computational complexity becomes extremely high with the increase in data points; requires efficient pruning strategy	Transport system, mobility analysis[64]
Density-based	Uses local density as the outlier score; partitions the data space instead of the data points	Mobility analysis, transport system[13]
Isolation-based	Explicitly isolates anomalies instead of profiling normal points; assumes that the abnormal trajectories are "few in number" and notably "different from the majority"[67]; e.g., "iBAT" model,"iBOAT" model, isolation forest or iForest model	Location-based services, transport system, mobility analysis[67]

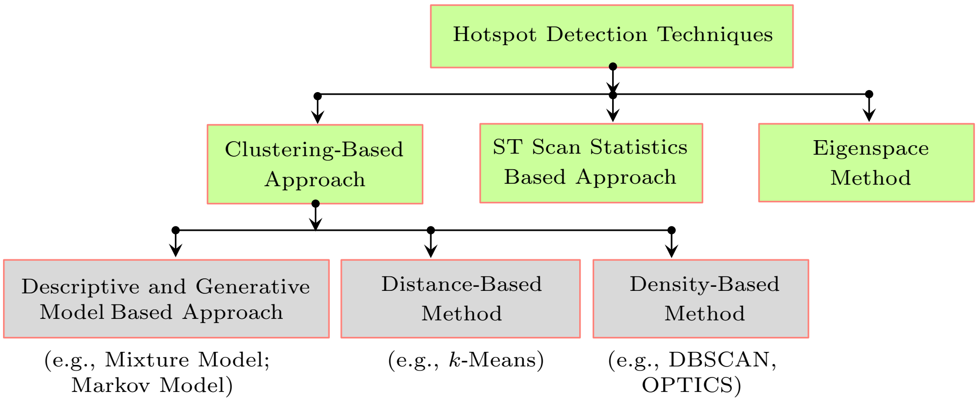

Fig.9. State-of-the-arts in conventional techniques for hotspot detection from ST data.

Table 7. Summary of the State-of-the-Arts in Conventional Techniques for ST Hotspot Detection

Technique	Key Feature	Primary Application Area
Descriptive and generative model based	Clustering based approach; learns global models, capable of describing the entire dataset in terms of some distribution function and a set of fitting parameters; mostly uses SVM and spatio-temporal neural network (STNN); suitable for ST events, spatial time series data, and moving object data	Homeland security, transport system, mobility analysis, urban planning and development, public health [70]
Distance-based	Clustering-based approach; clustering is performed based on some distance-functions that capture the resemblance between data items; suitable for moving object data	Transport system, mobility analysis [71]
Density-based	Clustering-based approach; sets a density-threshold for each object so as to distinguish relevant data-items from noise; suitable for moving object data	Mobility analysis, transport system
ST scan statistics	Based on exhaustive search over the whole space; suitable for ST events and spatial time series; makes restrictive assumptions over the shape of the hotspots, and the distribution and quality of data; sometimes become unrealistic for non-traditional data sources	Public health, epidemiology [72], homeland security [54, 68]
Eigenspace method	Instead of an exhaustive search over the space, the changes are tracked in a space-time occurrences structure; shows better computational efficiency than the standard ST scan statistics; makes no assumption about data distribution, hotspot shape, data quality, etc.; suitable for ST events, spatial time series data	Public health, epidemiology, homeland security [69]

these issues by tracking the changes in space-time correlation structure. There also exist ST hotspot detection approaches on density/distance-based clustering techniques. These are most applicable for moving object/trajectory data. The descriptive and generative models [70] for clustering techniques additionally help to predict the hotspot for future time stamps.

3.5 Spatio-Temporal Partitioning and Summarization Techniques

Spatio-temporal partitioning is closely related to spatio-temporal hotspot detection. However, the key difference is that, in the case of hotspots, the intensity of events/activities within partition is substantially higher than that of outside [8]. Contrarily, the spatio-temporal summarization is performed to obtain an aggregated statistics of the objects within each partition.

3.5.1 *Partitioning Techniques*

Depending on the type of underlying ST data, there exist several variants of partitioning techniques which can be broadly categorized as follows: global partitioning techniques, hierarchical approach, graph based approach, density based approach, and frequency based approach (refer to Fig.10). The summary of the various ST partitioning techniques is provided in Table 8.

The global partitioning and hierarchy-based methods are comparatively simpler than the other techniques. However, these are not very appropriate for partitioning the moving object data. The graph-based methods are more suitable for generating summary and rarely used for generating partitions only. The density-based partitioning methods are found to perform well for ST data portioning. However, because of the issue of high dimensionality, these are not suitable for spatial time series data. On the other side, the frequency-based

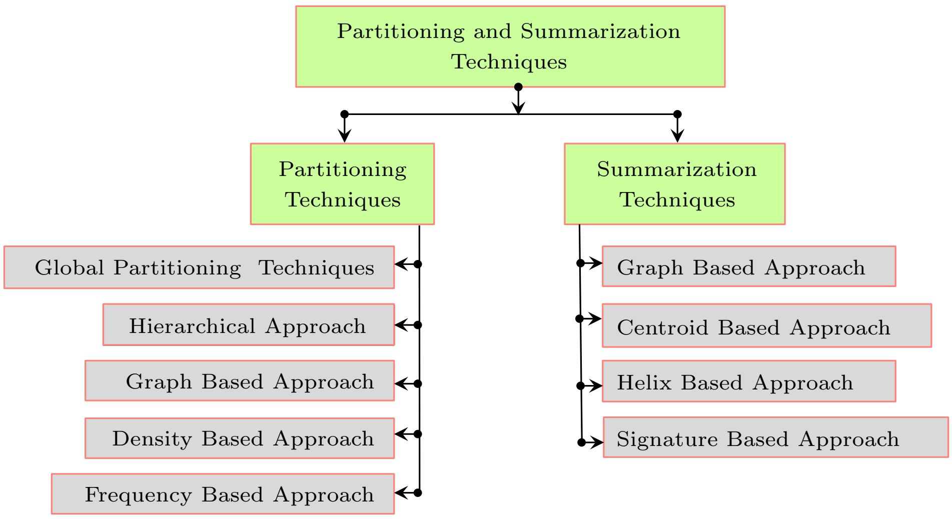

Fig.10. State-of-the-art techniques for partitioning and summarization of ST data.

Table 8. Summary of the State-of-the-Arts in Conventional Techniques for ST Partitioning

Technique	Key Feature	Primary Application Area
Global partitioning	Primary objective is to maximize the within-group similarity of the ST objects; K-means/medoids, EM algorithm, CLIQUE, CLARANS, BIRCH, etc. are used to serve the purpose; mainly used for partitioning ST events and spatial time series data	Public health, homeland security, urban planning and development [50]
Hierarchy-based	Partitions the ST data at different hierarchical levels; mostly used for ST events, spatial time series data, and sometimes for moving object data	Transport system, mobility analysis, public health, homeland security [50]
Graph-based	Represents the ST data-items in terms of sparse K-nearest neighbor graph, subsequently partitions them into segments, and finally, merges in hierarchical fashion; primarily used for ST events	Not so popular
Density-based	In the case of ST events, first identifies dense points and connects them to generate contiguous groups/clusters or partitions; in case of trajectory data, first performs trajectory segmentation and then applies density-based clustering techniques; suitable for ST events, and moving object data	Transport system, mobility analysis [73], public health, homeland security
Frequency-based	Identifies subsections of trajectories which have high frequencies; mostly based on association rule mining techniques; suitable for moving object data	Mobility analysis, transport system [74]

models are only applicable for moving object data. Incidentally, defining the matching prototype (or association) becomes a major issue in this case.

Apart from the techniques to directly partition raw ST data, recent research efforts are also found to learn cluster pattern hidden in ST data. The Autoencoder Regularized Network (ARNet) [75] proposed by Dong *et al.* and multi-task learning model T2INet [76] proposed by Kieu *et al.* are worth mentioning in this context. These models have unique property of partitioning even from incompletely labeled moving object data.

3.5.2 *Summarization Techniques*

Spatio-temporal (ST) summarization is often performed together with ST partitioning so that the objects within each partition/group can be summarized by some "aggregated statistics" [50]. The primary objective of ST summarization is to find a compact as well as informative description of an ST dataset. Considering all the different categories of underlying ST data (e.g., spatial time series data, moving object/trajectory data), the summarization techniques can be classified into: distance-based, graph-based, helix-based, and signature-based techniques (refer to Fig.10). Key properties of these techniques are presented in Table 9.

Though the graph-based techniques [77] are most widely used for ST summarization, these suffer from high computational cost for creating track similarity index which requires computing the distance of each node to the nearest node in the other tracks. This limitation of graph-based model is resolved by centroid methods [78] since these use matrix-view, instead

Table 9. Summary of the State-of-the-Arts in Conventional Techniques for ST Summarization

Technique	Key Feature	Primary Application Area
Graph-based	Suitable for moving object data; mainly focuses on determining "hot" trajectories via frequent subgraphs, hierarchical clustering, density clustering, etc.; needs to measure similarity between tracks	Mobility analysis, transport system [77]
Centroid-based	Suitable for moving object data; aims to generate summary of a group of moving objects in terms of a set of representative centroid objects; uses matrix-based view instead of a graph-based view	Mobility analysis, transport system [78, 80]
Helix-based	The spine of the helix describes the ST trajectory of an object's center, while the prongs describe the deformation of the object's outline at specific time instances; suitable for deformable objects, like spatial raster datasets	Urban planning and development, public health, video analysis, mobility analysis
Signature-based	Produces a family of data summaries or "signatures" that can effectively represent the actual data with much smaller storage footprint, while allowing for efficient querying; used for spatial time series, and moving object data	Transport system, mobility analysis [79]

of graph-view. The helix-based models are suitable for spatial raster time series, capturing a variety of phenomena at discrete temporal instances. These can also be used for developing complex and efficient knowledge bases. However, helix-based models suffer from complex calculation of contour modeling. The signature-based models [79] represent the ST summary in terms of a unique characteristic/pattern of the data. Eventually, these reduce the storage requirements and also support efficient query processing.

3.6 Spatio-Temporal Coupling and Tele-Coupling Techniques

3.6.1 *Coupling Techniques*

As stated earlier, the spatio-temporal coupling refers to the pattern of occurrence of two or more spatio-temporal object/event types in close spatial and temporal proximity. The common methods in this regard can be categorized as per the ordering pattern of the objects [8]. Some of these include mixed drove spatio-temporal co-occurrence pattern mining approaches, spatio-temporal sequential pattern mining approaches, and cascading spatio-temporal pattern mining approaches (refer to Fig.11). The key aspects of each of the techniques are briefly described in Table 10.

