Spatio-Temporal Data Mining: A Survey of Problems and Methods

GOWTHAM ATLURI*, University of Cincinnati
ANUJ KARPATNE* and VIPIN KUMAR, University of Minnesota

Large volumes of spatio-temporal data are increasingly collected and studied in diverse domains, including climate science, social sciences, neuroscience, epidemiology, transportation, mobile health, and Earth sciences. Spatio-temporal data differ from relational data for which computational approaches are developed in the data-mining community for multiple decades in that both spatial and temporal attributes are available in addition to the actual measurements/attributes. The presence of these attributes introduces additional challenges that needs to be dealt with. Approaches for mining spatio-temporal data have been studied for over a decade in the data-mining community. In this article, we present a broad survey of this relatively young field of spatio-temporal data mining. We discuss different types of spatio-temporal data and the relevant data-mining questions that arise in the context of analyzing each of these datasets. Based on the nature of the data-mining problem studied, we classify literature on spatio-temporal data mining into six major categories: clustering, predictive learning, change detection, frequent pattern mining, anomaly detection, and relationship mining. We discuss the various forms of spatio-temporal data-mining problems in each of these categories.

CCS Concepts: • **Computing methodologies** → *Artificial intelligence*; **Machine learning**; • **Applied computing** → *Physical sciences and engineering*;

Additional Key Words and Phrases: Data mining, spatiotemporal data, spatial data, geographic data, time-series, rasters

1 INTRODUCTION

Space and time are ubiquitous aspects of observations in a number of domains, including climate science, neuroscience, social sciences, epidemiology, transportation, criminology, and Earth sciences, that are rapidly being transformed by the deluge of data. Since the real-world processes being studied in these domains are inherently spatio-temporal in nature, a number of data collection methodologies have been devised to record the spatial and temporal information of every measurement in the data, hereby referred to as spatio-temporal (ST) data. For example,

---

in neuroimaging data, activity measured from the human brain is stored along with the spatial location from which the activity was measured and the time at which the measurement was made. Similarly, web-search requests arriving at Google's servers have a geographic location and time from which they are made. Effective analysis of such increasingly prevalent ST data holds great promise for advancing the state of the art in several scientific disciplines.

A unique quality of ST data that differentiates it from other data studied in classical data-mining literature (e.g., see Tan et al. (2017)) is the presence of dependencies among measurements induced by the spatial and temporal dimensions. For example, many of the widely used data-mining methods are founded on the assumption that data instances are independent and identically distributed (i.i.d.). However, this assumption is violated when dealing with ST data, where instances are structurally related to each other in the context of space and time and show varying properties in different spatial regions and time periods. Ignoring these dependencies during data analysis can lead to poor accuracy and interpretability of results (Eklund et al. 2016).

Apart from limiting the effectiveness of classical data-mining algorithms, the presence of spatial and temporal information also makes it possible to consider novel formulations for analyzing data in the emerging field of spatio-temporal data mining (STDM). Contrary to traditional data mining that deals with distinct objects (also referred to as data instances) having well-defined features, in STDM, one can define objects and features in a variety of ways. One scenario involves treating spatial locations as objects and using the measurements collected from a spatial location over time to define the features. For example, in climate science, one of the goals is to group locations that experience similar climatic phenomenon over time. In this case, locations are treated as instances/objects and features are defined based on climate variables measured over time (Steinbach et al. 2002). Another scenario involves treating time points as objects and using measurements collected from all the spatial locations under consideration to define features. For example, in the application of discovering patterns of human brain activity from neuroimages, the goal is to identify the time points at which similar brain activity is observed in the brain. In this case, time points are treated as objects/instances and features are defined using the observed spatial map of activity (Liu et al. 2013). There are also scenarios where events are treated as objects and features are defined based on the spatial and temporal information of events. For example, in the context of discovering crimes that are committed in close proximity in space and time, incidence of a crime is treated as an object and the location and time stamp of the crime are treated as features, in addition to other features such as nature of crime and number of victims involved (Tompson et al. 2015). Hence, the coupling of spatial and temporal information in ST data introduces novel problems, challenges, and opportunities for STDM research, with a broad scope of application in several domains of scientific and commercial significance.

There exists a vast literature on approaches for mining data that is purely spatial in nature, spanning multiple decades of research in spatial statistics (Cressie and Wikle 2015), spatial data mining (Aggarwal 2015; Shekhar et al. 2011, 2008), and spatial database management (Ester et al. 1997; Shekhar and Chawla 2003). An extensive taxonomy of spatial data types and representations has been explored in the field of spatial data mining for improving the efficiency and effectiveness of data-mining tasks such as clustering, prediction, anomaly detection, and pattern mining when dealing with spatial data (Shekhar et al. 2011). Another related area of research is time-series data mining (Esling and Agon 2012; Keogh and Kasetty 2003; Liao 2005), where approaches for mining useful information from time-series databases have been explored. Existing research in STDM includes foundational research in the statistics community (Cressie and Wikle 2015), e.g., research on spatio-temporal point processes (Diggle 2013). Approaches for handling spatial and temporal information have also been explored in the data-mining literature for problems such as spatio-temporal clustering (Kisilevich et al. 2009) and trajectory pattern mining (Giannotti et al. 2007).

There are a few recent surveys that have reviewed the literature on STDM in certain contexts from different perspectives. Articles by Vatsavai et al. (2012) and Chandola et al. (2015) discuss the computational issues for STDM algorithms in the era of "big-data" for application domains such as remote sensing, climate science, and social media analysis. The review by Cheng et al. (2014) covers STDM approaches for prediction, clustering, and visualization problems in several applications. An extensive survey of approaches for mining trajectory data, one of the many types of ST data, is presented in Li (2014), Zheng (2015), and Mamoulis (2009). A survey on STDM by Shekhar et al. (2015) provides a semantic categorization of ST data types and pattern families from a database-centric perspective.

Given the richness of problems and the variety of methods being explored in the rapidly advancing field of STDM, there is a need for developing an over-arching structure of research in STDM that highlights the similarities and differences of different problems and methods in diverse ST applications. This can enable the cross-pollination of ideas across disparate research areas and application domains by making it possible to see how a solution developed for a certain problem in a particular domain (e.g., identifying patterns in climate data) can be useful for solving a different problem in another domain (e.g., understanding the working of the brain). This can also help in connecting the traditional data-mining community with the challenges and opportunities in analyzing ST data, thus exposing some of the open questions and motivating future directions of research in STDM.

This review article on STDM problems and methods attempts to fulfill this need as follows. First, it builds a foundation of ST data types and properties that can help in identifying the relevant problems and methods for any class of ST data encountered in real-world applications. In particular, we provide a broad taxonomy of the different types of ST data, different ways of defining and describing ST data instances, and different ways of computing similarity among ST data instances. Second, it presents a survey of STDM approaches for a number of commonly studied data-mining problems such as clustering, predictive learning, frequent pattern mining, anomaly detection, change detection, and relationship mining. For every category of problems, we review the novel issues that arise in dealing with the unique properties of ST data types in classical data-mining frameworks. This article can be used as a guide by data-mining researchers and real-world practitioners working with ST data to identify STDM formulations fit for their data and to make most effective use of STDM research in their problem. In addition, by bridging the gap between classical data-mining literature and the novel aspects of spatio-temporal data, this article helps in opening novel possibilities of future research.

The rest of the article is organized as follows. In Section 2, we review the variety of application areas where analyzing ST data is important. In Section 3, we discuss the types and characteristics of ST data, and the different ways of defining instances and similarity measures using ST data types. Section 4 presents a survey of STDM methods developed for different types of ST data instances in the context of six major data-mining problems, viz., clustering, predictive learning, frequent pattern mining, anomaly detection, change detection, and relationship mining. Section 5 presents concluding remarks and discusses future research directions.

2   APPLICATIONS

Large volumes of ST data are collected in several application domains such as social media, healthcare, agriculture, transportation, and climate science. In this section, we briefly describe the different sources of ST data and the motivation for analyzing ST data in different application domains.

Climate Science: Data pertaining to historic and current atmospheric and oceanic conditions (e.g., temperature, pressure, wind-flow, and humidity) is collected and studied in climate science

(Karpatne et al. 2013). In addition to observational data[1] collected from weather stations and reanalysis data that is gridded in space (Kistler et al. 2001), simulated data generated using climate models (Voldoire et al. 2013) is also studied in this domain. The purpose in studying this data is to discover relationships and patterns in climate science that advance our understanding of the Earth's system and help us better prepare for future adverse conditions by informing adaptation and mitigation actions in a timely manner.

Neuroscience: Continuous neural activity captured using a variety of technologies such as Functional Magnetic Resonance Imaging (fMRI), Electroencephalogram (EEG), and Magnetoencephalography (MEG) is studied in neuroscience (Atluri et al. 2016). Spatial resolution of neural activity measured using these technologies is quite different from another. For example, neural activity is measured from millions of locations in fMRI data, while it is only measured from tens of locations in the case of EEG data. Temporal resolution of the data collected using these technologies is also quite different. For example, fMRI typically measures activity for every 2s, while the temporal resolution of EEG data is is typically 1ms. The purpose in studying these data is to understand the governing principles of the brain and thereby determine the disruptions to normal conditions that arise in the case of mental disorders (Atluri et al. 2013, 2015). Discovering such disruptions can be useful for designing diagnostic procedures and in developing therapeutic procedures for patients.

Environmental Science: Studying the data pertaining to the quality of air, water, and environment is one of the objectives of environmental science. While air quality is measured based on the presence of pollutants such as particles, carbon monoxide, nitrogen dioxide, sulphur dioxide, ozone, and so on, water quality is measured based on factors such as dissolved oxygen, conductivity, turbidity, and pH. Air quality sensors are typically placed on streets or on top of buildings, and water quality sensors are placed in lakes, rivers, and streams. In addition to air and water quality, data pertaining to sound pollution is also collected. These environmental datasets are studied to detect changes in levels of pollution, identify the causal factors that contribute to pollution, and design effective policies to reduce the different types of pollution (Thompson et al. 2014).

Precision Agriculture: Multi-band high-resolution (ranging from 0.25m to 1m) areal or remote-sensing images of large farms are being collected at regular intervals (e.g., daily to weekly). One of the purposes of collecting and studying these data is to detect plant diseases (Mahlein 2016) and understand the impact of several factors such as misapplication of fertilizer, compaction during planting and weeds on crop yield, as well as their inter-relationships. With the help of this knowledge, steps can be taken in future crop cycles to mitigate the risks due to the factors that adversely affect the crop yield.

Epidemiology/Health care: Electronic health record data that are widely stored in hospitals provide demographic information pertaining to patients as well diagnoses made on patients at different time points. This dataset can be represented as a spatio-temporal dataset where each diagnosis has a spatial location and a time-point associated with it. One can construct such spatio-temporal instances for different types of diseases such as cancers and diabetes, as well as for infectious diseases such as influenza. These data are studied to discover spatio-temporal patterns in different diseases (Matsubara et al. 2014) and to study the spread of an epidemic. These data are also used in conjunction with environmental, climate science datasets to discover relationships between environmental factors and public health (Ryan et al. 2007). Discovery of such relationships will allow policy-makers to develop effective policies that will ensure the well-being of the population.

---

[1]https://www.ncdc.noaa.gov/cdo-web/datasets.

Social media: Users of social media portals such as Twitter and Facebook post their experience at a given place and time. Each social media post captures the experience of a user at a given place and time. Using these data, one can study collective user experience at a given place for a given time period (Tang et al. 2014). One can also capture the spread of epidemics such as *Influenza* or *ebola* based on users' posts. More recently, there is also increased interest in studying the spread of social and political movements using social media data (Carney 2016). In addition, events such as earthquakes, tsunamis, and fires can also be automatically detected from this data.

Traffic Dynamics: Large-scale taxi pick-up/drop-off data are publicly available for several major cities across the world (Castro et al. 2013). These data contain information about each trip made by customers of the taxi service, including the time and location of pick-up and drop-off, and GPS locations for each second during the taxi ride. These data can be used to understand how the population in a city moves spatially as a function of time (e.g., see the City of Citizens[2] exploratory in the SoBigData infrastructure) and also the influence of extraneous factors such as traffic and weather. In addition, these data can also be studied to explore traffic dynamics based on the collective movement patterns of the taxis. This will enable transportation engineers to design effective policies to reduce traffic congestion. In addition, the behavior of taxi drivers can also be studied using these data so effective practices can be designed to detect abnormal behavior, increasing the likelihood of finding new passengers, and taking optimal routes to arrive at a destination.

Heliophysics: Heliophysics studies the events that occur in the Sun and their impact on the solar system. The publicly available Heliophysics Events Knowledgebase (Hurlburt et al. 2010) provides various observations that include solar events and their annotations on a daily basis. Examples of these events include Active region, Emerging flux, Filament, Flare, Sigmoid, and Sunspot. The time and the location of where these events were observed on the Sun are also provided in the Knowledgebase. The spatial and temporal information along with the different observations are studied to discover patterns in the solar events (Pillai et al. 2012). The Heliophysics Knowledgebase also enables the study of the impact of solar events and the Earth's climate system.

Crime data: Law enforcement agencies store information about reported crimes in many cities, and this information is made publicly available in the spirit of *open-data* (Tompson et al. 2015). These data typically have the type of crime (e.g., arson, assault, burglary, robbery, theft, and vandalism), as well as the time and location of the crime. Patterns in crime and the effect of law enforcement policies on the amount of crime in a region can be studied using these data with the goal of reducing crime.

3 DATA

The presence of space and time introduces a rich diversity of ST data types and representations, which leads to multiple ways of formulating STDM problems and methods. In this section, we first describe some of the generic properties of ST data and then describe the basic types of ST data available in different applications. Building on this discussion, we describe some of the common ways of defining and representing ST data instances and generic methods for computing similarity among different types of ST instances.

3.1 Properties

There are two generic properties of ST data that introduce challenges as well as opportunities for classical data-mining algorithms, as described in the following.