The co-occurrence pattern mining techniques are useful for identifying events or a subset of events that occur together in both space and time (e.g., solar events). However, defining the interest measure to find prevailing ST co-occurrence patterns often becomes computationally expensive when the spatio-temporal data is massive. The issue can be handled by employing filter and refine algorithms [81]. The cascading pattern mining techniques are applicable for events that occur serially but are found to be located together (e.g., drunk driving incidence after and nearby to a bar closing). However, these suffer from expensive ST neighborhood computation for evaluating interest measure, which needs further research. The sequential pattern mining techniques are used for mining totally ordered patterns (e.g., sequence of visiting places in a region of

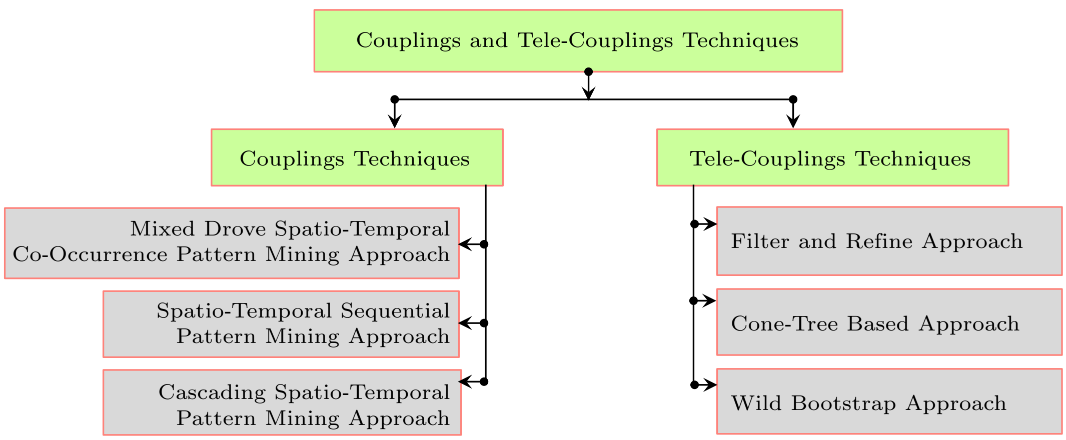

Fig.11. State-of-the-arts in conventional techniques for analyzing coupling/tele-coupling in ST data.

Table 10. Summary of the State-of-the-Arts in Conventional Techniques for ST Coupling

Techniques	Key Feature	Primary Application Area
Co-occurrence pattern mining	Used for unordered patterns and often useful for applications like tactic identification in games/battlefields, network planning, predator-prey interactions tracking, etc; primarily based on the concept of representing the events in terms of feature summary and grouping them based on some complex interest measure; applicable for ST events, ST time series, and also for moving object data	Environmental study, climatology [50, 81]
Cascading pattern mining	Meant for partially ordered subsets of events; mostly based on directed acyclic graph (DAG) based approaches; suffers from exponential cardinality of candidate patterns; suitable for ST events and spatial time series data	Understanding effect of climate change, global warming, spread of multiple infectious disease, etc. [82], homeland security
Sequential pattern mining	Used for mining totally ordered patterns; especially applied on moving object and transaction data; mostly based on the PrefixSpan algorithm or extension of it; needs to define an appropriate significance measure	Mobility analysis, transport system, transaction system [50, 83]

interest) from the ST data. However, defining meaningful significance measures and designing algorithm under these measures become serious challenges. "K-function statistics" can be used to overcome this issue.

3.6.2 Tele-Coupling Techniques

Discovering tele-coupling pattern is important especially in understanding global environmental change and climate oscillations [84, 85]. In a broader sense, spatio-temporal tele-coupling refers to "high correlation across spatial time series at a long distance" [8]. Various approaches in these regard are mostly based on ST auto-correlation analysis and can be broadly categorized into filter and refine based, cone-tree based, and wild bootstrap approaches (refer to Fig.11). The major challenges in tele-coupling pattern mining arise due to time series length, numerous candidate pairs, and spurious "high correlation" location-pairs. The filter and refine based and cone-tree based approaches are used to deal with the first two challenges by filtering out redundant pair-wise correlation computation and using efficient index structure, respectively. On the other side, the wild bootstrap approaches are used to address the challenge regarding spurious pairs of locations [11].

4 State-of-the-Art Deep Learning Techniques for Analyzing Spatio-Temporal Data

In addition to the conventional statistical and CI/AI-based techniques, recently deep-learning has gained increasing research interest in the field of spatio-temporal data analysis. Deep learning offers a set of machine learning algorithms that attempt to "learn in multiple levels, corresponding to different levels of abstraction or concepts" [86], and thereby can appropriately utilize the huge set of available ST data. However, to the best of our knowledge, none of the existing surveys has discussed on these techniques from the perspective of all the ST data analysis families. We, therefore, intend to present the state-of-the-art deep learning approaches for ST data analysis in this separate section.

4.1 Deep Learning Approaches for Spatio-Temporal Prediction

Most of the deep learning approaches for spatio-temporal prediction are based on convolutional neural networks (CNNs), deep neural networks (DNNs), recurrent neural networks (RNNs), and stacked auto-encoders (SAEs). A summary of all the articles employing these models is presented in Table 11.

The DNN models applied in the domain of ST prediction generally appear with additional functionalities like feature-level data fusion [91], convolution [90], etc. which are mainly utilized to learn the spatial and temporal dependencies. However, because of using a large number of parameters, these often suffer from large memory consumption and over fitting problems. Similar problems are also encountered in SAE models though these have advantage of representing features in a more compact way even from noisy inputs [94]. Contrarily, with the concept of weight sharing and pooling, the CNNs reduce the parameter as well as computational time requirements. In addition, these are translation-invariant and are intrinsically able to learn local dependencies using convolution. The recent research also shows that CNNs with added residual unit (e.g., ST-ResNet [89]) are also able to effectively capture large-scale ST dependencies through deep convolutional network architecture. Eventually, these models become suitable for ST prediction (classification) of imagery or video data [87]. Nevertheless, the vanilla CNN

Table 11. Summary of Articles Employing Deep Learning for ST Prediction

Technique	Task	Dataset
CNN-based multi-granular deep architecture (1D/2D/3D CNN, ST-ResNet)	Prediction of actions in videos [87]	UCF-101, Sports-1M
	Prediction of traffic speed [88]	Own (Beijing transportation network)
	Prediction of crowd flow [89]	TaxiBJ, BikeNYC
DNN	Prediction of traffic/crowd flow [90]	TaxiBJ15, TaxiGY16, LoopGY16, BikeNYC14
	Prediction of crime occurrence [91]	Own (from Chicago, Illinois)
RNN (LSTM, structured-RNN)	Travel speed prediction [92]	Own (from Beijing and Shanghai)
	Driver maneuver anticipation [93]	CAD-120
SAE	Traffic flow prediction [94]	Data from Caltrans Performance Measurement System (PeMS)
	Prediction of air quality [95]	Air quality data from the Ministry of Environmental Protection of China
DSN and its variants (Deep-STEP, Deep-STEP_FE)	Prediction of vegetation index [96,97]	Landsat TM-5 satellite imagery (United States Geological Survey)

models consider only the current input and cannot handle the sequential data, which certainly restricts their effectiveness during regression. The RNN-based models, on the other hand, can utilize their internal memory for remembering the sequence and can sometime perform better than CNNs, especially in case of ST regression. In addition, the structured RNNs can learn high-level spatio-temporal structures [93]. Though the standard RNN models are prone to become the victim of vanishing/exploding gradient problem which confines them learning long-term dependencies, the issue is resolved when RNN-LSTM (long-short term memory) is applied [92,98]. Another new deep learning model used in ST prediction is the deep stacking networks (DSNs). Though these models require a predefined neighborhood coverage to learn the spatial dependencies, the recent studies show that the spatio-temporal extensions of DSNs [96,97] can perform better than DNNs while predicting spatial raster time series.

4.2 Deep Learning Approaches for Spatio-Temporal Change Pattern Analysis

Most of the deep learning approaches for ST change pattern analysis have been developed for analyzing satellite remote sensing imagery to understand land-cover change [99–104] and for analyzing video data to recognize human action and various other events/objects [105,106], evolving over time and space. The deep learning approaches employed in analyzing remote sensing imagery are mostly based on convolutional neural networks (CNNs), stacked auto-encoders (SAEs), stacked denoising auto-encoder (SDAE), recurrent neural network (RNN) with long short-term memory (LSTM) architecture, ID-LSTM (Incremental Dual LSTM), and deep belief network with cellular automata based architecture (DBN-CA). In the other case, the majority of the deep learning techniques applied for analyzing the video data are based on convolutional neural networks (CNNs), very deep convolutional network (VGGNets), two-stream convolutional network, etc. More on deep network based video recognition can be found in the work of [87]. A summary of various articles using these techniques is provided in Table 12.

In the case of ST change pattern analysis, an appropriate modeling and simulation of the change plays a critical role, and accordingly, the majority of the deep learning based change pattern analysis techniques are found to have hybrid features inherited from two or more base techniques. For example, the DBN-CA [110] has excellent feature detection capability, achieved by means of unsupervised learning of DBN. Additionally, it also offers the best transition rules as captured using the cellular automata (CA). ID-LSTM [109] uses a combination of two LSTM architectures in order to keep account of long-term and short-term variations separately. This becomes extremely useful for dealing with newly evolving class/patterns. However, in this case, the complexity of the model increases since it needs to deal with a large number of parameters. On the other side, in order to overcome the limitation of vanilla CNN models in dealing with sequential changes in ST time series (e.g., video streams), the 3D CNNs [108], VGGNets [107], and two-stream ConvNets [106] are proposed. Though these models suffer from the drawback of increasing complexity and large computational time, they show promising performance to recognize change pattern from multi-frame optical flow of a video stream.

Table 12. Summary of Articles Employing Deep Learning for ST Change Pattern Analysis

Technique	Task	Dataset
CNN models	Land-cover change pattern learning [99]	Landsat-7 ETM+ satellite image
(Vanilla CNN [99],	Large-scale video classification [105]	CSports-1M, UCF101
ConvNet [105], 3D CNN [108],	Land-cover mapping from satellite images [100]	3A multispectral RapidEye imagery
Two-stream ConvNets [106],	Video classification, action recognition [106]	HMDB51, THUMOS14, ActivityNet
VGGNets [107])	Complex video action recognition [107]	UCF101, HMDB51
	Motion analysis in video data [108]	UCF101, HMDB51, ACT
SAE/SDAE	Change pattern analysis from large-scale remote sensing data [101]	Time series of Landsat imagery and MODIS imagery
	Multi-spatial-resolution change detection from remote sensing images [102]	Ottawa, Stone-Gate, Sardinia, and Yellow-River dataset
	Urban area classification [103]	UAVSAR L-Band dataset, ALOS-2 L-Band dataset
LSTM models (RNN-LSTM, ID-LSTM)	Binary and multi-class land-cover change detection [104]	Landsat 7 ETM+ satellite imagery and EO-1 Hyperion dataset
	Land cover classification, prediction [109]	MODIS dataset (NASA satellite)
DBN-CA	Urban growth pattern simulation [110]	Landsat 7 ETM+, Landsat TM, and Landsat 8 satellite imagery

4.3 Deep Learning Approaches for Rest of the Spatio-Temporal Data Analysis Families

Application of deep learning in the other families of ST data analysis (ST outlier and hotspot detection, ST partitioning and summarizaion, ST coupling/tele-coupling, etc.) is not very common. Only a very few research studies can be found in literature. The majority of the existing deep learning models for spatio-temporal outlier detection are based on autoencoder or its hybrid spatio-temporal extensions. For example, Xu *et al.* [111] used stacked autoencoder network to automatically learn both appearance and motion representations of scene activities for video anomaly detection. Zhao *et al.* [112] proposed a Spatio-Temporal Auto-Encoder model (STAE) which is extremely useful for detecting anomalies from real-world and complex video scenes with cluttered backgrounds. Kieu *et al.* [113] proposed a 2D convolutional autoencoder and an LSTM autoencoder for detecting outliers from multidimensional time series, like driving data. Zhu and Newsam [114] developed a deep convolutional network based classifier that detects sentiment hotspots and predicts the emotion conveyed by geotagged images. However, deep learning for ST data analysis is still at primitive stage and needs to be explored further. The statistics of various deep learning approaches used for analyzing ST data is depicted in Fig.12.

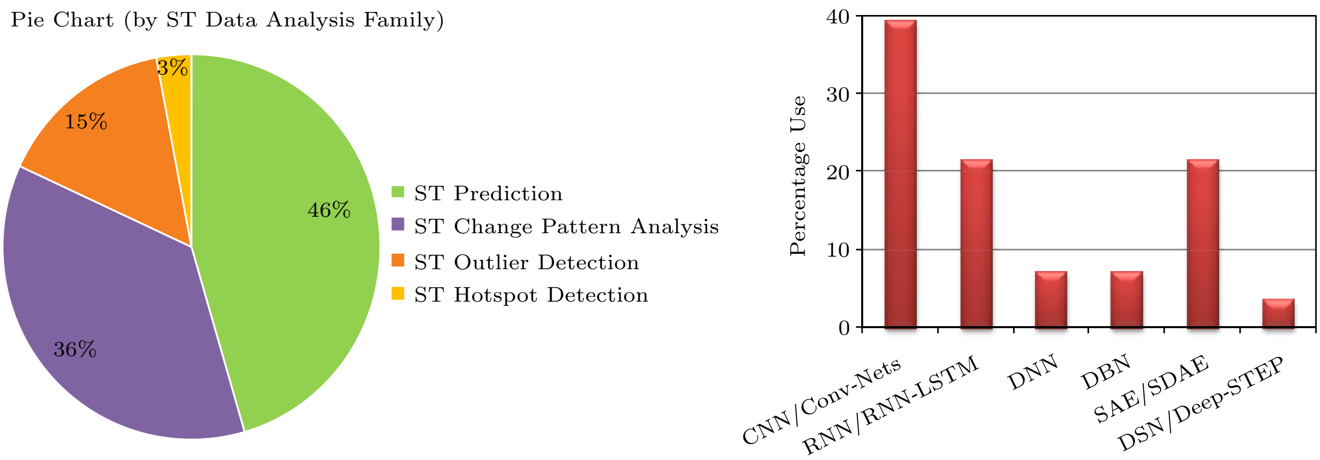

Fig.12. Statistics of the various deep learning approaches used for analyzing ST data.

5 Spatio-Temporal Data Analysis in Various Application Areas

Analyzing spatio-temporal data is crucial for many of the application domains that frequently need to take decisions based on large spatial and spatio-temporal datasets. This section provides an overview of some of such domain specific problems and the applications of spatio-temporal data analysis techniques to resolve the respective issues. The overall discussion may help in stimulating research initiatives and exploring prospective avenues in spatio-temporal data analysis which are still unexplored. Thirteen different application domains have been considered in this regard. For each application domain we discuss the following three aspects:

- nature/characteristics of the data;
- challenging issues in analyzing the data;
- existing data mining techniques.

1) *Climatology/Meteorology*. The climatological/meteorological data, either collected from the sensor networks or recorded by the spatially distributed meteorological stations or derived from the satellite remote sensing imagery, are by nature spatio-temporal data, more specifically, spatial time series data. The majority of the researches in this domain focus on the spatio-temporal prediction of climatological/meteorological time series and also on the detection and analysis of the pattern in climate change. Regarding spatio-temporal prediction of climatological/meteorological data, the approaches proposed in [19, 32, 36, 46, 96] are some significant and recent work. Most of these approaches are CI-based or spatio-temporally extended statistical approaches and are found to perform better than traditional time series prediction techniques. A comparative study of a few of these techniques is provided in [3].

On the other side, the recently proposed ST change pattern analyses on climatological/meteorological data mostly use ST scan statistics [52], image processing [115, 116], or geometric analysis based on morphological operators [58]. Besides, the recent research trend can also be found in identifying climate change hotspot, outlier, coupling, etc., by analyzing the climatological/meteorological data. For example, Wu *et al.* [52] proposed a scan statistics based ST outlier detection approach which can discover an outlier pattern of a weather phenomenon like El Nino Oscilliation; spatio-temporal analysis techniques have been utilized in [117] for modeling and analyzing urban heat island; Huang *et al.* [118] used ST analysis to detect coupling in climate data, in the form of sequential pattern. Further, the recent researches show that the key challenge in analyzing climatological data arises mainly due to inherent chaotic nature and uncertainty present in the data, which may be well tackled by incorporating appropriate domain knowledge [36] and integrating scientific theory [16].

2) *Hydrology*. Analyzing spatio-temporal variability of various hydrological processes, and modeling hydrological responses to the natural and anthropogenic activities at different spatio-temporal scales have gained increasing research interest in recent days. Apart from the existing hydrological models (e.g., SWAT), there is an emerging tendency to employ various data-driven approaches for these purposes [41]. The majority of these approaches are based on pure ANN or its combination with other intelligent techniques [119], like echo state network (ESN) [27], adaptive network-based fuzzy inference system (ANFIS) [120], and Bayesian neural network (BNN) [121]. The comparative study in [28] demonstrates that, compared with the traditional conceptual models, ANN can offer superior performance in modeling complex hydrological processes. However, the effectiveness of an ANN-based model is highly influenced by "proper understanding of the inter-variable dependency" and the "extent of knowledge regarding functionality of neural network" [39]. Therefore, recently, authors in [39] took an attempt to use probabilistic graph-based approach (based on spatial Bayesian network or SpaBN) to model reservoir water dynamics. More Bayesian and ANN-based approaches used in hydrology can be found in [122] and [123], respectively. However, since the hydrological processes are intrinsically dynamic and extremely complex, combining both physical and data-driven approaches for modeling hydrological process is expected to be more effective.

3) *Environment and Ecology*. With the aim of protecting and restoring the natural environment, the application of spatio-temporal analysis in environmental and ecological management has gained rising popularity in recent days. Some of the key objectives in these studies remain in finding the spatial change patterns of ecosystems [124, 125], analyzing the predatory impact and the mutualisms between various organisms, monitoring the dynamics (such as shrinking and expansion) of certain land cover types like forest/desert, studying the impact of climate change as well as human activities on eco-system [126], etc. Apart from all these studies on change pattern analysis, research is also going on to study the coupling and tele-coupling be-

tween globalization/urbanization and the state of eco-environment [85,127]. However, the advancement of ST analysis in this domain is significantly affected by different technical challenges due to heterogeneous and distributed nature of the data and also because of various sociological challenges like insufficient rewards for data collection or sharing.

4) *Medical Science and Public Health*. The applications of ST data analysis in the domain of medical science can be broadly classified into: applications on clinical medicine which deal with the health issues from individual perspective, and applications on community medicine or public health which deal with the health issues from the perspective of populations. The majority of the research studies on individual medical data intend to study the cross-sectional medical imaging (MRI scan, PET scan) or other kinds of spatial time series (e.g., ECG) for detecting chronic diseases such as Alzheimer's disease, multiple sclerosis, etc., diagnosing abnormal health conditions like arrhythmia, and monitoring abnormalities like growth of brain tumors, proliferation of cancer cells, etc. in the human body [128,129]. All these studies mainly fall under the category of ST change pattern analysis and ST outlier detection. On the other side, the mission of public health data analysis [130,131] is mainly to monitor the epidemic disease outbreak for identifying the regions with a high risk of infection and thereby to help in taking adequate measures accordingly [132,133]. These can also be viewed as ST pattern analysis and outlier discovery problems, and have gained growing research interest in present days. However, the heterogeneous unstructured nature of the data and various privacy/security concerns impose critical challenge in flourishing ST data analysis in this domain.

5) *Transport System*. Of late, the spatio-temporal analysis has gained increasing research interest in the field of transport system as well. The objective is to analyze the avalanche of data collected from global positioning system (GPS)-based receivers, cameras, inductive-loop detectors, microwave detectors, etc., and to generate useful insights for enhancing transportation system performance, strengthening travel security, and offering more options to travelers. Based on the ultimate objective, various applications in this regard can be classified into four broad categories, namely, travel management (e.g., travel time prediction, motion tracking, short-term traffic forecasting [134]), congestion control (e.g., congestion detection, traffic flow prediction, congestion propagation pattern analysis, abnormal event sensing [94,135]), route planning (e.g., personalized route planning [136,137], context-aware routing [138], path finding [139], path selection [140]), and accident management (e.g., bus route modification, bicycle corridor selection, analyzing driving behavior [141,142]). Heterogeneous traffic patterns at different road segments, data sparseness and distribution skewness with respect to large road network, causal influence from external factors, etc., are some key issues imposing significant challenges on ST analysis for transport system data.

6) *Urban Planning and Development*. The majority of the researches on spatio-temporal analysis in this domain focus on two key aspects, namely urban growth monitoring and public welfare. The existing studies on urban sprawl or growth monitoring are mostly based on ST analysis of satellite remote sensing imagery and the population data. In this regard, the studies in [99,143] are worth mentioning. Some of these studies also discuss on ST coupling between urbanization and state of the eco-environment [85]. On the other hand, the research studies on public welfare mainly use vector data, like GPS traces, thematic maps, as collected from mobile devices, sensor networks, or the respective monitoring stations/organizations, to generate new insights and aid in improving the quality of city life. Traffic management and transportation planning [144], power supply and energy management [145], water supply network monitoring, education and health management [146], etc. are some important applications in this respect. In all the above cases, the huge volume of the available data and the lack of proper validation mechanism impose substantial challenge during ST analysis.