3.1.1 Auto-Correlation. In domains involving ST data, the observations made at nearby locations and time stamps are not *independent* but are correlated with each other. This auto-correlation

---

[2]http://www.sobigdata.eu/exploratories/568.

in ST datasets results in a coherence of spatial observations (e.g., surface temperature values are consistent at nearby locations) and smoothness in temporal observations (e.g., changes in traffic activity occurs smoothly over time). As a result, classical data-mining algorithms that assume independence among observations are not well-suited for ST applications, often resulting in poor performance with *salt-and-pepper* errors (Jiang et al. 2015). Further, standard evaluation schemes such as cross-validation may become invalid in the presence of ST data, because the test error rate can be contaminated by the training error rate when random sampling approaches are used to generate training and test sets that are correlated with each other. We also need novel ways of evaluating the predictions of STDM methods, because estimates of the location/time of an ST object (e.g., a crime event) may be useful even if they are not exact but in the close ST vicinity of ground-truth labels. Hence, there is a need to account for the structure of auto-correlation among observations while analyzing ST datasets.

3.1.2 Heterogeneity. Another basic assumption that is made by classical data-mining formulations is the homogeneity (or stationarity) of instances, which implies that every instance belongs to the same population and is thus *identically distributed*. However, ST datasets can show heterogeneity (or non-stationarity) both in space and time in varying ways and levels. For example, satellite measurements of vegetation at a location on Earth show a cyclical pattern in time due to the presence of seasonal cycles. Hence, observations made in winter are differently distributed than the observations made in summer. There can also be inter-annual changes due to regime shifts in the Earth's climate, e.g., El Nino phase transitions, which can impact climate patterns to change on a global scale. As another example, different spatial regions of the brain perform different functions and hence show varying physiological responses to a stimuli. This heterogeneity in space and time requires the learning of different models for varying spatio-temporal regions.
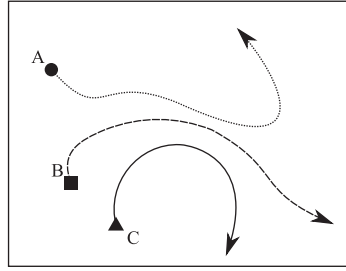
3.2 Data Types

There is a variety of ST data types that one can encounter in different real-world applications. They differ in the way space and time are used in the process of data collection and representation and lead to different categories of STDM problem formulations. For this reason, it is important to establish the type of ST data available in a given application to make the most effective use of STDM methods. In the following, we describe four common categories of ST data types: (i) event data, which comprise discrete events occurring at point locations and times (e.g., incidences of crime events in the city); (ii) trajectory data, where trajectories of moving bodies are being measured (e.g., the patrol route of a police surveillance car); (iii) point reference data (Cressie and Wikle 2015), where a continuous ST *field* is being measured at moving ST reference sites (e.g., measurements of surface temperature collected using weather balloons); and (iv) raster data, where observations of an ST field is being collected at fixed cells in an ST grid (e.g., fMRI scans of brain activity). While the first two data types (events and trajectories) record observations of discrete events and objects, the next two data types (point reference and rasters) capture information of continuous or discrete ST fields. We discuss the basic properties of these four data types using illustrative examples from diverse applications. Indeed, if an ST dataset is collected in a native data type that is different from the one we intend to use, in some cases, it is possible to convert from one ST data type to another, e.g., from point reference data to raster data. We briefly discuss some of the possible ways of converting an ST data type to other data types, for leveraging the STDM methods developed for those data types in a particular application.

3.2.1 Event Data. An ST event can generally be characterized by a *point* location and time, which denotes where and when the event occurred, respectively. For example, a crime event can be characterized by the location of the crime along with the time at which the crime activity occurred.
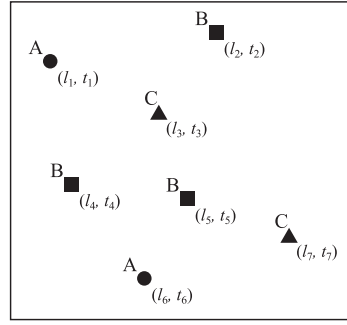
(a) Events belonging to three types: A (circles), B (squares), and C (triangles).

(b) Trajectories of three moving bodies, A, B, and C.

Fig. 1. Illustrations of event data and trajectory data types.

Similarly, a disease outbreak can be represented using the location and time where the patient was first infected. A collection of ST point events is called as a spatial point pattern (Gatrell et al. 1996) in the spatial statistics literature. Figure 1(a) shows an example of a spatial point pattern in a two-dimensional Euclidean coordinate system, where (l_i, t_i) denotes the location and time point of an event. Note that apart from the location and time information, every ST event may also contain non-ST variables, known as *marked variables*, that provide additional information about every ST event. For example, events can belong to different types, such as the type of disease or the nature of crime, which can be denoted by a categorical marked variable. In Figure 1(a), the marked variable over events can take three categorical values: A, B, and C. ST events are quite common in real-world applications such as criminology (incidence of crime and related events), epidemiology (disease outbreak events), transportation (road accidents), Earth science (land cover change events like forest fires and insect disease), and social media (Twitter activity or Google search requests).

While the spatial nature of most ST events can be represented using a Euclidean coordinate system (where every dimension is equally important), sometimes it is more relevant to explore alternative representations of these data. For example, accidents on freeways can be considered as events occurring on a spatial road network, where the distance between any two events is measured not by their Euclidean distance but by the shortest distance of the road segments connecting the events. Further, events may not always be point objects in space but instead be characterized by other geometric shapes such as lines and polygons. For example, a forest fire event can be represented as a spatial polygon that delineates the extent of damage due to the fire. Similarly, an event may not have an instantaneous time point of occurrence but instead be associated with a time period of appearance, denoting the birth and death of the event. For example, a music concert

(a) Reference points on time stamp 1.
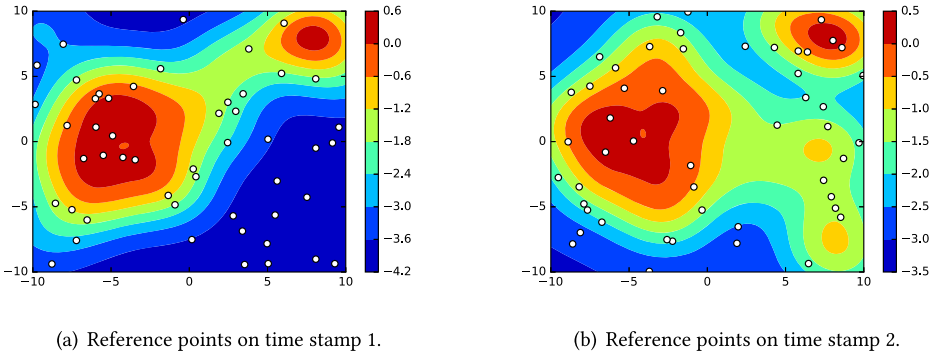
(b) Reference points on time stamp 2.

Fig. 2. An example of ST reference points (shown as dots) on two different time stamps. The colorbars show the distribution of the ST field on the two time stamps.

happening in the city can be represented using the start and end times of the event. While these simple extensions of ST events are quite common in real-world applications, most of the existing STDM methods are tailored for analyzing point ST events occurring in Euclidean spaces.

3.2.2 Trajectory Data. Trajectories denote the paths traced by bodies moving in space over time. Some examples of trajectories include the route taken by a taxi from the pick-up to the drop-off location or the migration patterns of animals traveling for better access to food, water, and shelter (Farine et al. 2016). Trajectory data are commonly collected by mounting sensors on the moving bodies that periodically transmit information about the location of the body over time. Figure 1(b) illustrates the trajectories of three moving bodies, *A*, *B*, and *C*. Apart from measuring the series of locations traversed by every moving body over time, trajectory data may also contain additional marked variables of the moving body, such as the heart-rate of a person running on a circuit track or the sequence of messages transmitted and received by a patrolling police car. Trajectories are common in applications such as transportation, ecology, and law enforcement.

3.2.3 Point Reference Data. Point reference data consist of measurements of a continuous ST field such as temperature, vegetation, or population over a set of moving *reference* points in space and time. For example, meteorological variables such as temperature and humidity are commonly measured using weather balloons floating in space, which continuously record weather observations at point locations in space and time. Another example is that of buoy sensors used for measuring ocean variables such as sea surface temperature at locations that keep changing over time. In both these examples, a finite sample of ST reference points is used to represent the behavior of a continuous spatio-temporal field. Figure 2 shows an example of the spatial distribution of a continuous ST field at two different time stamps that are being measured at reference locations (shown as dots) on the two time stamps. The observations at these discrete reference points can be used to reconstruct the ST field at any arbitrary location and time using data-driven methods (e.g., smoothing techniques) or physics-based methods (e.g., meteorological reanalysis approaches (Saha et al. 2010)). Point reference data are also known as geostatistical data in the spatial statistics literature.

3.2.4 Raster Data. In raster data, measurements of a continuous or discrete ST field are recorded at fixed locations in space and at fixed points in time. This is in contrast to point reference data where the ST reference sites may keep changing their location over time and collect recordings on different time stamps. To formally describe a raster, consider a set of fixed locations, $S = \{s_1, s_2, \ldots, s_m\}$, either distributed regularly in space with constant distance between adjacent locations, e.g., pixels in an image (see Figure 3(a)) or distributed in an irregular spatial pattern,
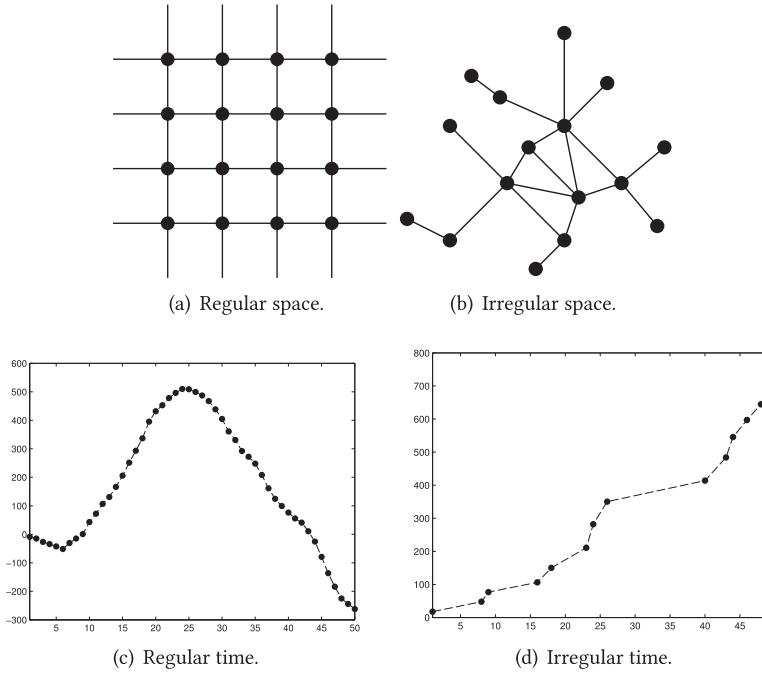
Fig. 3. Different aspects of ST grids used for representing a raster data. The set of locations in the ST raster can either be located regularly or irregularly in space. The set of time stamps can also be either spaced regularly or irregularly.

e.g., ground-based sensor networks (see Figure 3(b)). For every location, we record observations on a fixed set of time stamps, $\mathcal{T} = \{t_1, t_2, \ldots, t_n\}$, which can again be regularly spaced with equal delays between consecutive measurements (see Figure 3(c)) or irregularly spaced (see Figure 3(d)). It is the Cartesian product of $\mathcal{S}$ and $\mathcal{T}$ that results in the complete spatio-temporal grid, $\mathcal{S} \times \mathcal{T}$, where every vertex on the ST grid, (s_i, t_j), has a distinct measurement.

ST raster data are quite common in several real-world applications such as remote sensing, climate science, brain imaging, epidemiology, and demography. Some examples of ST raster data include measurements collected by ground-based sensors of ST fields such as air quality or weather information, geo-registered images of the Earth's surface collected by satellites at regular revisit times, and fMRI video sequences of brain activity. Note that while some examples of ST raster record observations at point vertices (e.g., measurements collected by a sensor network), others make aggregate measurements over the region at every grid cell. For example, demographic information is often collected at aggregate scales over political divisions such as cities, counties, districts, and states at annual scale. Another feature of an ST raster data is the resolution of the grid (both in space and in time) used for collecting measurements. In many applications, one commonly encounters ST raster datasets at varying resolutions of space and time, collected from different instruments or sensors. For example, satellite measurements of Earth's surface may be obtained via Landsat instruments at 30m spatial resolution every 16 days or via MODIS instruments at 500m spatial resolution on a daily scale. As another example, fMRI technology can be used to measure brain activity at each 1mm x 1mm x 1mm location, whereas EEG technology measures activity at a selected set of tens of locations. In problems involving ST rasters with varying resolutions, we often need to convert an ST raster from its native resolution to a finer or coarser resolution, so that a seamless analysis of all ST rasters can be performed at a common resolution.
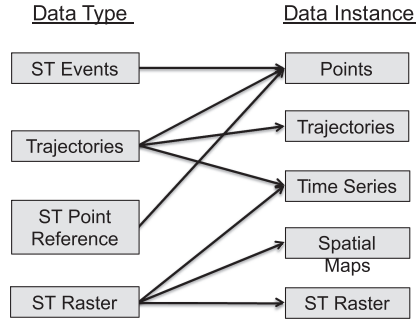
Fig. 4. A mapping showing the different categories of ST data instances that can be constructed from ST data types.

Interpolation techniques, also referred to as resampling methods in Geographic Information System (GIS) literature, are commonly used to convert a raster data to a finer resolution in space or time. Note that the computational requirements of STDM methods generally increase as we move to finer resolutions in space and time. Raster data can also be converted to a coarser resolution by aggregating over collections of ST cells. Aggregation generally helps in removing redundancies in the observations, especially when there is high spatial and temporal auto-correlation at finer resolutions. However, it is important to keep in mind that aggressive aggregation of data to coarser resolutions may result in loss of information about the ST field being measured.