7) *Finance and Economy*. The application of ST data analysis in the domain of finance and economy is mostly visible in analyzing real estate or housing price. In general, the financial data, like stock market price, share index, stock exchange index, etc. are considered to be purely temporal, and accordingly, a majority of the existing researches in this regard fall under the category of time series analysis without considering any spatial aspects. However, the real estate price or housing price is a category of financial data which is significantly affected by the recent selling prices of the nearby real estate/houses, and therefore, prominently shows spatio-temporal dependencies among such prices [147]. The studies in [148] and [149] are worth mentioning in this regard. These mainly belong to the category of ST change pattern analysis and ST prediction considering statistical regression models. How-

ever, since the individual real estate/housing sales occur at irregular time intervals, modeling such processes considering "standard discrete time series" becomes extremely difficult [147]. Similar to the real estate/housing price data, the economical data, such as the average annual income, extensively depends on the spatial aspects.

8) *Bio-Informatics and Molecular Biology*. Spatio-temporal analysis in bio-informatics and molecular biology is a highly promising but still an under-explored research area because of various challenging issues like rapidly growing volume and diversity of biomedical data, heterogeneous and ill-defined data structure, lack of appropriate tools/techniques for access and visualization of complex biological information, and so on. The majority of the researches in these domains are on analyzing the spatio-temporal dynamics at the cellular and molecular level [150, 151]. Besides, a few studies are also involved in learning the structural features of the molecules and predicting the same at certain context [152]. Moreover, recently deep learning approaches have gained popularity in these domains as well [153]. A compact discussion on the application of deep learning techniques in bio-informatics and molecular biology can be found in [154] and [155].

9) *Location-Based Services* (*LBSs*). LBSs are software-level services that utilize real-time location data, mostly from the mobile devices or smartphones, to provide a service or information that is relevant to the user at that location. LBS can be used in a variety of contexts, including transportation, health, work, entertainment, personal life, and so on. Typically, the data used by LBS can be broadly categorized into a) GPS data that is obtained through the global positioning functionality and internet technology embedded in the mobile device of the user, and b) participatory sensing data that is collected and shared by a group of active participants, through their personal mobile devices and web services [156]. Some typical examples of LBS include local business search and marketing (e.g., optimal location search for retail store placement [157]), point of interest searching (e.g., searching for ATMs, restaurants, cafes, within a user-specified range of distance [158]), e-marketing, social networking (e.g., sharing geo-tagged photos/messages among a group of people [159, 160]), automotive traffic monitoring (e.g., vehicle tracking, inferring traffic congestion [161]), route finding (e.g., shortest path finding, most visited route finding [162]), emergency management (e.g., accident management, disaster management, and health management [163]). The majority of these studies are involved in analyzing the location-related data or user trajectories, to get a better understanding of the spatial/spatio-temporal patterns/relationships.

However, one of the significant challenges faced by location-based services is the issue of privacy preservation which arises because of publishing the personal location information/trajectories to a third party or to the public for data analysis purpose [164]. The other challenges are imposed by the huge volume of data log, heterogeneous data from multiple data sources, issue of real-time data processing, etc.

10) *Mobility Analysis*. With the rapid dissemination of location-aware data generated by various technological infrastructures such as GPS positioning and wireless networks, the mobility analysis has been notably promoted in recent days. Huge volume of such spatio-temporal data, especially the tracking records of human activities, are available and these offer potential opportunities to assess the lifestyle, habits, and demands of citizens, in terms of mobility. Over the last few years, researchers and knowledge extraction communities have devised a number of techniques and models for analyzing movement patterns from raw GPS traces. Further, recently, a new promising area of research has been emerged to provide "applications with richer and more meaningful knowledge about movement", which is called semantic trajectory analysis [71]. This is accomplished through amalgamation of raw mobility data with relevant contextual information. In any case, the ultimate objective is to understand the individuals mobility behavior and aid in various location-based services. Diverse range of application of spatio-temporal analysis with movement data can be found in the literature. Next location prediction, passenger finding/taxi finding [165], user profiling/categorization [166], driving behavior analysis [167], crowd behavior/movement pattern analysis, transportation mode detection [168], real time detection of anomalous trajectory [67], etc. are some crucial examples in this regard. More on trajectory data mining and semantic trajectory data mining can be found in the work of [169] and [71] respectively.

However, various issues, like real-time processing of avalanche of data, efficient storage and retrieval of trajectory traces, scarcity of labeled and clean GPS data, privacy concern due to rapid GPS data sharing, impose critical challenge in extracting behavioral patterns and implicit information from mobility data.

11) *Online and Social Network*. With the increasing accessibility of mobile devices equipped with Internet connections, GPS sensors, and many other advanced

technologies, the number of users actively participating in creation, assembling, and dissemination of local knowledge or spatial/geographic information is growing rapidly. Various social network sites, like Twitter, Facebook, Flickr, and their evolution into location-based social networks play a significant role in attracting millions of users in this context. Now, the spatial information that can be harvested from social media feeds does not fall under the category of volunteered geographic information [170], since it remains embedded in the content of these feeds and not consciously volunteered by the users. The geotagged photo is a typical example in this regard. The key challenge in analyzing such spatio-temporal data arises because of real-time flow, constantly increasing volume, and heterogeneity of the data. The social media feeds are comprised of diverse categories of data (image/photo, text, video, etc.) from different platforms and the data are by nature unstructured and ill-defined, which impose significant challenge in data integration and extracting meaningful patterns/semantics out of the data. The majority of the research studies on social media data are on event detection [171] and abnormality investigation to aid in disaster and emergency management [172], traffic management, and disease/health management. Attempts have also been made to forecast ST event by analyzing social media data [173]. Recent research trend also shows an increasing interest in analyzing user behavioral pattern [174] and sentiment from the social media feeds [114]. A more systematic review of ST analysis on social media data can be found in [175].

12) *Homeland Security*. The application of spatio-temporal analysis for homeland security is mostly found in monitoring and controlling crime occurrence and terrorist attack in any region. With the increasing ability of collecting and storing detailed data tracking crime occurrence, a considerable amount of ST data is available with several countries and organizations. Analyzing the huge volume of crime data can help to better understand the patterns in criminal activities and to further predict crime hotspots in future, so that the police department can take adequate measures, like revising patrol strategies, improving street lighting, and investing surveillance cameras with night vision capability. The majority of the research studies on crime data are found in the form of crime pattern analysis [176], crime hotspot detection [70], and eventual prediction of the same for future [91, 177]. One of such interesting patterns as identified by the researchers is high concentration of vehicle crimes at night in residential neighborhoods, and during the mid-day time in nonresidential areas. Some studies also found an increase in domestic violence during the summer months and an increase in commercial robberies during the winter. However, the research challenges still arise due to lack of required data for conducting comprehensive analysis.

The overall statistics from the perspective of each considered domain has been presented in our supplementary document①. Further, a summary of popular publicly available ST datasets over diverse domains, along with the indication of suitable ST analysis family, is presented in Table 13.

6 Recent Trend and Future Scope of Work

After studying the data-driven approaches for spatio-temporal analysis, we can observe that ST analysis is still a widely open domain for research, and there are several relevant issues that require further exploration. In this section, we summarize a few of these research directions, including visual analytics, hierarchical modeling, data sparseness handling, deep learning, participatory sensing, theory-guided data science, etc., which have gained increasing research interest in recent days and have ample scopes to be explored in future.

6.1 Visual Analytics

Visual analytics can be defined as an approach that combines visualization with human perception and data analysis. The recent advent of visual analytics is mainly fostered by the need of appropriate tools to handle massive amounts of spatial/spatio-temporal data. Visualization gives a basic view of these huge volumes of data, from which the human gain further insights by applying the power of intuition. Then these human-generated insights are transformed into knowledge which helps to further carry out exploratory data analysis using available techniques. Thus, visual analytics has recently become a promising approach with a wide applicability in planning and decision making at various domains. Flow map, spatio-temporal graph [178], clustering [179], generalization, aggregation, etc. are some currently used approaches for visual analytics on moving object data. A huge scope remains in exploring visual analytics under various challenging issues including scala-

①http://cse.iitkgp.ac.in/~monidipa.das/Supplementary/JCST.pdf, Nov. 2019.

Table 13. Publicly Available Datasets for Spatio-Temporal (ST) Analysis

Data Category	Data Along with Source	ST Prediction	ST Change Pattern Analysis	ST Outlier Detection	ST Hotspot Detection	ST Partitioning and Summarization	ST Coupling and Tele-Coupling
ST event data	Crime②③	√	√		√		
	Disease②–④	√	√		√		
	Disaster②③	√	√		√		
	Climate②③⑤–⑦	√	√		√		
	Environment and ecology②③	√	√		√		
Spatial time series data	Climate data②③⑤–⑦	√	√	√	√		√
	Hydrology data②③	√	√	√			
	Environment and ecology②③	√	√	√	√		
	Transportation②③⑧	√	√	√		√	
	Geospatial image series⑤⑨	√	√	√	√		√
	Medical Sc. & public health③⑩	√	√	√			
	Urban planning③⑨	√	√			√	√
	Biology④	√	√	√			
	Finance and economy③	√	√				
	Location-based services⑪	√	√	√	√	√	
	Social network⑫	√	√	√	√	√	
	Crime③	√	√		√		
Moving object/ trajectory data	Cabspotting (USA)⑪	√	√	√	√	√	
	Geolife (China)⑪⑬	√	√	√	√	√	
	MDC (Switzerland)⑪	√	√	√	√	√	
	T-drive (China)⑪	√	√	√	√	√	
	Brightkite (Global)⑪	√	√	√	√	√	
	Gowalla (Global)⑪	√	√	√	√	√	

bility with data volume, data dimensionality, representation, integration of heterogeneous data, and so on.

6.2 Hierarchical Modeling

One of the significant challenges in analyzing ST data is that such data often contain variability at several spatial and temporal scales. The space-time variability is further complicated due to different spatial behaviors at different time instants and vice versa. In this regard, the hierarchical modeling is found to be an effective means of handling this issue [180]. The essence of hierarchical analysis remains in its ability to model real-world scenario and handling complexity in the underlying processes. Additionally, a Bayesian hierarchical model can deal with the uncertainty in data. From the mid of the last decade, hierarchical modeling has shown rising popularity, especially in the domains of environmental study [24] (e.g., study of formation patterns

②ArcGIS Hub. https://hub.arcgis.com/pages/open-data, Nov. 2019.
③US Government's Open Data. https://www.data.gov/, Nov. 2019.
④UCI Machine Learning Repository. https://archive.ics.uci.edu/ml/datasets.php, Nov. 2019.
⑤Bureau of Meteorology, Australia. www.bom.gov.au, Nov. 2019.
⑥Open Govt. Data India. https://data.gov.in/catalogs/ministry_department/india-meteorological-department-imd, Nov. 2019.
⑦Climate Research Unit. http://www.cru.uea.ac.uk/data/, Nov. 2019.
⑧California Open Data Portal. https://data.ca.gov/dataset/caltrans-traffic-volumes, Nov. 2019.
⑨USGS-EarthExplorer. https://earthexplorer.usgs.gov/, Nov. 2019.
⑩Centres for Disease Control and Prevention. https://www.cdc.gov/dhdsp/maps/gisx/resources/geo-spatial-data.html, Nov. 2019.
⑪Mobility Datasets-GitHub. https://privamov.github.io/accio/docs/datasets.html, Nov. 2019.
⑫Stanford Large Network Dataset. https://snap.stanford.edu/data/, Nov. 2019.
⑬Microsoft Research. https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/, Nov. 2019.

of floods/droughts, ozone layer), ecology, and public health and safety [180] (e.g., crime pattern discovery, medical image analysis).

6.3 Handling Data Sparseness

Spatio-temporal data are often found to be sparse because of missing measurements, and this is very common for the moving object data, like GPS traces, which is unlikely to cover all edges during all time intervals [181]. Therefore, effectively dealing with the sparseness in ST data has become one of the prominent research areas of interest in recent years. Weight propagation [182], Laplace smoothing [183], Latent Space Model [184], Graph Convolutional Weight Completion techniques [185], etc. are some state-of-the-art approaches proposed to address the challenges of data sparseness in trajectory data. However, many of these techniques ignore the temporal-dependencies in the data, and thus, huge scopes remain in further exploring this issue and coming up with improved models.

6.4 Deep Learning

Since the mid of the last decade, deep learning has emerged as a new area of machine learning research and within a few years it has been able to show significant impact on a wide range of applications, including signal processing, information retrieval, natural language processing, text processing, multimodal information processing, and so on. However, as discussed in Section 4, the effectiveness of these algorithms on ST data analysis is still under-explored, since only a few researchers have considered this direction mostly for ST prediction and change pattern analysis purpose [87, 88, 100–102, 106]. Ample scopes remain in exploring deep learning for the other kinds of ST data analysis, including ST hotspot and ourlier detection, ST partitioning and summarization, ST coupling and tele-coupling, etc.

6.5 Participatory Sensing

Participatory sensing is a process in which individuals or communities use their personal mobile devices and cloud/web services to collect, analyze, and interpret data for systematically exploring interesting aspects of the surrounding world. This is termed as “citizen sensing”, “human-centric sensing”, “community sensing”, etc., as well. The objective is to utilize the real-time and the high-resolution spatial information from social media, such as Twitter, for aiding scientific research and decision making process. Participatory sensing has already shown encouraging effect in several applications, including urban risk management, urban temperature analysis, environmental monitoring, chronic disease management, and climate assessment [186]. Huge scopes remain in utilizing participatory sensing for other applications of ST data analysis.

6.6 Modeling Based on Theory-Guided Data Science

Incidentally, though data science models have shown outstanding success in several domains, their application is severely restricted in scientific problems that involve complex physical phenomena. Further, since the data science models act as a black boxes, they lack the ability to deliver a mechanistic understanding of the underlying processes and therefore cannot be used as a basis for subsequent scientific developments. Therefore, recently, the theory-guided data science has emerged as a new paradigm with an aim to systematically integrate scientific theories with data science models in the process of knowledge discovery [187]. This new paradigm has already begun to show promising performance in solving diverse categories of complex scientific problems including novel patterns and relationship discovery in climate science [188], density functional designing in quantum chemistry, surface water dynamics estimation in hydrology [189], etc.

7 Conclusions

In this article, we presented an overview of the state-of-the-art spatio-temporal data mining techniques and their applications in various domains. More than 300 papers, mostly from the last 10 years (2009-2018), have been studied to provide a comprehensive review of the present trend of research on six major families of ST data analysis, and also to indicate various promising directions of work in future. The research articles have been considered from 13 different application areas and various data-driven approaches have been categorized into traditional/pure statistical, CI-based, and deep learning based techniques. A detailed statistical analysis for this survey and the list of all the papers studied for this purpose can be found in our online supplementary document (see Section 5). At the end of study, it is clear that the data-driven modeling for ST data analysis is still a widely opened field for research having ample scopes, especially in further enhancing the existing models with integrated scientific theories.

References

[1] Samadi S, Wilson C A, Moradkhani H. Uncertainty analysis of statistical downscaling models using Hadley Centre Coupled Model. *Theoretical and Applied Climatology*, 2013, 114(3/4): 673-690.

[2] Ivanovic R F, Valdes P J, Flecker R, Gregoire L J, Gutjahr M. The parameterisation of Mediterranean-Atlantic water exchange in the Hadley Centre model HadCM3, and its effect on modelled North Atlantic climate. *Ocean Modelling*, 2013, 62: 11-16.

[3] Das M, Ghosh S K. Data-driven approaches for meteorological time series prediction: A comparative study of the state-of-the-art computational intelligence techniques. *Pattern Recognition Letters*, 2018, 105: 155-164.

[4] Solomatine D, See L, Abrahart R. Data-driven modelling: Concepts, approaches and experiences. In *Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications*, Abrahart R J, See L M, Solomatine D P (eds.), Springer, 2009, pp.17-30.

[5] Pelekis N, Theodoulidis B, Kopanakis I, Theodoridis Y. Literature review of spatio-temporal database models. *The Knowledge Engineering Review*, 2004, 19(3): 235-274.

[6] Ding Z, Yang B, Güting R H, Li Y. Network-matched trajectory-based moving-object database: Models and applications. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(4): 1918-1928.

[7] Ding Z, Yang B, Chi Y, Guo L. Enabling smart transportation systems: A parallel spatio-temporal database approach. *IEEE Transactions on Computers*, 2016, 65(5): 1377-1391.

[8] Shekhar S, Jiang Z, Ali R Y, Eftelioglu E, Tang X, Gunturi V, Zhou X. Spatiotemporal data mining: A computational perspective. *ISPRS International Journal of Geo-Information*, 2015, 4(4): 2306-2338.

[9] Zhou X, Shekhar S, Ali R Y. Spatiotemporal change footprint pattern discovery: An inter-disciplinary survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2014, 4(1): 1-23.

[10] Kisilevich S, Mansmann F, Nanni M, Rinzivillo S. Spatiotemporal clustering. In *Data Mining and Knowledge Discovery Handbook*, Maimon O, Rokach L (eds.), Springer, 2009, pp.855-874.

[11] Kawale J, Chatterjee S, Ormsby D, Steinhaeuser K, Liess S, Kumar V. Testing the significance of spatio-temporal teleconnection patterns. In *Proc. the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2012, pp.642-650.

[12] Cheng T, Haworth J, Anbaroglu B, Tanaksaranond G, Wang J. Spatiotemporal data mining. In *Handbook of Regional Science*, Fischer M M, Nijkamp P (eds.), Springer, 2014, pp.1173-1193.

[13] Aggarwal C C. Outlier analysis. In *Data Mining*, Aggarwal C C (ed.), Springer, 2015, pp.237-263.

[14] Atluri G, Karpatne A, Kumar V. Spatio-temporal data mining: A survey of problems and methods. *ACM Computing Surveys*, 2018, 51(4): Article No. 83.

[15] Das M, Ghosh S K. A cost-efficient approach for measuring Moran's index of spatial autocorrelation in geostationary satellite data. In *Proc. the 2016 IEEE International Conference on Geoscience and Remote Sensing Symposium*, July 2016, pp.5913-5916.

[16] Faghmous J H, Banerjee A, Shekhar S, Steinbach M, Kumar V, Ganguly A R, Samatova N. Theory-guided data science for climate change. *IEEE Computer*, 2014, 47(11): 74-78.

[17] Das M, Ghosh S K. FB-STEP: A fuzzy Bayesian network based data-driven framework for spatio-temporal prediction of climatological time series data. *Expert Systems with Applications*, 2019, 117: 211-227.

[18] Benmouiza K, Cheknane A. Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. *Theoretical and Applied Climatology*, 2019, 137(1/2): 31-43.

[19] Yaseen Z M, Ghareb M I, Ebtehaj I, Bonakdari H, Siddique R, Heddam S, Yusif A A, Deo R. Rainfall pattern forecasting using novel hybrid intelligent model based ANFIS-FFA. *Water Resources Management*, 2018, 32(1): 105-122.

[20] Kim H Y, Won C H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 2018, 103: 25-37.

[21] Duan P, Mao G, Liang W, Zhang D. A unified spatiotemporal model for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 20(9): 3212-3223.

[22] Cressie N, Wikle C K. Statistics for Spatio-Temporal Data (1st edition). John Wiley & Sons, 2011.

[23] Aryaputera A W, Yang D, Zhao L, Walsh W M. Very short-term irradiance forecasting at unobserved locations using spatio-temporal kriging. *Solar Energy*, 2015, 122: 1266-1278.

[24] Sahu S K, Bakar K S. Hierarchical Bayesian auto-regressive models for large space time data with applications to ozone concentration modelling. *Applied Stochastic Models in Business and Industry*, 2012, 28(5): 395-415.

[25] Kouziokas G N, Chatzigeorgiou A, Perakis K. Multilayer feed forward models in groundwater level forecasting using meteorological data in public management. *Water Resources Management*, 2018, 32(15): 5041-5052.

[26] Larsen P E, Field D, Gilbert J A. Predicting bacterial community assemblages using an artificial neural network approach. *Nature Methods,* 2012, 9(6): 621-625.

[27] Wang L, Wang Z, Qu H, Liu S. Optimal forecast combination based on neural networks for time series forecasting. *Applied Soft Computing*, 2018, 66: 1-17.

[28] Daliakopoulos I N, Tsanis I K. Comparison of an artificial neural network and a conceptual rainfall-runoff model in the simulation of ephemeral streamflow. *Hydrological Sciences Journal*, 2016, 61(15): 2763-2774.

[29] Raza A, Zhong M. Hybrid artificial neural network and locally weighted regression models for lane-based short-term urban traffic flow forecasting. *Transportation Planning and Technology*, 2018, 41(8): 901-917.

[30] Bazartseren B, Hildebrandt G, Holz K P. Short-term water level prediction using neural networks and neuro-fuzzy approach. *Neurocomputing*, 2003, 55(3/4): 439-450.

[31] Gerek I H. House selling price assessment using two different adaptive neuro-fuzzy techniques. *Automation in Construction*, 2014, 41: 33-39.

[32] Semero Y K, Zhang J, Zheng D, Wei D. A GA-PSO hybrid algorithm based neural network modeling technique for short-term wind power forecasting. *Distributed Generation & Alternative Energy Journal*, 2018, 33(4): 26-43.

[33] Wang X, Wen J, Zhang Y, Wang Y. Real estate price forecasting based on SVM optimized by PSO. *Optik — International Journal for Light and Electron Optics*, 2014, 125(3): 1439-1443.

[34] Gu J, Zhu M, Jiang L. Housing price forecasting based on genetic algorithm and support vector machine. *Expert Systems with Applications*, 2011, 38(4): 3383-3386.