   3.2.5   Converting Data Types. Even if ST data are naturally collected in a particular data type in a certain application, it is possible to *transform* it to a different ST type so that the relevant family of STDM tools are used for their analyses. We provide some examples of inter-conversion among ST data types in the following. An event data type can be converted to a raster data type by aggregating the counts of events at every cell of an ST grid. For example, crime events can be counted at the levels of counties in a city at an hourly scale, thus producing an ST raster of crime occurrences. In some cases, raster data can also be converted to ST events by using special algorithms for event extraction, e.g., techniques to find ST regions with abnormal activity. As an example, ecosystem disturbance events such as forest fires can be extracted from geo-registered satellite images of vegetation cover (Mithal et al. 2011a). Another common type of conversion among ST data types is between point reference data and raster data. Observations at ST reference points can be transformed into an ST raster format by interpolating or aggregating over an ST grid. Raster data can also be converted to a point reference data where every vertex of the ST grid is viewed as an ST reference point.

3.3   Data Instances

The basic unit of data that a data-mining algorithm operates on is called a *data instance.* In classical data-mining settings, a data instance is unambiguously represented as a set of observed features with optional supervised labels. However, in the context of ST data, there are multiple ways of defining instances for a given data type, each resulting in a different STDM formulation. In this section, we review five common categories of ST instances that one encounters in STDM problems, namely, points, trajectories, time series, spatial maps, and ST rasters. These data instances form the canonical building blocks of analysis for a broad range of problems and methods in STDM that we will discuss in Section 4.

   Figure 4 summarizes the different types of data instances that can be constructed for each ST data type. While ST events can be naturally represented as point instances, a trajectory data type

can either be represented as a collection of point instances (ordered list of locations visited by the moving object), a trajectory instance, or as a time series of spatial identifiers (e.g., location coordinates). A ST reference data type comprises point instances, where every instance is a *reference* point of the ST field in space and time. There are three different ways of constructing instances for ST raster data. First, we can define an instance to be the set of measurements at any location, s_i, represented as a time series, T_i. Second, we can also define an instance to be the set of measurements at any time stamp, t_j, represented as a spatial map, S_j. A third approach is to consider the entire ST raster (collection of observations over the entire ST grid) as a single data instance. It is the plurality of ways for defining instances for every ST data types that results in multiple STDM problems and methods. The choice of the right approach for constructing ST instances from a ST data type depends on the nature of the question being investigated and the family of available STDM methods that can be used.

In the following, we introduce the five categories of ST data instances and discuss some of the common questions that can be investigated for each category. Since ST data instances can be summarized using a number of ways that go beyond standard dimensionality reduction techniques such as principal component analysis, we also describe some of the common ways of representing and summarizing the features for every instance category.

3.3.1 Points. An ST point can be represented as a tuple containing the spatial and temporal information of a discrete observation, as well as any additional variables associated with the observation. ST points are frequently used as data instances in STDM analyses involving event data, e.g., point events such as the occurrence of crimes at certain locations and times can be treated as basic instances to group similar instances or to find anomalous instances. ST points are also used as data instances when dealing with point reference data, where the measurements at ST reference points are used as instances to estimate the ST field at unseen instances. Additionally, a trajectory can also be viewed as an ordered collection of ST points, which are the locations visited by the moving object.

Some of the common questions that can be asked using ST points as data instances include the follows: *How are ST points clustered in space and time? What are the frequently occurring patterns of ST points? Can we identify ST points that do not follow the general behavior of other ST points? Can we estimate a target variable of interest at an ST point that has not been seen during training?*

A collection of ST points can be summarized by measures that capture the strength of interaction (or auto-correlation) among the points. For example, the Ripley's *K* function (Dixon 2002) is a commonly used statistic for describing the amount of attraction or repulsion among spatial locations beyond its expected value. Spatio-temporal extensions of the Ripley's *K* function have also been explored to measure the strength of interaction among ST point events (Lynch and Moorcroft 2008). The strength of auto-correlation among ST reference points can also be measured using the Moran's I function (Li et al. 2007), local Moran's I (Anselin 1995), and its ST extension (Hardisty and Klippel 2010).

3.3.2 Trajectories. Trajectories are a different class of data instances that can be used in STDM analyses involving moving bodies. Trajectories can be represented as multi-dimensional sequences that contain a temporally ordered list of locations visited by the moving object, along with any other information recorded by the object.

Some of the common questions that can be asked using trajectories as data instances include the following: *Can we cluster a collection of trajectories into a small set of representative groups? Are there frequent sequences of locations within the trajectories that are traversed by multiple moving bodies?*

A common approach for representing and extracting features from trajectories is by the use of generative models (Gaffney and Smyth 1999), where a parametric model is used to approximate the behavior of every trajectory. The learned parameters can then used as succinct representations of the trajectories. A trajectory can also be represented using the frequent sub-sequences of locations that are visited by the moving body. Techniques for identifying frequent trajectory patterns are discussed in detail later in Section 4.3.3. Distributed and efficient indexing structures for answering trajectory-related similarity queries have been developed in Zeinalipour-Yazti et al. (2006) and Al-Naymat et al. (2007). Apart from these methods, different schemes such as semantic trajectories (Bogorny et al. 2014; Li 2017), symbolic trajectories (Güting et al. 2015), and spatio-textual trajectories (Damiani 2016) are also recently being explored for representing trajectory data.

3.3.3 Time Series. Time series can be used as data instances in two different scenarios involving ST data. First, given an ST raster data, we can consider the set of observations at every spatial cell in the ST grid as a time series that can be used as a data instance in an STDM analysis. Second, trajectory data can also be treated as a multi-dimensional time-series data, where the multiple dimensions correspond to the spatial identifiers (e.g., location coordinates) traversed by the moving objects over time and any other variables recorded by the moving object in the course of its trajectory. While representing spatial identifiers as multiple (and independent) dimensions of a time series may not preserve the spatial context among the identifiers, it opens up the vast literature on time-series data mining that can be used for analyzing trajectories in novel ways, as will be described in Section 4.

Some of the common questions that can be asked using time series as data instances include the following: *Can we identify groups of time series that show similar temporal activity and are located nearby in space? Are there some temporal patterns that commonly repeat in a number of time series? Can we identify time stamps where the time series deviate from their normal behavior for a short period of time? Can we discover time stamps where the time series show a change in its profile? Can we use time series as input features to predict a target variable? Can we predict the value of a time series at a future time stamp using its historical values? Can we find distant groups of spatially contiguous time series that are related to each other?*

A number of approaches exist for extracting useful features from a collection of time series. This include methods that can identify temporally frequent sub-sequences that occur in a majority of time series (e.g., temporal motifs (Mueen 2014)) or contain discriminatory information about a particular time-series class (e.g., shapelets (Hills et al. 2014; Ye and Keogh 2009)). A review of techniques for identifying such frequent patterns in time series is presented in Fu (2011) and Esling and Agon (2012).

3.3.4 Spatial Maps. ST raster data can be viewed as a collection of spatial maps observed at every time stamp, which can also be used as data instance in analyses involving ST rasters.

Some of the common questions that can be asked using spatial maps as data instances in STDM analyses include the following: *Can we cluster the spatial maps to find groups of time stamps showing similar spatial activity? Can we identify spatial patterns that are observed in a number of spatial maps? Can we use spatial maps as input variables to predict a target variable? Can we predict the value at a certain location using observed values at other locations in the map?*

A common approach for extracting features among spatial maps is using image segmentation techniques (Haralick and Shapiro 1985). The presence or absence of different types of image segments can then be used as features to represent spatial maps.

3.3.5 ST Rasters. ST raster data in their entirety, with measurements spanning the entire set of locations and time stamps, can also be treated as individual data instances in STDM analyses.

Some of the common questions that can be asked using ST rasters as data instances include the following: *Can we cluster ST rasters into groups that show similar behavior in space and time? Can we find frequent spatio-temporal behavior that occurs in a number of ST raster datasets? Can we find ST rasters that show distinctly different behavior than other ST rasters? Can we find timestamps where the behavior of an ST raster changes over time? Can we use ST rasters as input features for predicting a target variable of interest? Can we predict the value at a certain location and time stamp in the ST grid using observed values at other locations and time stamps? Can we find subsets of locations in the ST grid that show interesting relationships in their temporal activity?*

A basic approach for representing ST rasters is using N-way arrays also called as *tensors*. In a tensor representation of an ST raster data, some dimensions are used to represent the set of locations while the remaining dimension is used to represent the set of time stamps available in the ST grid. For example, precipitation data are represented as a three-dimensional (3D) array where the first two dimensions capture 2D space and the third dimension captures time. Similarly, fMRI data are represented as a 4D array where the first three dimensions capture 3D space and the third dimension captures time. A tensor representation of ST raster data can then be summarized using space–time subspaces that have similar values, which are the equivalent of image segmentation in ST domains. Existing image segmentation techniques that have been extended to work with videos to discover moving objects are relevant to address this problem (Haritaoglu et al. 2000; Prati et al. 2003).

One way to summarize or extract features from ST rasters is by using network-based representations, where the nodes correspond to the locations and the edges denote the similarity among the time series at locations (Atluri et al. 2016; Feldhoff et al. 2015). Techniques for computing similarity among time series are discussed in detail later in Section 3.4.3. The topological properties of nodes such as degree and different variants of centrality measures can then be used to characterize the "role" and "influence" of locations. Such properties have been found to be useful in characterizing nodes in different types of networks such as social, biological, and transportation networks. For example, the use of network-based properties such as coherence among a set of locations (De Martino et al. 2007; Sui et al. 2009) or relationship between distant locations (Lynall et al. 2010; Pettersson-Yeo et al. 2011) has been explored in previous studies.

3.4 Similarity Among Instances

Determining similarity (or dissimilarity) between data instances is key to many data-mining problems such as clustering, classification, pattern discovery, and relationship mining. Similarity measures have been extensively studied for data instances in traditional data-mining settings that are able to deal with varying types and formats of attributes. However, in the presence of space and time, there is a variety of ways we can define similarity among ST data instances such as points, trajectories, time series, spatial maps, and ST rasters. Our ability to capture similarities among ST data instances using proximity measures underpins the effectiveness of STDM methods that build on them. In the following, we describe some of the common ways for computing similarity among different types of ST data instances.

3.4.1 Point Similarity. Two points are considered close if they lie within the ST neighborhoods of each other. The ST neighborhood of a point can be defined using a fixed distance threshold in space and time, e.g., within a 1km radius and 1 hour time difference. Alternatively, the ST neighborhood of every point can also be defined in terms of a fixed number, k, of closest points. The choice of the right notion of locality depends on the application context and can be decided by the domain analyst.

3.4.2 Trajectory Similarity. Similarity among trajectories is often measured in terms of the co-location frequency, which is the number of times two moving bodies appear spatially close to one another. Other approaches for measuring similarity among trajectories include subsequence similarity metrics such as the length of the longest common subsequence, Fréchet distance, dynamic time warping (DTW), and edit distance (Toohey and Duckham 2015). Trajectory similarity can also be computed using feature-based representations such as the frequent trajectory patterns extracted from the data.

3.4.3 Time-Series Similarity. If we consider every time series as a 1D array of observations, then the similarity among two time series can be simply computed using proximity measures such as the Euclidean distance and the correlation strength, which consider a one-to-one correspondence between the elements of the two arrays. However, sometimes it is the case that two similar time series are not exactly aligned with one another but show the same pattern of activity over time. Measures such as DTW (Keogh and Ratanamahatana 2005) and Fréchet distance (Alt and Godau 1995) are able to capture such forms of similarity among time series. We can also compute the similarity among two time series based on their closeness in time-series features such as temporal motifs and shapelets. When it is expected to observe a certain delay or time lag between the observations of two time series, a common approach is to translate one of the time series with a range of candidate values of time lag and then choose the time lag that provides maximum similarity (e.g., highest absolute correlation).

While most measures of time-series similarity consider the entire time duration into account, it is possible that the similarity structure among time series is subject to variation over time. For example, when a subject's mental activity is switching between planning their day (i.e., an executive task) and reminding themselves of a recent meeting (i.e., a memory task), the similarity structure among the time series at brain regions could be different. In such cases, a desired pattern of similarity among time series may only be exhibited for short periods of time, which need to be determined from the data. An example of an approach that simultaneously identifies the relevant time windows for computing time-series similarity and uses this metric to cluster the time series can be found in Atluri et al. (2014).

3.4.4 Spatial Map Similarity. Two spatial maps can be considered similar if they show similar values at corresponding locations, which can be captured using standard proximity measures such as Euclidean distance. However, spatial maps can often suffer from small misalignments due to geo-registration errors, which can result in misleading distance metrics. Further, it is often more useful to compute similarity over smaller sub-regions in the map that contain foreground objects than considering the similarity over the entire map. The Earth's mover distance (Rubner et al. 2000) is one such metric that is robust to changes in the alignment of images, which is based on the minimal cost that must be paid to transform one image into the other. Spatial map similarity can also be computed on the basis of features such as image segments extracted from the data.

3.4.5 ST Raster Similarity. Network-based representations of ST rasters can be used to assess if two rasters are similar or not. For example, link and node similarity scores (Berg and Lässig 2006; Li and Yang 2009) can be used to compute the similarity among the corresponding links and nodes of ST rasters. ST raster similarity can also be computed on the basis of features extracted from their network representations (Soundarajan et al. 2013).

4 PROBLEMS AND METHODS

4.1 Clustering

Clustering refers to the grouping of instances in a dataset that share similar feature values. When clustering ST data, novel challenges arise due to the spatial and temporal aspects of different types

of ST instances. For example, clustering locations based on their time series of observations in ST raster data has to ensure that the discovered clusters are spatially contiguous. Ignoring this spatial information can lead to fragmented clusters of locations that are difficult to interpret and have salt-and-pepper errors. Since many clustering algorithms use similarity measures for identifying groups of similar instances, techniques for computing similarity among ST instances (described in Section 3.4) will come in handy while clustering ST data. In the following, we describe some of the common methods for clustering different types of ST data instances.