[35] Nandar A. Bayesian network probability model for weather prediction. In *Proc. the 2009 International Conference on the Current Trends in Information Technology*, December 2009, Article No. 21.

[36] Das M, Ghosh S K. semBnet: A semantic Bayesian network for multivariate prediction of meteorological time series data. *Pattern Recognition Letters*, 2017, 93: 192-201.

[37] Madadgar S, Moradkhani H. Spatio-temporal drought forecasting within Bayesian networks. *Journal of Hydrology*, 2014, 512: 134-146.

[38] Wu Y, Xu W, Feng J, Palaiahnakote S, Lu T. Local and global Bayesian network based model for flood prediction. In *Proc. the 24th International Conference on Pattern Recognition*, August 2018, pp.225-230.

[39] Das M, Ghosh S, Gupta P, Chowdary V, Nagaraja R, Dadhwal V. FORWARD: A model for FOrecasting Reservoir WAteR dynamics using spatial Bayesian network (SpaBN). *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(4): 842-855.

[40] Das M, Ghosh S K. BESTED: An exponentially smoothed spatial Bayesian analysis model for spatio-temporal prediction of daily precipitation. In *Proc. the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, November 2017, Article No. 55.

[41] Das M, Ghosh S K, Chowdary V, Saikrishnaveni A, Sharma R. A probabilistic nonlinear model for forecasting daily water level in reservoir. *Water Resources Management*, 2016, 30(9): 3107-3122.

[42] Das M, Ghosh S K. Spatio-temporal prediction under scarcity of influencing variables: A hybrid probabilistic graph-based approach. In *Proc. the 9th International Conference on Advances in Pattern Recognition*, December 2017, Article No. 27.

[43] Yang B, Guo C, Jensen C S. Travel cost inference from sparse, spatio temporally correlated time series using Markov models. *Proceedings of the VLDB Endowment*, 2013, 6(9): 769-780.

[44] Yuan J, Zheng Y, Xie X, Sun G. Driving with knowledge from the physical world. In *Proc. the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2011, pp.316-324.

[45] Lu Q, Chang N B, Joyce J, Chen A S, Savic D A, Djordjevic S, Fu G. Exploring the potential climate change impact on urban growth in London by a cellular automata-based Markov chain model. *Computers, Environment and Urban Systems*, 2018, 68: 121-132.

[46] Das M, Ghosh S K. Short-term prediction of land surface temperature using multifractal detrended fluctuation analysis. In *Proc. the 2014 Annual IEEE India Conference*, 2014, Article No. 330.

[47] Kazem A, Sharifi E, Hussain F K, Saberi M, Hussain O K. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. *Applied Soft Computing*, 2013, 13(2): 947-958.

[48] Kisi O, Cimen M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *Journal of Hydrology*, 2011, 399(1/2): 132-140.

[49] Liu S, Yamada M, Collier N, Sugiyama M. Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks*, 2013, 43: 72-83.

[50] Zhou X, Shekhar S, Mohan P. Spatiotemporal change pattern mining: A multi-disciplinary perspective. In *Space-Time Integration in Geography and GIScience*, Kwan M P, Richardson D, Wang D, Zhou C H (eds.), Springer, 2015, pp.301-326.

[51] Liang S, Banerjee S, Carlin B P. Bayesian wombling for spatial point processes. *Biometrics*, 2009, 65(4): 1243-1253.

[52] Wu E, Liu W, Chawla S. Spatio-temporal outlier detection in precipitation data. In *Proc. the 2nd International Workshop on Knowledge Discovery from Sensor Data*, August 2010, pp.115-133.

[53] Huang L, Kulldorff M, Gregorio D. A spatial scan statistic for survival data. *Biometrics*, 2007, 63(1): 109-118.

[54] Nakaya T, Yano K. Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics. *Transactions in GIS*, 2010, 14(3): 223-239.

[55] Hsu K C, Li S T. Clustering spatial-temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources*, 2010, 33(2): 190-200.

[56] de Oliveira Silveira E M, de Mello J M, Júnior F W A, de Carvalho L M T. Object-based land-cover change detection applied to Brazilian seasonal savannahs using geostatistical features. *International Journal of Remote Sensing*, 2018, 39(8): 2597-2619.

[57] Das M, Ghosh S K. Spatio-temporal autocorrelation analysis for regional land-cover change detection from remote sensing data. In *Proc. the 4th ACM IKDD Conferences on Data Sciences*, March 2017, Article No. 8.

[58] Das M, Ghosh S K. Modeling spatio-temporal change pattern using mathematical morphology. In *Proc. the 3rd IKDD Conference on Data Science*, March 2016, Article No. 4.

[59] Sagar B S D. Visualization of spatiotemporal behavior of discrete maps via generation of recursive median elements. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(2): 378-384.

[60] Schubert E, Zimek A, Kriegel H P. Local outlier detection reconsidered: A generalized view on locality with applications to spatial, video, and network outlier detection. *Data Mining and Knowledge Discovery*, 2014, 28(1): 190-237.

[61] Sibolla B H, Coetzee S, van Zyl T L. A framework for visual analytics of spatio-temporal sensor observations from data streams. *ISPRS International Journal of Geo-Information*, 2018, 7(12): Article No. 475.

[62] Turchini F, Seidenari L, Bimbo A D. Convex polytope ensembles for spatio-temporal anomaly detection. In *Proc. the 19th International Conference on Image Analysis and Processing*, September 2017, pp.174-184.

[63] Liu C, Xiong H, Ge Y, Geng W, Perkins M. A stochastic model for context-aware anomaly detection in indoor location traces. In *Proc. the 12th IEEE International Conference on Data Mining*, December 2012, pp.449-458.

[64] Bu Y, Chen L, Fu A W C, Liu D. Efficient anomaly monitoring over moving object trajectory streams. In *Proc. the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, June 2009, pp.159-168.

[65] Li X, Han J, Kim S, Gonzalez H. ROAM: Rule- and motif-based anomaly detection in massive moving object data sets. In *Proc. the 2007 SIAM International Conference on Data Mining*, April 2007, pp.273-284.

[66] Wu H, Sun W, Zheng B. A fast trajectory outlier detection approach via driving behavior modeling. In *Proc. the 2017 ACM Conference on Information and Knowledge Management*, November 2017, pp.837-846.

[67] Chen C, Zhang D, Castro P S, Li N, Sun L, Li S, Wang Z. iBOAT: Isolation-based online anomalous trajectory detection. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(2): 806-818.

[68] Linton S L, Jennings J M, Latkin C A, Gomez M B, Mehta S H. Application of space-time scan statistics to describe geographic and temporal clustering of visible drug activity. *Journal of Urban Health*, 2014, 91(5): 940-956.

[69] Fanaee-T H, Gama J. Eigenspace method for spatiotemporal hotspot detection. *Expert Sys.: J. Knowl. Eng.*, 2015, 32(3): 454-464.

[70] Zhuang Y, Almeida M, Morabito M, Ding W. Crime hot spot forecasting: A recurrent model with spatial and temporal information. In *Proc. the 2017 IEEE International Conference on Big Knowledge*, August 2017, pp.143-150.

[71] Parent C, Spaccapietra S, Renso C *et al*. Semantic trajectories modeling and analysis. *ACM Computing Surveys*, 2013, 45(4): Article No. 42.

[72] Takahashi K, Kulldorff M, Tango T, Yih K. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *International Journal of Health Geographics*, 2008, 7(1): Article No. 14.

[73] Birant D, Kut A. ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data & Knowledge Engineering*, 2007, 60(1): 208-221.

[74] Schreck T, Tekušová T, Kohlhammer J, Fellner D. Trajectory-based visual analysis of large financial time series data. *ACM SIGKDD Explorations Newsletter*, 2007, 9(2): 30-37.

[75] Dong W, Yuan T, Yang K, Li C, Zhang S. Autoencoder regularized network for driving style representation learning. arXiv:1701.01272, 2017. https://arxiv.org/pdf/1701.01272.pdf, Oct. 2019.

[76] Kieu T, Yang B, Guo C, Jensen C S. Distinguishing trajectories from different drivers using incompletely labeled trajectories. In *Proc. the 27th ACM International Conference on Information and Knowledge Management*, October 2018, pp.863-872.

[77] Chen Z, Shen H T, Zhou X. Discovering popular routes from trajectories. In *Proc. the 27th IEEE International Conference on Data Engineering*, April 2011, pp.900-911.

[78] Jiang Z, Evans M, Oliver D, Shekhar S. Identifying K primary corridors from urban bicycle GPS trajectories on a road network. *Information Systems*, 2016, 57: 142-159.

[79] Pan B, Demiryurek U, Banaei-Kashani F, Shahabi C. Spatiotemporal summarization of traffic data streams. In *Proc. the ACM SIGSPATIAL International Workshop on GeoStreaming*, November 2010, pp.4-10.

[80] Evans M R, Oliver D, Shekhar S, Harvey F. Summarizing trajectories into k-primary corridors: A summary of results. In *Proc. the 20th International Conference on Advances in Geographic Information Systems*, November 2012, pp.454-457.

[81] Pillai K G, Angryk R A, Aydin B. A filter-and-refine approach to mine spatiotemporal co-occurrences. In *Proc. the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, November 2013, pp.104-113.

[82] Mohan P, Shekhar S, Shine J A, Rogers J P. Cascading spatio-temporal pattern discovery. *IEEE Transactions on Knowledge and Data Engineering*, 2012, 24(11): 1977-1992.

[83] Bermingham L, Lee I. Spatio-temporal sequential pattern mining for tourism sciences. *Procedia Computer Science*, 2014, 29: 379-389.

[84] Carlson A, Zaehringer J, Garrett R, Silva R F B, Furumo P, Rey A N, Torres A, Chung M G, Li Y, Liu J. Toward rigorous telecoupling causal attribution: A systematic review and typology. *Sustainability*, 2018, 10(12): Article No. 4426.

[85] Fang C, Liu H, Li G. International progress and evaluation on interactive coupling effects between urbanization and the eco-environment. *Journal of Geographical Sciences*, 2016, 26(8): 1081-1116.

[86] Deng L, Liu Y. Deep Learning in Natural Language Processing (1st edition). Springer, 2018.

[87] Kong Y, Gao S, Sun B, Fu Y. Action prediction from videos via memorizing hard-to-predict samples. In *Proc. the 32nd AAAI Conference on Artificial Intelligence*, February 2018, pp.7000-7007.

[88] 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*, 2017, 17(4): Article No. 818.

[89] Zhang J, Zheng Y, Qi D. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Proc. the 31st AAAI Conference on Artificial Intelligence*, February 2017, pp.1655-1661.

[90] Zhang J, Zheng Y, Qi D, Li R, Yi X. DNN-based prediction model for spatio-temporal data. In *Proc. the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, October 2016, Article No. 92.

[91] Kang H W, Kang H B. Prediction of crime occurrence from multi-modal data using deep learning. *PLoS ONE*, 2017, 12(4): Article No. e0176244.

[92] Lv Z, Xu J, Zheng K, Yin H, Zhao P, Zhou X. LC-RNN: A deep learning model for traffic speed prediction. In *Proc. the 27th International Joint Conference on Artificial Intelligence*, July 2018, pp.3470-3476.

[93] Jain A, Zamir A R, Savarese S, Saxena A. Structural-RNN: Deep learning on spatio-temporal graphs. In *Proc. the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, June 2016, pp.5308-5317.

[94] Lv Y, Duan Y, Kang W, Li Z, Wang F Y. Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 865-873.

[95] Li X, Peng L, Hu Y, Shao J, Chi T. Deep learning architecture for air quality predictions. *Environmental Science and Pollution Research*, 2016, 23(22): 22408-22417.

[96] Das M, Ghosh S K. Deep-STEP: A deep learning approach for spatiotemporal prediction of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(12): 1984-1988.

[97] Das M, Ghosh S K. A deep-learning-based forecasting ensemble to predict missing data for remote sensing analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(12): 5228-5236.

[98] Liao B, Zhang J,Wu C, McIlwraith D, Chen T, Yang S, Guo Y, Wu F. Deep sequence learning with auxiliary information for traffic prediction. In *Proc. the 24th ACM SIGKDD International Conference on Knowledge Discovery* & *Data Mining*, August 2018, pp.537-546.

[99] Pal S, Chowdhury S, Ghosh S K. DCAP: A deep convolution architecture for prediction of urban growth. In *Proc. the 2016 IEEE International Geoscience and Remote Sensing Symposium*, July 2016, pp.1812-1815.

[100] Mahdianpari M, Salehi B, Rezaee M, Mohammadimanesh F, Zhang Y. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sensing*, 2018, 10(7): Article No. 1119.

[101] Li W, Fu H, Yu L, Gong P, Feng D, Li C, Clinton N. Stacked autoencoder-based deep learning for remote-sensing image classification: A case study of African land-cover mapping. *International Journal of Remote Sensing*, 2016, 37(23): 5632-5646.

[102] Zhang P, Gong M, Su L, Liu J, Li Z. Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 116: 24-41.

[103] De S, Bruzzone L, Bhattacharya A, Bovolo F, Chaudhuri S. A novel technique based on deep learning and a synthetic target database for classification of urban areas in PolSAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 11(1): 154-170.

[104] Lyu H, Lu H, Mou L. Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sensing*, 2016, 8(6): Article No. 506.

[105] Zhuo L, Jiang L, Zhu Z, Li J, Zhang J, Long H. Vehicle classification for large-scale traffic surveillance videos using convolutional neural networks. *Machine Vision and Applications*, 2017, 28(7): 793-802.

[106] Zhu Y, Lan Z, Newsam S, Hauptmann A. Hidden two-stream convolutional networks for action recognition. In *Proc. the 14th Asian Conference on Computer Vision*, December 2018, pp.363-378.

[107] Zhao S, Liu Y, Han Y, Hong R, Hu Q, Tian Q. Pooling the convolutional layers in deep convNets for video action recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 28(8): 1839-1849.

[108] Wang X, Gao L, Wang P, Sun X, Liu X. Two-stream 3-D convNet fusion for action recognition in videos with arbitrary size and length. *IEEE Transactions on Multimedia*, 2017, 20(3): 634-644.

[109] Jia X, Khandelwal A, Nayak G, Gerber J, Carlson K, West P, Kumar V. Incremental dual-memory LSTM in land cover prediction. In *Proc. the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2017, pp.867-876.

[110] Zhou Y, Zhang F, Du Z, Ye X, Liu R. Integrating cellular automata with the deep belief network for simulating urban growth. *Sustainability*, 2017, 9(10): Article No. 1786.

[111] Xu D, Yan Y, Ricci E, Sebe N. Detecting anomalous events in videos by learning deep representations of appearance and motion. *Computer Vision and Image Understanding*, 2017, 156: 117-127.

[112] Zhao Y, Deng B, Shen C, Liu Y, Lu H, Hua X S. Spatio-temporal autoencoder for video anomaly detection. In *Proc. the 25th ACM International Conference on Multimedia*, October 2017, pp.1933-1941.

[113] Kieu T, Yang B, Jensen C S. Outlier detection for multidimensional time series using deep neural networks. In *Proc. the 19th IEEE International Conference on Mobile Data Management*, June 2018, pp.125-134.

[114] Zhu Y, Newsam S. Spatio-temporal sentiment hotspot detection using geotagged photos. In *Proc. the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, October 2016, Article No. 76.

[115] Chaudhuri G, Mishra N B. Spatio-temporal dynamics of land cover and land surface temperature in ganges-brahmaputra delta: A comparative analysis between India and Bangladesh. *Applied Geography*, 2016, 68: 68-83.

[116] Das M, Ghosh S K. Measuring Moran's I in a cost-efficient manner to describe a land-cover change pattern in large-scale remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(6): 2631-2639.

[117] Buyantuyev A, Wu J. Urban heat islands and landscape heterogeneity: Linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landscape Ecology*, 2010, 25(1): 17-33.

[118] Huang Y, Zhang L, Zhang P. A framework for mining sequential patterns from spatio-temporal event data sets. *IEEE Transactions on Knowledge and Data Engineering*, 2008, 20(4): 433-448.

[119] Kisi Ö, Shiri J, Nikoofar B. Forecasting daily lake levels using artificial intelligence approaches. *Computers* & *Geosciences*, 2012, 41: 169-180.

[120] Chang F J, Chang Y T. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources*, 2006, 29(1): 1-10.

[121] Zhou X, Yang T, Shi P, Yu Z, Wang X, Li Z. Prospective scenarios of the saltwater intrusion in an estuary under climate change context using Bayesian neural networks. *Stochastic Environmental Research and Risk Assessment*, 2017, 31(4): 981-991.

[122] Han S, Coulibaly P. Bayesian flood forecasting methods: A review. *Journal of Hydrology*, 2017, 551: 340-351.

[123] Fotovatikhah F, Herrera M, Shamshirband S, Chau K W, Faizollahzadeh Ardabili S, Piran M J. Survey of computational intelligence as basis to big flood management: Challenges, research directions and future work. *Engineering Applications of Computational Fluid Mechanics*, 2018, 12(1): 411-437.

[124] Haest B, Hüppop O, Bairlein F. The influence of weather on avian spring migration phenology: What, where and when? *Global Change Biology*, 2018, 24(12): 5769-5788.

[125] Castaño C, Lindahl B D, Alday J G, Hagenbo A, de Aragón J, Parladé J, Pera J, Bonet J A. Soil microclimate changes affect soil fungal communities in a Mediterranean pine forest. *New Phytologist*, 2018, 220(4): 1211-1221.

[126] Thom D, Rammer W, Dirnböck T, Müller J, Kobler J, Katzensteiner K, Helm N, Seidl R. The impacts of climate change and disturbance on spatio-temporal trajectories of biodiversity in a temperate forest landscape. *Journal of Applied Ecology*, 2017, 54(1): 28-38.

[127] Bruckner M, Fischer G, Tramberend S, Giljum S. Measuring telecouplings in the global land system: A review and comparative evaluation of land footprint accounting methods. *Ecological Economics*, 2015, 114: 11-21.

[128] Schalk S G, Demi L, Smeenge M, Mills D M, Wallace K D, Rosette J J, Wijkstra H, Mischi M. 4-D spatiotemporal analysis of ultrasound contrast agent dispersion for prostate cancer localization: A feasibility study. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2015, 62(5): 839-851.

[129] Rahman M M, Feng Y, Yankeelov T E, Oden J T. A fully coupled space-time multiscale modeling framework for predicting tumor growth. *Computer Methods in Applied Mechanics and Engineering*, 2017, 320: 261-286.

[130] Macharia P M, Giorgi E, Noor A M, Waqo E, Kiptui R, Okiro E A, Snow R W. Spatio-temporal analysis of *plasmodium falciparum* prevalence to understand the past and chart the future of malaria control in Kenya. *Malaria Journal*, 2018, 17(1): Article No. 340.

[131] Guinat C, Nicolas G, Vergne T, Bronner A, Durand B, Courcoul A, Gilbert M, Guérin J L, Paul M C. Spatiotemporal patterns of highly pathogenic avian influenza virus subtype H5N8 spread, France, 2016 to 2017. *Eurosurveillance*, 2018, 23(26): Article No. 1.

[132] Lai P C, Chow C B, Wong H T, Kwong K H, Kwan Y W, Liu S H, Tong W K, Cheung W K, Wong W L. An early warning system for detecting H1N1 disease outbreak — A spatio-temporal approach. *International Journal of Geographical Information Science*, 2015, 29(7): 1251-1268.

[133] Bisanzio D, Dzul-Manzanilla F, Gomez-Dantés H *et al*. Spatiotemporal coherence of dengue, chikungunya and Zika outbreaks in Merida, Mexico. *PLoS: Neglected Tropical Diseases*, 2018, 12(3): Article No. 0006298.

[134] Vlahogianni E I, Karlaftis M G, Golias J C. Short-term traffic forecasting: Where we are and where were going. *Transportation Research Part C: Emerging Technologies*, 2014, 43: 3-19.

[135] Nguyen H, Liu W, Chen F. Discovering congestion propagation patterns in spatio-temporal traffic data. *IEEE Transactions on Big Data*, 2017, 3(2): 169-180.

[136] Lim K H, Chan J, Leckie C, Karunasekera S. Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. *Knowledge and Information Systems*, 2018, 54(2): 375-406.

[137] Dai J, Yang B, Guo C, Ding Z. Personalized route recommendation using big trajectory data. In *Proc. the 31st IEEE International Conference on Data Engineering*, April 2015, pp.543-554.

[138] Yang B, Guo C, Ma Y, Jensen C S. Toward personalized, context-aware routing. *The VLDB Journal*, 2015, 24(2): 297-318.

[139] Yang B, Dai J, Guo C, Jensen C S, Hu J. PACE: A PathCEntric paradigm for stochastic path finding. *The VLDB Journal*, 2018, 27(2): 153-178.

[140] Hu J, Yang B, Guo C, Jensen C S. Risk-aware path selection with time-varying, uncertain travel costs: A time series approach. *The VLDB Journal*, 2018, 27(2): 179-200.

[141] Wang P, Fu Y, Zhang J, Wang P, Zheng Y, Aggarwal C. You are how you drive: Peer and temporal-aware representation learning for driving behavior analysis. In *Proc. the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, August 2018, pp.2457-2466.

[142] Morton J, Wheeler T A, Kochenderfer M J. Analysis of recurrent neural networks for probabilistic modeling of driver behavior. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 18(5): 1289-1298.