4.1.1 Clustering Points. There are two different objectives that are of interest when clustering ST points. One objective involves finding clusters that have an unusually high density of ST points, also termed as *hot-spots*. This can be used for finding outbreaks of diseases or social movements, where there is a dense conglomeration of events both in space and time (Gomide et al. 2011; Kulldorff et al. 2005). This problem is also referred to as "event detection" in the literature on mining social media data (Feng et al. 2015). The problem of finding hot-spots has been studied in the spatial statistics literature using methods such as the spatial scan statistic (Kulldorff 1997), which explores all possible circular shaped regions of different sizes to determine regions where the incidence of points is significantly higher than expected. Generalization of the scan statistic for ST data, termed as space–time scan statistics, have been explored for studying disease outbreaks, crime hot-spots, and events in twitter data (Cheng and Wicks 2014; Kulldorff 2001; Kulldorff et al. 2005). While these early works provide a promising framework, open questions pertaining to shapes of regions (Eftelioglu et al. 2016; Takahashi et al. 2008), background distribution (Tango et al. 2011), and speed of search (Agarwal et al. 2006) are being investigated. Approaches for clustering spatial objects such as CLARANS (Ng and Han 2002) can also be extended in the ST domain for finding groups of objects that are related in both space and time.

The second objective involves finding clusters of ST points that also have similar non-ST attributes. For example, given a collection of crime events, we may be interested in finding regions in space and time with similar crime activities. This clustering objective has been studied in the context of crime data (Eftelioglu et al. 2014), twitter data (Abdelhaq et al. 2013; Chierichetti et al. 2014; Ihler et al. 2006; Walther and Kaisser 2013; Weng and Lee 2011), geo-tagged photos (Zheng et al. 2012), traffic accidents (Zheng et al. 2012), and epidemiological data (Glatman-Freedman et al. 2016). A number of techniques for clustering ST points are based on the DBSCAN algorithm (Ester et al. 1996), which is a widely used method for finding arbitrarily shaped clusters of spatial points based on the density of points. ST-DBSCAN (Birant and Kut 2007) is one of the popular extensions of DBSCAN that defines two separate distances between ST points: one that captures spatial attributes and another that captures temporal and non-ST attributes. Computing distance based on spatial, temporal and non-ST attributes separately and having separate thresholds for them provides the user a flexibility to determine the desired spatial density that is relevant to the problem at hand. However, several challenges including heterogeneity in space and time, varying densities of clusters, and sampling bias inherent in the data are yet to be addressed.

4.1.2 Clustering Trajectories. When clustering trajectory data, we are often interested in finding groups of trajectories that are similar to each other across the entire duration of trajectories. For example, clustering of trajectories has been used to find groups of hurricanes with similar trajectories (Lee et al. 2007), groups of taxi traces that follow a similar route (Liu et al. 2010), or groups of moving objects that follow the same motion in video streams (Gaffney and Smyth 1999). A review of techniques for clustering trajectories is presented in Kisilevich et al. (2009). Two important aspects of trajectory clustering methods are the choice of distance measure (see Section 3.4.2) and the choice of clustering technique (Morris and Trivedi 2009). One category of methods for clustering trajectories involves using mixture modeling approaches (Alon et al. 2003; Chudova et al.
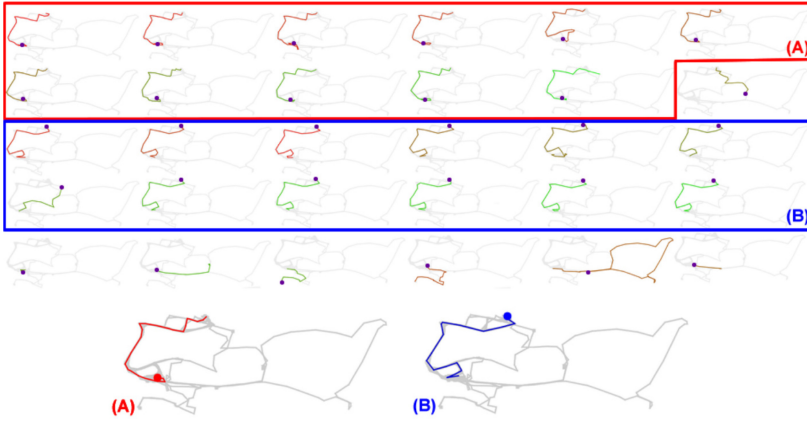
Fig. 5. Mobility profiles discoverd from a user's GPS traces based on clustering of trips. (Figure taken from Trasarti et al. (2011).)

2003; Gaffney and Smyth 1999), e.g., mixtures of regression models where a different regression model is learned for every cluster of trajectories (Gaffney and Smyth 1999). A different approach has been explored in Trasarti et al. (2011), where a two-step clustering method was proposed to find mobility profiles of users based on their GPS traces. Figure 5 shows two sets of discovered profiles (A and B), along with several noisy trips that do not conform to these profiles.

In some trajectory clustering problems, we are interested in finding groups of trajectories that share similarity in only a short duration of the trajectory. One of the methods developed for this problem is a partition-and-group framework called TRACLUS (Lee et al. 2007), which first partitions each trajectory into smaller line segments based on a minimum description length principle and then groups line segments based on their similarity using a DBSCAN-based approach.

A related area of research is the identification of "moving clusters" of trajectories, where moving bodies may join or leave a cluster as it progresses in space over time. This is a common pattern that is observed in several applications, e.g., migrating flocks of animals or convoys of cars. Algorithms for detecting moving clusters of trajectories have been developed in Kalnis et al. (2005) and in subsequent studies (Dodge et al. 2008; Jeung et al. 2008; Li et al. 2010). More recently, Zhang et al. (2016) proposed an iterative framework, GMove, that alternates between two tasks: assigning users to groups and modeling group-level mobility, using an ensemble of hidden Markov models.

4.1.3 Clustering Time Series. A central objective when clustering time series derived from ST raster data (using the series of measurements at every location) is to find spatially coherent groups of locations with similar temporal activity. This problem has been approached from multiple directions. One direction is to use traditional clustering schemes such as k-means clustering (Mezer et al. 2009), hierarchical clustering (Goutte et al. 1999), shared nearest neighbor (SNN) clustering (Steinbach et al. 2003), and normalized-cut spectral clustering (Van Den Heuvel et al. 2008) to cluster time series (using the time-series similarity measures discussed in Section 3.4.3). The clustering of Sea Level Pressure time series, studied in climate science, from all locations on the Earth's surface is shown in Figure 6. Data from 1982 to 1993 was used to find these clusters and Pearson's correlation coefficient was used to assess similarity between time series.

Since traditional approaches for time-series clustering do not guarantee that the resultant clusters are spatially contiguous, this is typically addressed in a post-processing step either by increasing the number of clusters (Smith et al. 2009) or by separating clusters with non-contiguous sets
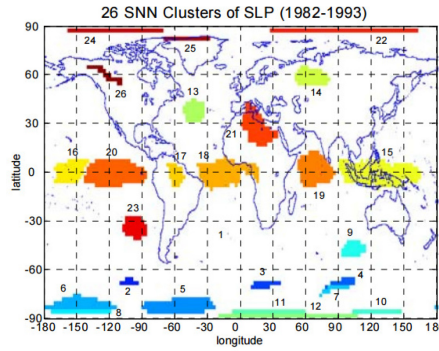
Fig. 6. Clusters found using SNN clustering of Sea Level Pressure data (1982–1993). (Figure taken from Steinbach et al. (2002).)

into multiple contiguous clusters. Another direction is to directly incorporate spatial contiguity in the clustering process, e.g., by using "region growing" approaches or by enforcing contiguity constraints in the clustering technique. Region growing approaches (Bellec et al. 2006; Heller et al. 2006; Lu et al. 2003b) work by merging spatially adjacent locations that are highly similar to each other (defined using similarity measures discussed in Section 3.4.3) into a single cluster until a minimum number of clusters has been achieved or no cluster can be further grown without violating a similarity criterion. While region growing approaches ensure that every cluster has similar locations, they do not ensure that the different clusters are dissimilar. On the other hand, clustering approaches that utilize contiguity constraints (Blumensath et al. 2013; Craddock et al. 2012) do not have this problem as these approaches ensure that locations within a cluster are highly similar to each other than locations that are in two different clusters.

4.1.4 *Clustering Spatial Maps.* When clustering spatial maps (sets of measurements from all locations in an ST raster on different times), the central objective is to find groups of time stamps that have similar spatial maps, e.g., time stamps with similar maps of brain activity (Liu and Duyn 2013). Similarity measures among spatial maps, such as those discussed in Section 3.4.4, can help in discovering meaningful clusters of spatial maps irrespective of data artifacts such as changes in alignment and registration among the locations of different maps (Liu et al. 2013). While these methods are useful in producing temporally contiguous segments with similar spatial activity (due to the temporal auto-correlation in the data), in some cases, we may be interested in identifying non-contiguous groups of time stamps. One way to ensure that a given cluster of spatial maps is not due to temporal auto-correlation is to evaluate this in a post-processing step or use a smaller k such that distant time points are grouped into clusters.

4.1.5 *Finding Dynamic ST Clusters.* A common clustering problem that is of interest when dealing with ST raster data is to identify sub-regions of space and time that show coherent measurements, termed "dynamic ST clusters" (Chen et al. 2015). The discovery of dynamic ST clusters can help in detecting phenomena that only influence a subset of locations during a subset of time points. For example, water bodies that grow and shrink in space over time, e.g., lakes and reservoirs, can be identified as dynamic ST clusters in remote sensing data, as they appear as coherent observations in subsets of space and time. Note that a dynamic ST cluster may evolve over time and thus change its shape, size, and appearance as we progress in time. Hence, while some locations are retained across consecutive time stamps, the cluster assignments of locations are dynamic and the clusters can grow and shrink over time. A recent work (Chen et al. 2015) explored a novel

approach to identify dynamic ST clusters in raster datasets for detecting surface water dynamics, which used an iterative algorithm to first identify the set of "core" locations that are part of a dynamic cluster across all time stamps, and then growing around the core locations at every time stamps to capture the dynamic behavior occurring at the boundaries. Similar formulations can be developed for identifying other types of dynamic clusters in ST raster data, e.g., a moving cluster of locations such as the evolution of a hurricane.

4.1.6 Clustering ST Rasters. Given a collection of ST raster datasets, possibly collected over different spatial regions and time periods, it is useful to identify groups of ST raster datasets considered as individual instances. This has applications in several domains dealing with ST raster data such as climate science and neuroimaging. For example, in climate science, different climate models produce simulations of global climate variables as ST rasters, which when clustered sheds light on the similarity of climate models (Steinhaeuser and Tsonis 2014). ST rasters can be clustered by using similarity measures among ST rasters as basic building blocks, which typically involve extracting features from network representations of ST rasters (discussed in Section 3.4.5). For example, Yu et al. (2015b) found groups of ST rasters by first constructing networks from ST rasters and then grouping the resultant networks using a network module-finding approach from Newman (2006). However, extracting such high higher-order features from ST rasters is non-trivial and often requires domain expertise.

4.2 Predictive Learning

The basic objective of predictive learning methods is to learn a mapping from the input features (also called as independent variables) to the output variables (also called as dependent variables) using a representative training set. In spatio-temporal applications, both the input and output variables can belong to different types of ST data instances, thus resulting in a variety of predictive learning problem formulations. In the following, we discuss some of the predictive learning problems that are commonly encountered in ST applications based on the type of ST data instance used as input variables.

4.2.1 Time Series. Given an ST raster data, we can consider the problem of predicting an output variable (continuous or categorical) at every location of the raster using the time series at every location as input variables. The use of time series as input features is quite common in a number of classification and regression problems. For example, the temporal dynamics of audio frequencies is used to classify words or sentences in human speech recognition problems. This can be achieved by using recurrent neural networks (Graves and Schmidhuber 2009; Mikolov et al. 2010), which are extensions of artificial neural networks with appropriate skip connections among the neural nodes to model information delay. Another approach to incorporate the temporal context of time-series features in classification problems is the use of shapelets (Hills et al. 2014; Ye and Keogh 2009). Shapelets are time-series subsequences that are discriminative in nature, i.e., their occurrence is selective to certain classes. While these techniques provide an ability to model the temporal characteristics of input features, there is a need to develop novel methods that can take into account the spatial information among the time series in ST rasters. For example, instead of predicting the class label at every location using its time series independently, we can leverage information about spatial neighborhoods to enforce spatial consistence among the labels at nearby locations. Variants of recurrent neural networks that include spatial features for spatio-temporal prediction have been explored in Jia et al. (2017a, 2017b) and Jain et al. (2016). Latent space models that use topological as well as temporal attributes of locations have also been developed for real-time traffic prediction using time-varying information from sensor recordings (Deng et al. 2016). Time series can also be constructed from trajectory data, where the objective is to

predict the future location of a moving object (or a group of objects), given their past history of visited locations (Horton et al. 2017; Li et al. 2016).

4.2.2   Spatial Maps. In this class of predictive problems, a scalar output variable has to be predicted at a time step of the ST raster, using the spatial map at the same time step as input variables. Some examples of applications that use spatial maps as input features include image classification and object recognition problems, where a categorical value has to be assigned to every image or sub-region in an image using the information contained in spatial maps. A classification approach that has recently gained widespread recognition in the computer vision community is *deep* convolutional neural networks (CNN) (Krizhevsky et al. 2012; LeCun and Bengio 1995), which uses the spatial nature of inputs to share model parameters and provide robust generalization performance. In spatio-temporal applications, a promising research direction is to use the temporal auto-correlation among consecutive spatial maps to share the parameters of CNN models over time. Spatio-temporal extensions of CNN-based learning frameworks have been developed in Taylor et al. (2010) and Karpathy et al. (2014). Extensions of neural networks that use both the spatial and temporal information of data have also been developed in Stiles and Ghosh (1997) and Ghosh and Deuser (1995), where the network design was inspired by the biological information of habituation mechanisms in neuroscience.

4.2.3   ST Rasters. Another class of predictive learning problems is to use the entire information in an ST raster as input variables to predict a scalar output variable. An example of this is to predict if a subject has a mental disorder or note based on their fMRI scan, stored as an ST raster. This has tremendous applications in diagnosing mental disorders, which is currently done in a very subjective manner. Another application that is increasingly becoming popular in the realm of brain imaging is that of "brain reading" (Norman et al. 2006), where the objective is to determine the nature of activity (e.g., planning, memorizing, recollecting, etc.) based on measured spatio-temporal activity from the brain, represented as an ST raster.

A naïve approach to this problem could involve representing every ST raster with $|\mathcal{S}|$ locations and $|\mathcal{T}|$ time stamps as a vector of size $|\mathcal{S}| \times |\mathcal{T}|$ and employing traditional classification schemes on these vector representations, e.g., linear discriminant analysis and support vector machines (Ku et al. 2008). Note that this is only possible when the ST grid of every grid is perfectly aligned with each other, such that their sets of locations and time stamps are identical. This is not usually the case with resting state fMRI scans, because the time points in one subject's scan cannot be matched with those of another subject. Furthermore, especially in fMRI data, the large number of spatial locations (typically hundreds of thousands) and time points (typically hundreds) leads to millions of potential features, where the number of instances are often tens to hundreds, leading to the phenomena of model overfitting. Hence, it is important to use *derived* features from ST rasters that summarize the ST activity of every raster, as described in Section 3.3.5. Alternatively, tensor learning–based approaches (Bahadori et al. 2014; Yu et al. 2015a; Zhou et al. 2013) provide a way to reduce the model complexity by making use of the spatial and temporal dependencies among the input features, thus showing a promise in predictive learning problems involving input ST rasters.

4.2.4   ST Reference Points. A common predictive learning problem in spatio-temporal applications is to predict the response at a certain location and time using observations collected at other locations and time stamps (often in ST neighborhoods). This is important in a number of domains, e.g., while estimating an ecological variable over every location and time using remote sensing observations at nearby locations and time stamps. The problem of land cover classification (estimating a categorical label at every location and time indicating its propensity to belong to a land cover type) has been heavily studied in the remote sensing literature for a variety of problems

(DeFries and Chan 2000; Jun and Ghosh 2011; Li et al. 2014b; Vatsavai 2008), e.g., the mapping of surface water dynamics using multi-spectral remote sensing data (Karpatne et al. 2016b; Khandelwal et al. 2017). As another example, the outbreak of influenza at a given location and time can be predicted based on web searches (Ginsberg et al. 2009) and twitter messages (Culotta 2010) at neighboring locations and times. There are two classes of methods that are relevant for making predictions at ST reference points: methods that use the temporal information to predict values at nearby time points and methods that use the spatial information to estimate values at nearby spatial points. We discuss both these classes of methods in the following.

Using Temporal Information: In many domains such as climate and health, estimation (or forecasting) of the future conditions based on present and past conditions is desired. For example, the sea surface pressure and temperature for the present month can be predicted based on values in the previous months. Similar problems are studied in predicting closing stock prices at the New York Stock Exchange and in predicting sales in the manufacturing industry (Montgomery et al. 2015). Some of the widely used methods for time-series forecasting problems include exponential smoothing techniques (Gardner 2006), ARIMA models (Box and Jenkins 1976), and state-space models (Aoki 2013). Another type of methods for making predictions at time stamps include dynamic Bayesian networks such as hidden Markov models and Kalman filters (Harvey 1990; Rabiner and Juang 1986), which estimate the most likely sequence of latent values at time stamps using the temporal auto-correlation structure. Techniques that make predictions in time need to be modified to include the spatial context in spatio-temporal applications. As an example spatio-temporal Granger causality models, that use both spatial and temporal information in the regression models, have been explored to model relationships among ST variables (Lozano et al. 2009b; Luo et al. 2013).

Using Spatial Information: There is a vast body of literature on spatial prediction methods that take into account the spatial auto-correlation structure in the data to ensure spatially coherent results. This includes the use of spatial auto-regressive (SAR) models (Kelejian and Prucha 1999), geographically weighted regression (GWR) models (Brunsdon et al. 1998), and Kriging (Oliver and Webster 1990) in the spatial statistics literature. Markov random field–based approaches that are naturally suited to handle the spatial auto-correlation in the data have also been widely studied (Kasetkasem and Varshney 2002; Schroder et al. 1998; Zhao et al. 2007). There is promise in informing such techniques with the temporal nature of ST points in spatio-temporal applications, such that both spatial and temporal auto-correlation are incorporated in the modeling framework. For example, spatio-temporal Kriging approaches have been used in a number of applications in climate and environmental modeling (Cressie and Wikle 2015), where time is treated as another dimension while learning covariance structures in space and time. A related area of research is spatial item recommendation (Wang et al. 2017), where the preference over spatial items (e.g., restaurants or tourist attractions) have to be predicted in a time-varying manner, using social network information and the history of preferences of every user.

4.3 Frequent Pattern Mining

Frequent pattern mining is the process of discovering patterns in a dataset that occur frequently over multiple instances in a dataset, e.g., frequently bought groups of items in market-basket transactions. Given the rich variety of data types and instances in ST applications, there are several categories of frequent pattern-mining problems that can be formulated in the presence of spatial and temporal components, as described in the following.

4.3.1 Co-occurrence Patterns in ST Points. Given a collection of points such as ST events of varying types, a common pattern of interest is co-occurrence patterns, which are basically subsets

of ST event types that occur in close spatial and temporal proximity of each other. For example, given a dataset of crime and other events in a city, we may be interested in finding ST event types that occur together (e.g., bar closing and drunk driving). Co-occurrence patterns has been studied in the context of spatial datasets as "co-location" patterns for more than a decade (Huang et al. 2004; Wang et al. 2013; Xiao et al. 2008). In the spatial statistics literature, co-locations have been studied using measures such as the Ripley's cross-K function (Dixon 2002), which capture statistical patterns of attraction or repulsion among pairs of spatial point types.

One of the first approaches for discovering co-occurrence patterns includes an Apriori-principle-based approach developed by Pillai et al. (2012). This approach relies on an interestingness measured termed as *spatio-temporal co-occurrence co-efficient* to capture co-occurrence in spatial and temporal dimensions. A filter and refine approach was later proposed in Pillai et al. (2013) to tackle the problem of discovering co-occurrence patterns in large event databases. Constructing rules from co-occurrence patterns have also been studied in Pillai et al. (2014). Spatiotemporal extensions of the cross-K function have been proposed in Lynch and Moorcroft (2008), where the ST neighborhood of ST points is used for identifying attraction or repulsion patterns among ST point types in both space and time.

4.3.2 Sequential Patterns in ST Points. Sequential patterns have been studied in the context of ST event data, where the occurrence of ST events of a particular type can trigger a sequence of ST events of other types. For example, a car accident on a freeway could trigger a traffic jam in its ST neighborhood. Sequential patterns have been originally defined in the context of market-basket transactions where sequences of transactions from every customer are available and the goal is to discover ordered list of items appearing with high frequency (Agrawal and Srikant 1995). An approach for discovering sequential patterns of ST event types was presented in Huang et al. (2008). This approach was able to discover ordered lists of event types such as $f_1 \rightarrow f_2 \rightarrow \cdots \rightarrow f_n$, where events belonging to type f_1 trigger events of type f_2, that further triggers events of type f_3 and a series of events resulting in events of type f_k. They developed an Slicing-STS-Miner approach to efficiently discover statistically significant sequential patterns of events. A partially ordered subsets of event types, referred to as cascading spatial temporal patterns (Mohan et al. 2010, 2012), have also been studied to capture event sequences whose instance are located together and occur in successive stages. Some of the key challenges in mining ST sequential patterns include defining interesting measures that capture meaningful non-spurious patterns and developing efficient approaches to discover interesting patterns from an exponentially large space of candidate patterns.

4.3.3 Sequential Patterns in Trajectories. Sequential patterns in trajectories are sequences of spatial locations or regions that are visited by multiple moving objects in the same order and are thus covered by multiple trajectory instances. For example, given the trajectories of tourists visiting a city, we can find sequential patterns that represent the typical behavior of tourists as they visit an attraction and move to another. Although trajectories can be treated as multi-dimensional sequences where the spatial identifiers (e.g., latitude and longitude) of visited locations are used as dimensions, traditional sequential pattern-mining approaches are unsuitable for finding sequential patterns of trajectories because of their inability to handle the spatial nature of information contained in trajectories. For example, the locations visited by moving objects may not exactly match but be in close spatial vicinity to be considered as part of the same pattern. Moreover, trajectories belonging to a common sequential pattern are expected to follow similar spatial path but not exactly the same path. To handle the spatial aspect of trajectories in the process of mining sequential patterns, Tsoukatos and Gunopulos (2001) used ordered sequences of spatial elements such as rectangular regions to define sequential patterns. A trajectory is then considered to be part of a sequential pattern if every line segment joining consecutive locations of the
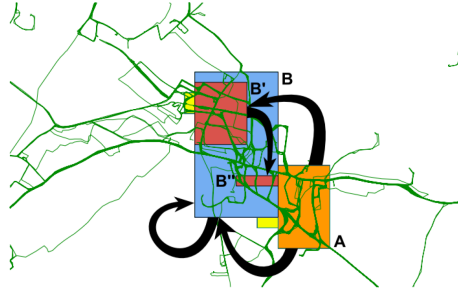
Fig. 7. Example of a sequential patterns found in GPS traces of trucks in Athens, Greece. (Figure taken from Giannotti et al. (2007).)

trajectory are enclosed within the rectangular regions of the pattern. A similar approach for sequential pattern mining was developed in Cao et al. (2005), where pattern elements were defined as spatial regions around frequent line segments of trajectories. Spatio-temporal extensions of association rules, termed as spatio-temporal association rules (Verhein and Chawla 2006), have also been explored to find frequent patterns of geographic regions visited by moving objects (Verhein and Chawla 2008). Approaches for mining periodic patterns in trajectories, where moving objects periodically visit sequences of locations in time, have been developed in Li and Han (2014). Approaches to handle uncertainty have also been studied to handle the incompleteness of the data and artificial noise introduced in privacy-sensitive applications (Chen et al. 2016; Li et al. 2013).

Another aspect of trajectories that is important to consider while mining sequential patterns in some applications is the presence of timing constraints among visited locations. For example, we may be interested in finding sequential patterns where the time duration between consecutive locations are similar across the trajectory instances, in addition to the similarity of paths followed by the moving objects (Giannotti et al. 2007). This approach was used to discover sequential patterns in GPS traces of trucks in Athens, Greece. One resultant pattern $A \xrightarrow{\Delta t1} B' \xrightarrow{\Delta t2} B''$ is shown in Figure 7 that captures how the trucks visit region A first, then visit region B', and then visit region B'', with fixed time gaps between the locations. Another way of introducing timing constraints is to consider sequential patterns of trajectories that are in close temporal proximity of each other, i.e., the moving objects visit a sequence of locations at similar points in time. Such patterns are referred to as "flock patterns" (Vieira et al. 2009) or "convoys" (Jeung et al. 2008) that are used for finding groups of objects moving together.

4.3.4 Motif Patterns in Time Series. In the context of ST raster data, time-series motifs represent repeated temporal measurements observed across multiple spatial locations. For example, in vegetation time-series data of agricultural farms, the harvesting cycles of crops can be discovered as time-series motifs that recur at multiple farm locations belonging to the same crop type. Time-series motifs have been extensively studied in applications such as ecology, medicine, finance, and the production industry (Mueen 2014; Torkamani and Lohweg 2017). While the naïve approach for discovering motifs is quadratic in the length of the time series, approximate approaches (Chiu et al. 2003; Meng et al. 2008; Minnen et al. 2007) have also been designed. Recently, the development of efficient approaches for constructing "matrix profile" (Yeh et al. 2016), a vector comprised of minimum non-trivial distance for each subsequence in time series A with subsequences in time series B, has enabled an efficient approach for motif discovery (Zhu et al. 2016). At the heart of this approach for computing the matrix profile is a fast Fourier transform–based algorithm for efficiently computing z-normalized Euclidean distance between two time-series subsequences. Using

this approach, subsequence similarity in any given time series can be efficiently computed, and the subsequences that are highly similar correspond to motifs.

While traditional formulations for motif discovery assume independence among the time series, in ST applications, there is a strong spatial auto-correlation among the time series at nearby locations, limiting the usefulness of existing formulations. This is because any time-series motif observed over a set of locations may also show up in the neighborhood of those locations with minor variations, resulting in a number of redundant patterns that look almost similar. One approach to address this challenge is to enforce motif discovery algorithms to identify recurring temporal patterns that appear synchronously in spatially coherent regions, in contrast to isolated points in space. This would result in the discovery of physically relevant processes and events that span both space and time, e.g., droughts or floods affecting an ST region in climate datasets. We can call such ST motifs "spatially coherent time-series motifs." It may also be useful to explore a slightly modified version of this problem where locations belonging to the same region do not synchronously show a common pattern but instead there is a gradual evolution of temporal activity across them. Such "evolving ST motifs" can help us detect frequent ST phenomena that gradually move across locations in a region.

An active research direction is to discover structural patterns in ST data that elucidate complex spatial and temporal dynamics. For example, Zhou and Matteson (2015) developed approaches to discover ST dynamics in ambulance demand data for ambulance fleet management. A major challenge in this direction is to handle the heterogeneity in the presentation of different patterns in space and time. Sparsity of the underlying data also poses a challenge to capture patterns of interest. While early efforts in using tensor-based factorization (Takahashi et al. 2017), multi-task learning (Zhao et al. 2015), and spatio-temporal kernel density estimation (Zhou and Matteson 2015) approaches have shown promise, their effectiveness in handling heterogeneity and sparsity is yet to be investigated.