[143] Subasinghe S, Estoque R C, Murayama Y. Spatiotemporal analysis of urban growth using GIS and remote sensing: A case study of the Colombo Metropolitan Area, Sri Lanka. *ISPRS International Journal of Geo-Information*, 2016, 5(11): Article No. 197.

[144] Scheepens R, Hurter C, van de Wetering H, van Wijk J J. Visualization, selection, and analysis of traffic flows. *IEEE Transactions on Visualization and Computer Graphics*, 2016, 22(1): 379-388.

[145] Voulis N, Warnier M, Brazier F M. Understanding spatio-temporal electricity demand at different urban scales: A data-driven approach. *Applied Energy*, 2018, 230: 1157-1171.

[146] Padilla C M, Kihal-Talantikite W, Vieira V M, Rossello P, Le Nir G, Zmirou-Navier D, Deguen S. Air quality and social deprivation in four French metropolitan areas — A localized spatio-temporal environmental inequality analysis. *Environmental Research*, 2014, 134: 315-324.

[147] Smith T E, Wu P. A spatio-temporal model of housing prices based on individual sales transactions over time. *Journal of Geographical Systems*, 2009, 11(4): 333-355.

[148] Holly S, Pesaran M H, Yamagata T. The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics*, 2011, 69(1): 2-23.

[149] Dubé J, Legros D. A spatio-temporal measure of spatial dependence: An example using real estate data. *Papers in Regional Science*, 2013, 92(1): 19-30.

[150] Chang H, Smallwood P M, Williams J, Nathans J. The spatio-temporal domains of Frizzled6 action in planar polarity control of hair follicle orientation. *Developmental Biology*, 2016, 409(1): 181-193.

[151] Bindea G, Mlecnik B, Tosolini M *et al*. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*, 2013, 39(4): 782-795.

[152] Lena P D, Nagata K, Baldi P F. Deep spatio-temporal architectures and learning for protein structure prediction. In *Proc. the 26th Annual Conference on Neural Information Processing Systems*, December 2012, pp.512-520.

[153] Cao C, Liu F, Tan H, Song D, Shu W, Li W, Zhou Y, Bo X, Xie Z. Deep learning and its applications in biomedicine. *Genomics, Proteomics & Bioinformatics*, 2018, 16(1): 17-32.

[154] Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of deep learning in biomedicine. *Molecular Pharmaceutics*, 2016, 13(5): 1445-1454.

[155] Lan K, Wang D T, Fong S, Liu L S, Wong K K, Dey N. A survey of data mining and deep learning in bioinformatics. *Journal of Medical Systems*, 2018, 42(8): Article No. 139.

[156] Brovelli M A, Minghini M, Zamboni G. Public participation in GIS via mobile applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 114: 306-315.

[157] Hsieh H P, Lin S D, Zheng Y. Inferring air quality for station location recommendation based on urban big data. In *Proc. the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2015, pp.437-446.

[158] Liu Y, Pham T A N, Cong G, Yuan Q. An experimental evaluation of point-of-interest recommendation in location-based social networks. *Proceedings of the VLDB Endowment*, 2017, 10(10): 1010-1021.

[159] Zhao P, Xu X, Liu Y, Sheng V S, Zheng K, Xiong H. Photo2Trip: Exploiting visual contents in geo-tagged photos for personalized tour recommendation. In *Proc. the 25th ACM International Conference on Multimedia*, October 2017, pp.916-924.

[160] Jiang S, Qian X, Mei T, Fu Y. Personalized travel sequence recommendation on multi-source big social media. *IEEE Transactions on Big Data*, 2016, 2(1): 43-56.

[161] Li Y, Yiu M L. Route-saver: Leveraging route APIs for accurate and efficient query processing at location-based services. *IEEE Transactions on Knowledge and Data Engineering*, 2015, 27(1): 235-249.

[162] Guo J, Wu Y, Zhang X, Zhang L, Chen W, Cao Z, Zhang L, Guo H. Finding the 'faster' path in vehicle routing. *IET Intelligent Transport Systems*, 2017, 11(10): 685-694.

[163] Wang Z, Ye X. Social media analytics for natural disaster management. *International Journal of Geographical Information Science*, 2018, 32(1): 49-72.

[164] Kong X, Li M, Ma K, Tian K, Wang M, Ning Z, Xia F. Big trajectory data: A survey of applications and services. *IEEE Access*, 2018, 6: 58295-58306.

[165] Huang Z, Zhao Z, E S J, Yu C, Shan G, Li T, Cheng J, Sun J, Xiang Y. PRACE: A taxi recommender for finding passengers with deep learning approaches. In *Proc. the 13th International Conference on Intelligent Computing*, August 2017, pp.759-770.

[166] Li P, Yao J, Wang L, Lin X. SPOT: Selecting occuPations from trajectories. In *Proc. the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, August 2017, pp.813-816.

[167] Wang Y, Qin K, Chen Y, Zhao P. Detecting anomalous trajectories and behavior patterns using hierarchical clustering from taxi GPS data. *ISPRS International Journal of Geo-Information*, 2018, 7(1): Article No. 25.

[168] Dabiri S, Heaslip K. Inferring transportation modes from GPS trajectories using a convolutional neural network. *Transportation Research Part C: Emerging Technologies*, 2018, 86: 360-371.

[169] Zheng Y. Trajectory data mining: An overview. *ACM Transactions on Intelligent Systems and Technology*, 2015, 6(3): Article No. 29.

[170] Harvey F. To volunteer or to contribute locational information? Towards truth in labeling for crowdsourced geographic information. In *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, Sui D, Elwood S, Goodchild M (eds.), Springer, 2013, pp.31-42.

[171] Zhang C, Zhou G, Yuan Q, Zhuang H, Zheng Y, Kaplan L, Wang S, Han J. GeoBurst: Real-time local event detection in geo-tagged tweet streams. In *Proc. the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, July 2016, pp.513-522.

[172] Martín Y, Li Z, Cutter S L. Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew. *PLoS ONE*, 2017, 12(7): Article No. e0181701.

[173] Zhao L, Sun Q, Ye J, Chen F, Lu C T, Ramakrishnan N. Multi-task learning for spatio-temporal event forecasting. In *Proc. the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2015, pp.1503-1512.

[174] Paul D, Li F, Teja M K, Yu X, Frost R. Compass: Spatio temporal sentiment analysis of US election what Twitter says! In *Proc. the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2017, pp.1585-1594.

[175] Steiger E, Albuquerque J P, Zipf A. An advanced systematic literature review on spatiotemporal analyses of Twitter data. *Transactions in GIS*, 2015, 19(6): 809-834.

[176] Khalid S, Shoaib F, Qian T, Rui Y, Bari A I, Sajjad M, Shakeel M, Wang J. Network constrained spatio-temporal hotspot mapping of crimes in Faisalabad. *Applied Spatial Analysis and Policy*, 2018, 11(3): 599-622.

[177] Yu C H, Ding W, Chen P, Morabito M. Crime forecasting using spatio-temporal pattern with ensemble learning. In *Proc. the 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, May 2014, pp.174-185.

[178] von Landesberger T, Brodkorb F, Roskosch P, Andrienko N, Andrienko G, Kerren A. MobilityGraphs: Visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering. *IEEE Transactions on Visualization and Computer Graphics*, 2016, 22(1): 11-20.

[179] Bak P, Ship H, Yaeli A, Nardi Y, Packer E, Saadoun G, Bnayahu J, Peterfreund L. Visual analytics for movement behavior in traffic and transportation. *IBM Journal of Research and Development*, 2015, 59(2/3): Article No. 10.

[180] Yu C H, Ding W, Morabito M, Chen P. Hierarchical spatio-temporal pattern discovery and predictive modeling. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(4): 979-993.

[181] Guo C, Yang B, Hu J, Jensen C. Learning to route with sparse trajectory sets. In *Proc. the 34th IEEE International Conference on Data Engineering*, April 2018, pp.1073-1084.

[182] Idé T, Sugiyama M. Trajectory regression on road networks. In *Proc. the 35th AAAI Conference on Artificial Intelligence*, August 2011, Article No. 32.

[183] Yang B, Kaul M, Jensen C S. Using incomplete information for complete weight annotation of road networks. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(5): 1267-1279.

[184] Deng D, Shahabi C, Demiryurek U, Zhu L, Yu R, Liu Y. Latent space model for road networks to predict time-varying traffic. In *Proc. the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August 2016, pp.1525-1534.

[185] Hu J, Guo C, Yang B, Jensen C S. Stochastic weight completion for road networks using graph convolutional networks. In *Proc. the 34th IEEE International Conference on Data Engineering*, April 2019, pp.1274-1285.

[186] Henriksen H J, Roberts M J, Keur P, Harjanne A, Egilson D, Alfonso L. Participatory early warning and monitoring systems: A Nordic framework for web-based flood risk management. *International Journal of Disaster Risk Reduction*, 2018, 31: 1295-1306.

[187] Karpatne A, Atluri G, Faghmous J H, Steinbach M, Banerjee A, Ganguly A, Shekhar S, Samatova N, Kumar V. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(10): 2318-2331.

[188] Faghmous J H, Frenger I, Yao Y, Warmka R, Lindell A, Kumar V. A daily global mesoscale ocean eddy dataset from satellite altimetry. *Scientific Data*, 2015, 2: Article No. 150028.

[189] Khandelwal A, Karpatne A, Marlier M E, Kim J, Lettenmaier D P, Kumar V. An approach for global monitoring of surface water extent variations in reservoirs using MODIS data. *Remote Sensing of Environment*, 2017, 202: 113-128.

Monidipa Das is currently working as a postdoctoral research fellow in the Computational Intelligence Laboratory, in the School of Computer Science and Engineering (SCSE), Nanyang Technological University (NTU), Singapore. She received her Ph.D. degree in computer science and engineering from the Indian Institute of Technology (IIT) Kharagpur, Kharagpur, in 2018, and her M.E. degree in computer science and engineering from the Indian Institute of Engineering Science and Technology (IIEST), Shibpur, in 2013. Her research interests include spatial informatics, spatio-temporal data mining, soft computing, and machine learning. Dr. Das is a member of IEEE.

Soumya K. Ghosh received his Ph.D. and M.Tech. degrees in computer science and engineering from the Indian Institute of Technology (IIT) Kharagpur, Kharagpur. Presently, he is a professor with the Department of Computer Science and Engineering, IIT Kharagpur. Before joining IIT Kharagpur, Kharagpur, he worked for the Indian Space Research Organization in the area of satellite remote sensing and geographic information systems. He has been awarded the National Geospatial Chair Professorship by the Department of Science and Technology, Government of India, in 2017. He has more than 300 research papers in reputed journals and conference proceedings. His research interests include spatial data science, spatial web services, and cloud computing. Dr. Ghosh is a senior member of IEEE.