4.3.5 Network Patterns in ST Rasters. Given a collection of ST rasters, we are interested in discovering patterns of structure in network-based representations of ST data (described in Section 3.3.5), which occur frequently in multiple ST rasters. A commonly studied type of network structure is the arrangement of *communities* in the network, where a community represents a group of entities (nodes or links) that interact among themselves more than they do with the rest of the entities (Leskovec et al. 2010). Another type of network structure is network motifs (Milo et al. 2002), which are the basic building blocks of the connectivity patterns that frequently occur across multiple network instances. The use of network motifs has been explored in several domains, e.g., bioinformatics (Shen-Orr et al. 2002) and social networks (Milo et al. 2002).

In ST applications, network motif discovery has been used to identify the inherent ST structures of brain fMRI data (Sporns and Kötter 2004). Network communities in ST rasters can reveal interesting high-level patterns. For example, in climate data, a community can represent a set of distant locations that are experiencing similar climatic conditions and show consistent temporal activity. While the discovery of network patterns can prove to be a valuable tool in the analysis of ST raster datasets, care must be taken while using them as the spatial auto-correlation in the data can result in a number of spurious edges that can lead to misleading results. Hence, during the evaluation of patterns in network-based representations of ST rasters, it is important to filter out patterns arising from spatially neighboring locations for the discovery of long-range dependencies between distant locations.

4.4 Anomaly Detection

Anomalies are traditionally defined as instances that are remarkably different from the majority of instances in the dataset. In ST applications, detecting anomalies (also referred to as outliers)

can help us identify interesting but rare phenomena, e.g., an anomalous trajectory taken by a taxi vehicle or the changing frequencies of floods and droughts in certain areas due to climate change or human actions that change the ecology of a region. Given a method for computing pairwise similarities among ST instances (discussed in Section 3.4), anomalies can be identified as instances that are highly dissimilar from the rest of the instances. Detecting anomalies can also be seen as a by-product of clustering, since instances that do not easily become part of a cluster or form an isolated cluster that is quite far from the other clusters can be termed as anomalies. There are two areas of work that are relevant in the context of detecting anomalies in ST data. First, a number of techniques have been developed for identifying anomalies in time-series data, which are also called "discords." This includes approaches for univariate time series (Bishop 1994; Keogh et al. 2005, 2002; Wei et al. 2005) as well as multivariate time series (Baragona and Battaglia 2007; Chen et al. 2008; Galeano et al. 2006; Takeishi and Yairi 2014). A survey of anomaly detection approaches for temporal data is presented in Gupta et al. (2014). Second, approaches for detecting anomalies in spatial data, also termed "spatial outliers," have also been developed (Shekhar and Chawla 2003). Spatial outliers are defined as points whose non-spatial attributes are different from those that are in its close spatial neighborhood. Techniques for detecting spatial outliers vary in their choice of method for constructing local neighborhood and assigning anomaly scores, e.g., using spatial distance measures (Knorr and Ng 1997; Knorr et al. 2000; Lu et al. 2003a), graphical-distance measures (Kou et al. 2007; Shekhar et al. 2001), and visualization tools such as the Moran's I scatter plot, pocket plots, and the Variogram Cloud plot (Anselin 1994; Haslett et al. 1991). A review of techniques for spatial outlier detection is presented in Aggarwal (2017). However, the joint presence of spatial and temporal aspects introduces novel ways of describing anomalies in ST data, resulting in unique anomaly detection problems, as described below.

4.4.1 ST Point Anomalies. The concept of spatial outliers can be easily extended to spatio-temporal domains, where the neighborhood of an ST point is defined with respect to both space and time. Thus, a spatio-temporal outlier is basically an ST point that breaks the natural ST auto-correlation structure of the normal points. One approach to find ST outliers is to first cluster the normal points using approaches such as ST-DBSCAN and then report points that did not conform well to the discovered clusters (Kut and Birant 2006). Another category of approaches presented in Cheng and Li (2004, 2006) is to aggregate ST clusters at a coarser scale so that the effect of outliers on the clustering is reduced. ST outliers are then detected by comparing the original clustering with the coarsened clustering. Note that most ST outlier detection algorithms assume homogeneity in neighborhood properties across space and time, which can be violated in the presence of ST heterogeneity. This can be handled by methods that model the variance of normal instances in the local neighborhood of every point, along with its expected value (Sun and Chawla 2004).

4.4.2 Trajectory Anomalies. There are several ways we can identify a trajectory to be anomalous. A common approach is to compute pairwise similarities among trajectories (discussed in Section 3.4.2) and identify trajectories that are spatially distant from the others. For example, Lee et al. (2008) explored varying notions of distances between trajectory sub-sequences (such as perpendicular distance, parallel distance, and angle distance) and identified a trajectory to be anomalous if it had very few neighbors within a certain threshold. Distance-based methods for detecting anomalous trajectories have also been explored in Bu et al. (2009), where trajectories that appear in distant spatial regions than the rest of the trajectories are considered anomalous. Another class of approaches is to aggregate the spatial shapes of trajectories using a coarse spatial grid and then compute useful characteristics of a trajectory in every spatial grid cell. For example, we can compute the average direction of trajectories in every spatial grid and identify a trajectory as anomalous if it deviates from the expected behavior in a number of grid cells that it covers, as

proposed in Ge et al. (2010). A graph-based approach for detecting outliers in traffic data streams was developed in Liu et al. (2011), where nodes are regions and edge weights represent the traffic flow between regions. Traffic outliers were discovered as edge anomalies in this graph, which were subsequently analyzed for causal interactions using causal outlier trees. Supervised methods for detecting anomalous trajectory shapes (or motifs) have also been explored in Li et al. (2006, 2007).

While the techniques discussed above for identifying trajectory anomalies focused on the spatial pattern (e.g., shape, distance, or angle) of trajectories, we can also consider a trajectory to be anomalous if it deviates from its local neighbors as it switches from one cohort to other, e.g., a zebra moving from one group to another. To identify such anomalies, the approach presented in Li et al. (2009) computed the spatial neighbors of every moving object at any given time stamp (using trajectory information in the immediate history) and designed an anomaly score based on deviations from the trajectories of neighbors. Such an approach can enable the discovery of trajectories that are contextually anomalous in their local neighborhoods and can thus be robust to aggregate changes such as population shifts. It can also be useful for detecting anomalies in an on-line fashion, as trajectories that begin to drift apart from their local neighbors will start to accumulate large anomaly scores in a short number of time stamps.

4.4.3 Group Anomalies in ST Rasters. While techniques described in Section 4.4.1 are well suited for discovering ST point anomalies, it is often the case that anomalies appear in ST raster data as spatially contiguous groups of locations (regions) that show anomalous values consistently for a short duration of time stamps. Some examples of such group anomalies in ST raster data include rare events such as cyclones, floods, and droughts that result in abnormally high or low precipitation in a given region for a certain duration of time or an abnormal number of tweets or emergency calls from a spatial region in a small time window.

Most approaches for detecting group anomalies in ST raster data decompose the anomaly detection problem by first treating the spatial and temporal properties of the outliers independently, which are then merged together in a post-processing step. For example, the approach presented in Wu et al. (2010) made use of the spatial scan statistic to find the top-k contiguous groups of locations that showed anomalous activity at every time stamp. These groups of locations were then stitched together in time to find ST groups of anomalous values. A similar line of work has been explored in Lu et al. (2007), where the anomalous spatial regions were discovered using an image segmentation algorithm and a regression technique was applied to track the movement of the centers of the anomalous regions across consecutive time stamps. Another approach for detecting anomalies in ST raster data was explored in Faghmous et al. (2013a) for discovering ocean eddies in climate data. Ocean eddies are revolving masses of water in the ocean that appear as local extremas (depressions or elevations) in a snapshot of sea surface height data on any given time stamp, and these extremas keep moving over time. The approach presented in Faghmous et al. (2013a) identified eddies as local extremas in space using a bottom-up thresholding scheme, which could then be stitched in time using multiple hypothesis object tracking procedures (Faghmous et al. 2013b). An orthogonal approach to detecting anomalies in ST rasters is to find anomalies in time-series data that can then be merged across space (Faghmous et al. 2012).

Approaches for detecting anomalies in ST raster that jointly use information about the spatial and temporal aspects of the data have also been developed. For example, in the application of detecting abnormal activities in crowded scenes from surveillance camera videos, anomaly detection approaches using models of normal activity have been developed in Li et al. (2014a) and Kratz and Nishino (2009). In this problem, a major difficulty is to come up with a notion of normal motion activity that can be differentiated from anomalous activities, which can be quite complex in crowded atmospheres. To address this challenge, Li et al. (2014a) used a mixture of dynamic textures models

to capture the spatial and temporal salience of normal activities, borrowing from the vast literature on activity recognition in the area of computer vision. Spatial and temporal anomaly maps are then constructed at multiple spatial scales, which are then integrated together using a conditional random field framework. On the other hand, the approach presented in Kratz and Nishino (2009) models the steady-state motion behavior of normal space–time volumes that can capture the variations in the ST data and can compactly represent the overall video volume. These models are then used to detect unusual motion activities as statistical deviations from the normal model.

A special type of group anomaly in ST rasters is *bursts* of activity in the time series of locations, detected for short time periods. For example, Twitter queries involving specific terms such as "earthquake" can be discovered as spatio-temporal burst events related to natural disasters. One of the early works on detecting ST burst events includes the work by Lappas et al. (2012), where both the spatial and temporal nature of burst patterns were jointly taken into consideration. A system for mining ST burst events, termed the Spatio-Temporal Miner, has been developed in Lappas et al. (2013).

4.5   Change Detection

The problem of change detection involves identifying the time point when the behavior of a system undergoes a significant deviation from its past behavior. Detecting changes in ST data is useful in several applications, e.g., discovering transitions from an El Nino phase to a La Nina phase in climate science (Miralles et al. 2014) or understanding the switching of the state of the brain from planning to cognition in brain imaging data (Tang et al. 2012).

Change detection has been extensively studied in the context of time-series data, where the objective is to determine time intervals (segments) that exhibit homogeneous properties (Huang et al. 2013; Keogh et al. 1993; Liu et al. 2008; Mithal et al. 2012). Different types of homogeneity in segments such as mean (Bernaola-Galván et al. 2001; Horváth 2001), variance (Andreou and Ghysels 2002; Inclan and Tiao 1994; Whitcher et al. 2000), and distribution statistics (Carpena and Bernaola-Galván 1999; Grosse et al. 2002) have been used for identifying time-series segments and their changes. Apart from time-series segmentation, other approaches has also been explored for identifying changes in time series. For example, variational approaches for switching state space models have been used to discover abnormal breathing episodes during sleep (also known as sleep apnea) based on chest volume measurements (Ghahramani and Hinton 2000). Another category of approaches is change detection methods for periodic or semi-periodic data. Boriah et al. (2010) defined an iterative bottom-up approach for change detection in periodic time series for identifying forest fires, where adjoining annual segments are merged if they are highly similar, after accounting for the noise and seasonal variability in the data. Recently, Chandola and Vatsavai (2011) defined different types of changes in "periodic" time series (which exhibits annual cycles) and proposed a Gaussian Process–based approach for discovering such changes. Other techniques for time-series change detection have been explored in remote sensing applications (Karpatne et al. 2016a), e.g., using anomaly detection approaches (Lunetta et al. 2006; Mithal et al. 2011b), forecasting-based approaches (Chandola and Vatsavai 2011; Liang et al. 2014), and sub-sequence pattern matching approaches (Salmon et al. 2011; Zhu et al. 2012).

In ST applications involving raster data, it is important to consider the spatial context of time series at every location to identify changes in both space and time. A recent work by Chen et al. (2013) defined a new type of change called *contextual change*, where a time series is considered to be changing if it deviates from other time series in its local context. The context of a time series can be defined in a number of ways, e.g., by considering the group of time series that are similar to the given time series for a period of time or the time series observed at locations in close spatial vicinity. Another generalization of change detection problems in ST applications is to determine

the spatial extent and temporal window in an ST raster where a change has likely manifested. For example, the set of geographic locations and time intervals where loss in vegetation occurred due to deforestation in a forest area is of interest to ecologists. Recent work has explored some efficient approaches to capture specific type of events in global vegetation index data. An efficient approach that enumerates and prunes candidate spatio-temporal windows has been proposed by Zhou et al. (2013). This approach is referred to as a space–time window enumeration and pruning approach and was shown to be efficient for discovering the space and time subspaces where there is a consistent decrease in vegetation. Zhou et al. (2011) defined change detection in the context of a path (i.e., a highway or a longitude), where their goal is to determine sub-paths where abrupt changes are seen. They proposed a a sub-path enumeration and pruning approach and have shown its promise in discovering abrupt changes in vegetation across longitudes in Africa.

Approaches that are relevant to spatio-temporal change detection problem have also been studied in the domain of image analysis where the objective is to determine objects and segments of videos where they are present (Grundmann et al. 2010; Lhermitte et al. 2008; Moscheni et al. 1998). Suitability of these approaches to spatio-temporal data such as remote sensing and magnetic resonance imaging is yet to be explored. A taxonomy of possible generalization of "change" in spatio-temporal data is provided by Zhou et al. (2014), where they discuss applications that require the spatial extent of change to be defined as a point, line-segment, polygon, or network, and the temporal extent of changed to be defined as a time point or a time interval.

4.6 Relationship Mining

In the domain of time-series data mining, relationships among pairs of time series can be discovered using any of the similarity measures defined over time-series instances as described in Section 3.4.3. However, in ST applications where the time series are also associated with a spatial information about the location of their observation, relationships in ST data can be defined in numerous ways. One type of relationship is defined as a similarity/dissimilarity between two distant groups of contiguous locations. Such long-range relationships are of interest in a number of applications such as climate science and brain imaging. For example, while positive correlations have generally been studied in brain imaging (Bassett and Bullmore 2006), negative correlations (also referred to as teleconnections) among distant locations have been studied in climate datasets (Yang and DelSole 2012). In such problems, the goal is to determine both the sets of contiguous locations, not just a pair of locations, that participate in a relationship. Hence, it is important to discover the pairs of regions as well as their relationships simultaneously. This is because if the regions are discovered independently of the relationships between them, e.g., using a clustering algorithm, then we may miss a number of relationships whose one or both regions fall at the boundary of the clusters. Davidson et al. (2013) designed a tensor-based approach to discover regions and relationships among them simultaneously, with applications in brain fMRI Data. Figure 8 shows the network discovered from the resting-state fMRI data of a healthy subject, which appears to be highly similar to a commonly found and widely studied network known as the *default-mode network* in resting subjects' fMRI scans (Raichle and Snyder 2007).

An important consideration when mining relationships in ST data is that the strength of relationships among pairs of regions may vary with time. For example, Handwerker et al. (2012) demonstrated that the correlation between time series from the posterior cingulate region and other locations in the brain change with time. Due to this, it becomes necessary to determine the pair of interacting regions as well as the time window within which they interact. Atluri et al. (2014) designed a pattern-mining-based approach to study dynamic relationships among time series from different regions (with homogeneous time series). This approach was used in the context of brain fMRI data to discover sets of intermittently synergistic brain regions. An example of such
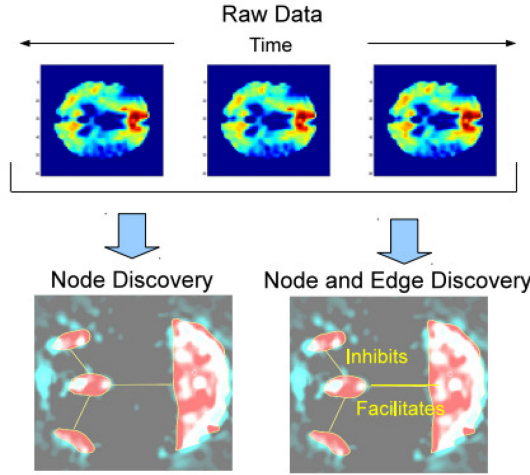
Fig. 8. Example of a set of nodes and edges that are simultaneously discovered using a tensor decomposition of spatio-temporal data proposed in Davidson et al. (2013). (Figure taken from Davidson et al. (2013).)
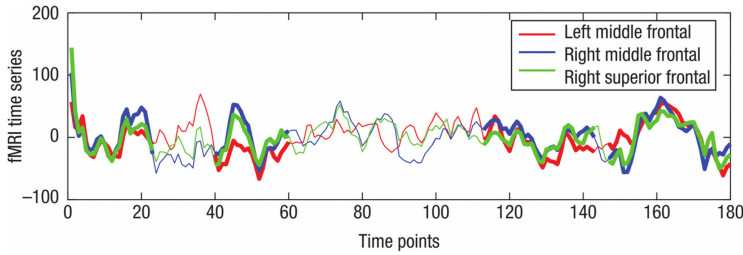

Fig. 9. Example of a time-series pattern that shows three different brain regions exhibiting high similarity in different time intervals. This patterns is discovered using the pattern-mining approach proposed in Atluri et al. (2014). (Figure taken from Atluri et al. (2016).)

a set is shown in Figure 9. Another approach was explored in Kawale et al. (2011) using graph-based methods for climate science applications, where a new graph was constructed for each time interval and the relationships within a time interval are discovered from the corresponding graph.

Another important consideration when discovering relationships in ST data is that the relationships among time series could exist with time lags. This can be because one region has an influence on the other, and the lag is due to the time it takes to conduct the influence through the system. Such relationships are referred to as lagged relationships, and they have been studied in several applications such as climate science (Chen et al. 2011; Lu et al. 2016). A more complex version of lagged relationships is dynamic lagged relationships, where the lagged relationships lasts for a small interval rather than the entire time. In this case, the goal is to determine the two sets of locations as well as the intervals in each of them that exhibit the desired similarity. A central challenge in discovering lagged relationships is the increase in the number of degrees of freedom available for finding relationships, which can result in spurious detections unless statistical corrections for multiple hypothesis testing have been carefully performed (Bland and Altman 1995).

While lagged relationships suggest some type of causal association between the two regions, causal relationships that have been defined in time-series data are known to capture truly causal

associations (Granger 1969). Clive Granger, who received a Nobel prize for this work, defined causality based on the notion that a cause (x) can predict the effect (y) significantly better than an auto-regressive model of the effect (y) itself. This work can be generalized for multivariate time-series data where the effect of all variables on a target variable can be studied. Such an approach does not take into account the causal relationships that could exist between the causal factors themselves. This observation motivated the need to develop Granger causal maps for multivariate time series where the causal relationships between all variables are assessed (Eichler 2013). These are further extended to capture sparsity (Arnold et al. 2007) and group structure (Lozano et al. 2009a) in the underlying model. The framework of Pearl causality has also been explored for identifying causal interactions in climate science (Ebert-Uphoff and Deng 2012, 2017; Hannart et al. 2016), where probabilistic graphical models were constructed using causal edges among the nodes.

5 CONCLUSIONS AND FUTURE WORK

In this article, we attempted to provide a broad overview of the field of spatio-temporal data mining that covers research spanning more than several decades. However, given the vast literature on this topic and space limitations, we have only been able to cover a small fraction of work in this fast growing area of research. Still, our hope is that this article provides the right foundation for consolidating the rich and diverse literature on STDM research, which has been explored in different application contexts for different types of ST data, under a common over-arching structure. The organization of STDM research presented in this article can be used as a stepping stone for future reviews on this topic that provide detailed empirical comparisons of different STDM methods, e.g., in terms of accuracy, efficiency, and scalability. While such evaluations have been performed in limited contexts for specific topics, e.g., for trajectory pattern mining (Giannotti et al. 2007; Mamoulis et al. 2004) and time-series data mining (Ding et al. 2008; Esling and Agon 2012), an extensive evaluation covering the full breadth of STDM problems and methods is missing.

We also hope that this article helps the reader in identifying the relevant set of questions that can be posed for a given ST dataset, and the right choice of methods for pursuing a research question in STDM. This can be particularly helpful in translating STDM research ideas developed in one community to another, so that new problems in novel application domains can be solved. For example, in flow dynamics, data pertaining to the breakup of liquid droplets is generated using simulations where droplets are present in some spatial locations in arbitrary shapes. The goal in these studies is to study droplet dynamics, i.e., changes in shape and position of the droplets with time under various flow and ambient conditions. Such patterns can be discovered using approaches for finding moving clusters of migratory animals, discussed in Section 4.1.2 in the context of clustering trajectories.

The variety of data types, problems, and methods, and the range of emerging application areas where ST data are being increasingly collected and new scientific questions explored, makes spatio-temporal data mining a quintessential melting pot for new research in data mining. One of the major emerging themes in STDM research is of studying novel representations of ST raster datasets. Most recent work in this direction involves defining novel types of edges or relationships between spatial entities and developing effective approaches for discovering them (Agrawal et al. 2017). Much of the work in this area focuses on capturing "static" edges, while there is increasing recognition that "dynamic" edges are better suited for representing ST raster data. Approaches for effectively capturing such edges in spatial-temporal varying graphs have been explored in Chen et al. (2010) for applications in climate science.

Another emerging direction where there is tremendous interest is to integratively mine multi-modal spatio-temporal datasets. Multi-modal ST datasets are available in domains such as

neuroimaging and climate science. In neuroimaging, fMRI and ECG capture the same underlying brain activity using very different technologies that offer different spatial and temporal resolution. However, in climate science, different variables such as temperature, pressure, humidity, and precipitation are available at the same spatial and temporal resolution. In the former case, the objective of interest is to construct a best possible image of brain activity based on the images from different modalities. Recent work in the form of BrainZoom (Fu et al. 2017) is one of the first efforts in this direction where they developed an optimization-based approach to construct the best possible brain activity image based on fMRI and MEG images. In the latter case, the objective of interest is to learn a model of the climate system by taking into account the different climate variables and their inter-relationships (Tye et al. 2016).

Yet another direction that is most relevant in particular to ST data involves tackling the problem of determining the level of granularity or resolution at which a phenomenon of interest needs to be searched for, be it clusters, patterns, or anomalies. Most of the existing work in ST data mining overlooks this granularity problem. A closely related contribution is made in the context of network science where it was found that existing community detection approaches might miss important substructures in the network as they over-partition or under-partition the network (Fortunato and Barthélemy 2007). They also noted that the definition of modularity that these techniques rely on can have an inherent resolution limit and argued for the need of developing approaches without such limits (Fortunato and Barthélemy 2007). To address this problem, Delvenne et al. (2010) proposed an approach to discover clusters that are relevant at different granularities. Similarly, in the context of ST data, there is a need for understanding the inherent limitations of existing approaches in tackling the granularity problem and for developing new approaches that can overcome such limitations.

Since many problems in physical sciences involve the study of processes that are spatio-temporal in nature, e.g., the dynamics of turbulent flow or the evolution of climate states and weather patterns, there is a growing interest to incorporate the scientific basis of such processes in the STDM framework. While traditional STDM methods typically rely only on the information contained in the data, the complex nature of problems and the paucity of observations in scientific applications requires a principled way of integrating data science methods with the wealth of scientific knowledge, often encoded as physics- (or theory-) based models. This is the paradigm of theory-guided data science (Karpatne et al. 2017) that is gaining attention in several disciplines for accelerating scientific discovery from data and is particularly relevant for STDM applications. Overall, considering the emerging problems and promising research directions, we anticipate the relatively young field of STDM to grow significantly over the next decade.

REFERENCES

Hamed Abdelhaq, Christian Sengstock, and Michael Gertz. 2013. Eventweet: Online localized event detection from twitter. *VLDB* 6, 12 (2013), 1326–1329.

Deepak Agarwal, Andrew McGregor, Jeff M. Phillips, Suresh Venkatasubramanian, and Zhengyuan Zhu. 2006. Spatial scan statistics: Approximations and performance study. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'06)*. ACM, 24–33.

Charu C. Aggarwal. 2015. Mining spatial data. In *Data Mining*. Springer, 531–555.

Charu C. Aggarwal. 2017. Spatial outlier detection. In *Outlier Analysis*. Springer, 345–368.

Rakesh Agrawal and Ramakrishnan Srikant. 1995. Mining sequential patterns. In *Proceedings of the IEEE International Conference on Data Engineering (ICDE'95)*. IEEE, 3–14.

Saurabh Agrawal, Gowtham Atluri, Anuj Karpatne, William Haltom, Stefan Liess, Snigdhansu Chatterjee, and Vipin Kumar. 2017. Tripoles: A new class of relationships in time series data. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'17)*. ACM, 697–706.

Ghazi Al-Naymat, Sanjay Chawla, and Joachim Gudmundsson. 2007. Dimensionality reduction for long duration and complex spatio-temporal queries. In *Proceedings of the Symposium on Applied Computing*. ACM, 393–397.

Jonathan Alon, Stan Sclaroff, George Kollios, and Vladimir Pavlovic. 2003. Discovering clusters in motion time-series data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'03)*, Vol. 1. IEEE, 375–381.

Helmut Alt and Michael Godau. 1995. Computing the fréchet distance between two polygonal curves. *International J. Comput. Geometry & Appl.* 5, 01n02 (1995), 75–91.

Elena Andreou and Eric Ghysels. 2002. Detecting multiple breaks in financial market volatility dynamics. *J. Appl. Econom.* 17, 5 (2002), 579–600.

Luc Anselin. 1994. Exploratory spatial data analysis and geographic information systems. *New Tools Spatial Anal.* 17 (1994), 45–54.

Luc Anselin. 1995. Local indicators of spatial association LISA. *Geograph. Anal.* 27, 2 (1995), 93–115.

Masanao Aoki. 2013. *State Space Modeling of Time Series.* Springer Science & Business Media.

Andrew Arnold, Yan Liu, and Naoki Abe. 2007. Temporal causal modeling with graphical granger methods. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'07)*. ACM, 66–75.

Gowtham Atluri, Angus MacDonald III, Kelvin O. Lim, and Vipin Kumar. 2016. The brain-network paradigm: Using functional imaging data to study how the brain works. *Computer* 49, 10 (2016), 65–71.

Gowtham Atluri, Kanchana Padmanabhan, Gang Fang, Michael Steinbach, Jeffrey R. Petrella, Kelvin Lim, Angus MacDonald, Nagiza F. Samatova, P. Murali Doraiswamy, and Vipin Kumar. 2013. Complex biomarker discovery in neuroimaging data: Finding a needle in a haystack. *NeuroImage: Clin.* 7, 3 (2013), 123–131.

Gowtham Atluri, Michael Steinbach, Kelvin Lim, Angus MacDonald III, and Vipin Kumar. 2014. Discovering groups of time series with similar behavior in multiple small intervals of time. In *Proceedings of the SIAM International Conference on Data Mining.*

Gowtham Atluri, Michael Steinbach, Kelvin O. Lim, Vipin Kumar, and Angus MacDonald. 2015. Connectivity cluster analysis for discovering discriminative subnetworks in schizophrenia. *Hum. Brain Map.* 36, 2 (2015), 756–767.

Mohammad Bahadori, Qi Rose Yu, and Yan Liu. 2014. Fast multivariate spatio-temporal analysis via low rank tensor learning. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS'14)*. 3491–3499.

Roberto Baragona and Francesco Battaglia. 2007. Outliers detection in multivariate time series by independent component analysis. *Neur. Comput.* 19, 7 (2007), 1962–1984.

Danielle Bassett and E. D. Bullmore. 2006. Small-world brain networks. *Neuroscientist* 12, 6 (2006), 512–523.

Pierre Bellec, Vincent Perlbarg, Saâd Jbabdi, Mélanie Pélégrini-Issac, Jean-Luc Anton, Julien Doyon, and Habib Benali. 2006. Identification of large-scale networks in the brain using fMRI. *NeuroImage* 29, 4 (2006), 1231–1243.

Johannes Berg and Michael Lässig. 2006. Cross-species analysis of biological networks by Bayesian alignment. *Proc. Natl. Acad. Sci. USA* 103, 29 (2006), 10967–10972.

Pedro Bernaola-Galván, Plamen Ch. Ivanov, Luís A. Nunes Amaral, and H. Eugene Stanley. 2001. Scale invariance in the nonstationarity of human heart rate. *Phys. Rev Lett.* 87, 16 (2001), 168105.

Derya Birant and Alp Kut. 2007. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data Knowl. Eng.* 60, 1 (2007), 208–221.

Christopher Bishop. 1994. Novelty detection and neural network validation. In *Vision, Image and Signal Processing*, Vol. 141. IET, 217–222.

J. Martin Bland and Douglas G. Altman. 1995. Multiple significance tests: The Bonferroni method. *Br. Med. J.* 310, 6973 (1995), 170.

Thomas Blumensath, Saad Jbabdi, Matthew F. Glasser, David C. Van Essen, Kamil Ugurbil, Timothy E. J. Behrens, and Stephen M. Smith. 2013. Spatially constrained hierarchical parcellation of the brain with resting-state fMRI. *NeuroImage* 1, 76 (2013), 313–324.

Vania Bogorny, Chiara Renso, Artur Ribeiro Aquino, Fernando Lucca Siqueira, and Luis Otavio Alvares. 2014. Constant–A conceptual data model for semantic trajectories of moving objects. *Trans. GIS* 18, 1 (2014), 66–88.

Shyam Boriah, Varun Mithal, Ashish Garg, Vipin Kumar, Michael Steinbach, Christopher Potter, and Steven A Klooster. 2010. A comparative study of algorithms for land cover change. In *Proceedings of the Conference on Intelligent Data Understanding (CIDU'10)*. 175–188.

George E. P. Box and Gwilym M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control.* Holden-Day.

Chris Brunsdon, Stewart Fotheringham, and Martin Charlton. 1998. Geographically weighted regression. *J. Roy. Stat. Sci. D* 47, 3 (1998), 431–443.

Yingyi Bu, Lei Chen, Ada Wai-Chee Fu, and Dawei Liu. 2009. Efficient anomaly monitoring over moving object trajectory streams. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'09)*. 159–168.

Huiping Cao, Nikos Mamoulis, and David W. Cheung. 2005. Mining frequent spatio-temporal sequential patterns. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'05)*. IEEE, 82–89.

Nikita Carney. 2016. All lives matter, but so does race: Black lives matter and the evolving role of social media. *Human. Soc.* 40, 2 (2016), 180–199.

P. Carpena and P. Bernaola-Galván. 1999. Statistical characterization of the mobility edge of vibrational states in disordered materials. *Phys. Rev. B* 60, 1 (1999), 201.

Pablo Samuel Castro, Daqing Zhang, Chao Chen, Shijian Li, and Gang Pan. 2013. From taxi GPS traces to social and community dynamics: A survey. *ACM Comput. Surv.* 46, 2 (2013), 17.

Varun Chandola and Ranga Raju Vatsavai. 2011. A scalable gaussian process analysis algorithm for biomass monitoring. *Stat. Anal. Data Min.* 4, 4 (2011), 430–445.

Varun Chandola, Ranga Raju Vatsavai, Devashish Kumar, and Auroop Ganguly. 2015. Analyzing big spa-tial and big spatiotemporal data: A case study of methods and ap-plications. *Big Data Anal.* 33 (2015), 239.

Haifeng Chen, Haibin Cheng, Guofei Jiang, and Kenji Yoshihira. 2008. Exploiting local and global invariants for the management of large scale information systems. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'08)*. IEEE, 113–122.

Xi Chen, Yan Liu, Han Liu, and Jaime G. Carbonell. 2010. Learning spatial-temporal varying graphs with applications to climate data analysis. In *Proceedings of the Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence (AAAI'10)*.

Xi C. Chen, James H. Faghmous, Ankush Khandelwal, and Vipin Kumar. 2015. Clustering dynamic spatio-temporal patterns in the presence of noise and missing data. In *IJCAI*. 2575–2581.

Xi C. Chen, Karsten Steinhaeuser, Shyam Boriah, Snigdhansu Chatterjee, and Vipin Kumar. 2013. Contextual time series change detection. In *Proceedings of the SIAM International Conference on Data Mining (SDM'13)*. SIAM, 503–511.

Yang Chen, James T. Randerson, Douglas C. Morton, Ruth S. DeFries, G. James Collatz, Prasad S. Kasibhatla, Louis Giglio, Yufang Jin, and Miriam E. Marlier. 2011. Forecasting fire season severity in south america using sea surface temperature anomalies. *Science* 334, 6057 (2011), 787–791.

Yu Chi Chen, En Tzu Wang, and Arbee L. P. Chen. 2016. Mining user trajectories from smartphone data considering data uncertainty. In *Proceedings of the International Conference on Big Data Analytics and Knowledge Discovery*. Springer, 51–67.

Tao Cheng, James Haworth, Berk Anbaroglu, Garavig Tanaksaranond, and Jiaqiu Wang. 2014. Spatiotemporal data mining. In *Handbook of Regional Science*. 1173–1193.

Tao Cheng and Zhilin Li. 2004. A hybrid approach to detect spatial-temporal outliers. In *Intl. Conf. on Geoinformatics Research*. 173–178.

Tao Cheng and Zhilin Li. 2006. A multiscale approach for spatio-temporal outlier detection. *Trans. GIS* 10, 2 (2006), 253–263.

Tao Cheng and Thomas Wicks. 2014. Event detection using Twitter: A spatio-temporal approach. *PloS One* 9, 6 (2014), e97807.

Flavio Chierichetti, Jon M. Kleinberg, Ravi Kumar, Mohammad Mahdian, and Sandeep Pandey. 2014. Event detection via communication pattern analysis. In *ICWSM*.

Bill Chiu, Eamonn Keogh, and Stefano Lonardi. 2003. Probabilistic discovery of time series motifs. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'03)*. ACM, 493–498.

Darya Chudova, Scott Gaffney, Eric Mjolsness, and Padhraic Smyth. 2003. Translation-invariant mixture models for curve clustering. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'03)*. 79–88.

R. Cameron Craddock, G. Andrew James, Paul E. Holtzheimer, Xiaoping P. Hu, and Helen S. Mayberg. 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Map.* 33, 8 (2012), 1914–1928.

Noel Cressie and Christopher K. Wikle. 2015. *Statistics for Spatio-Temporal Data*. John Wiley & Sons.

Aron Culotta. 2010. Towards detecting influenza epidemics by analyzing Twitter messages. In *Proceedings of the 1st Workshop on Social Media Analytics*. ACM, 115–122.

Maria Luisa Damiani. 2016. Spatial trajectories segmentation: Trends and challenges. In *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems*. ACM, 1–1.

Ian Davidson, Sean Gilpin, Owen Carmichael, and Peter Walker. 2013. Network discovery via constrained tensor analysis of fmri data. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'13)*. 194–202.

Federico De Martino, Francesco Gentile, Fabrizio Esposito, Marco Balsi, Francesco Di Salle, Rainer Goebel, and Elia Formisano. 2007. Classification of fMRI independent components using IC-fingerprints and support vector machine classifiers. *Neuroimage* 34, 1 (2007), 177–194.

R. S. DeFries and Jonathan Chan. 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. *Remote Sens. Environ.* 74, 3 (2000), 503–515.

J.-C. Delvenne, Sophia N. Yaliraki, and Mauricio Barahona. 2010. Stability of graph communities across time scales. *Proc. Natl. Acad. Sci. USA* 107, 29 (2010), 12755–12760.

Dingxiong Deng, Cyrus Shahabi, Ugur Demiryurek, Linhong Zhu, Rose Yu, and Yan Liu. 2016. Latent space model for road networks to predict time-varying traffic. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1525–1534.

Peter J. Diggle. 2013. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press.

Hui Ding, Goce Trajcevski, Peter Scheuermann, Xiaoyue Wang, and Eamonn Keogh. 2008. Querying and mining of time series data: Experimental comparison of representations and distance measures. *Proc. VLDB Endow.* 1, 2 (2008), 1542–1552.

Philip M. Dixon. 2002. Ripley's K function. *Encyclopedia of Environmetrics* (2002).

S. Dodge, Robert Weibel, and Anna-Katharina Lautenschütz. 2008. Towards a taxonomy of movement patterns. *Info. Vis.* 7, 3-4 (2008), 240–252.

Imme Ebert-Uphoff and Yi Deng. 2012. Causal discovery for climate research using graphical models. *Journal of Climate* 25, 17 (2012), 5648–5665.

I. Ebert-Uphoff and Y. Deng. 2017. Causal discovery in the geosciences—Using synthetic data to learn how to interpret results. *Comput. Geosci.* 99 (Feb. 2017), 50–60.

Emre Eftelioglu, Shashi Shekhar, James M. Kang, and Christopher C. Farah. 2016. Ring-shaped hotspot detection. *IEEE Trans. Knowl. Data Eng.* 28, 12 (2016), 3367–3381.

Emre Eftelioglu, Shashi Shekhar, Dev Oliver, Xun Zhou, Michael R. Evans, Yiqun Xie, James M Kang, Renee Laubscher, and Christopher Farah. 2014. Ring-shaped hotspot detection: A summary of results. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'14)*. IEEE, 815–820.

Michael Eichler. 2013. Causal inference with multiple time series: Principles and problems. *Phil. Trans. R. Soc. Lond. A.* 371, 1997 (2013), 20110613.

Anders Eklund, Thomas E. Nichols, and Hans Knutsson. 2016. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl Acad. Sci. USA* (2016), 201602413.

Philippe Esling and Carlos Agon. 2012. Time-series data mining. *ACM Comput. Surv.* 45, 1 (2012), 12.

Martin Ester, Hans-Peter Kriegel, and Jörg Sander. 1997. Spatial data mining: A database approach. In *ISSD*. Springer, 47–66.

Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, and others. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'96)*, Vol. 96. 226–231.

James H. Faghmous, Yashu Chamber, Shyam Boriah, Frode Vikebø, Stefan Liess, Michel dos Santos Mesquita, and Vipin Kumar. 2012. A novel and scalable spatio-temporal technique for ocean eddy monitoring. In *Proceedings of the Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence (AAAI'12)*.

James H. Faghmous, Matthew Le, Muhammed Uluyol, Vipin Kumar, and Snigdhansu Chatterjee. 2013a. A parameter-free spatio-temporal pattern mining model to catalog global ocean dynamics. In *Proceedings of the 2013 IEEE 13th International Conference on Data Mining (ICDM'13)*. IEEE, 151–160.

James H. Faghmous, Muhammed Uluyol, Luke Styles, Matthew Le, Varun Mithal, Shyam Boriah, and Vipin Kumar. 2013b. Multiple hypothesis object tracking for unsupervised self-learning: An ocean eddy tracking application. In *Proceedings of the Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence (AAAI'13)*.

Damien R. Farine, Ariana Strandburg-Peshkin, Tanya Berger-Wolf, Brian Ziebart, Ivan Brugere, Jia Li, and Margaret C. Crofoot. 2016. Both nearest neighbours and long-term affiliates predict individual locations during collective movement in wild baboons. *Sci. Rep.* 6 (2016), 27704.

Jan H. Feldhoff, Stefan Lange, Jan Volkholz, Jonathan F. Donges, Jürgen Kurths, and Friedrich-Wilhelm Gerstengarbe. 2015. Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate. *Clim. Dynam.* 44, 5–6 (2015), 1567–1581.

Wei Feng, Chao Zhang, Wei Zhang, Jiawei Han, Jianyong Wang, Charu Aggarwal, and Jianbin Huang. 2015. STREAMCUBE: Hierarchical spatio-temporal hashtag clustering for event exploration over the twitter stream. In *Proceedings of the 2015 IEEE 31st International Conference on Data Engineering (ICDE'15)*. IEEE, 1561–1572.

Santo Fortunato and Marc Barthélemy. 2007. Resolution limit in community detection. *Proc. Natl. Acad. Sci. USA* 104, 1 (2007), 36–41.

Tak-chung Fu. 2011. A review on time series data mining. *Eng. App. AI* 24, 1 (2011), 164–181.

Xiao Fu, Kejun Huang, Otilia Stretcu, Hyun Ah Song, Evangelos Papalexakis, Partha Talukdar, Tom Mitchell, Nicholas Sidiropoulo, Christos Faloutsos, and Barnabas Poczos. 2017. BrainZoom: High resolution reconstruction from multi-modal brain signals. In *Proceedings of the 2017 SIAM International Conference on Data Mining*. SIAM, 216–227.

Scott Gaffney and Padhraic Smyth. 1999. Trajectory clustering with mixtures of regression models. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'99)*. ACM, 63–72.

Pedro Galeano, Daniel Peña, and Ruey S. Tsay. 2006. Outlier detection in multivariate time series by projection pursuit. *J. Am. Stat. Assoc.* 101, 474 (2006), 654–669.

Everette S. Gardner. 2006. Exponential smoothing: The state of the art–Part II. *Int. J. Forecast.* 22, 4 (2006), 637–666.

Anthony C. Gatrell, Trevor C. Bailey, Peter J. Diggle, and Barry S. Rowlingson. 1996. Spatial point pattern analysis and its application in geographical epidemiology. *Trans. Inst. Br. Geogr.* 21, 1 (1996), 256–274.

Yong Ge, Hui Xiong, Zhi-hua Zhou, Hasan Ozdemir, Jannite Yu, and Kuo Chu Lee. 2010. Top-eye: Top-k evolving trajectory outlier detection. In *Proceedings of the Conference on Information and Knowledge Management (CIKM'10)*. ACM, 1733–1736.

Zoubin Ghahramani and Geoffrey E. Hinton. 2000. Variational learning for switching state-space models. *Neur. Comput.* 12, 4 (2000), 831–864.

Joydeep Ghosh and Larry Deuser. 1995. Classification of spatio-temporal patterns with applications to recognition of sonar sequences. *Neural Representation of Temporal Patterns* (1995), 221–250.

Fosca Giannotti, Mirco Nanni, Fabio Pinelli, and Dino Pedreschi. 2007. Trajectory pattern mining. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'07)*. ACM, 330–339.

Jeremy Ginsberg, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. Detecting influenza epidemics using search engine query data. *Nature* 457, 7232 (2009), 1012–1014.

Aharona Glatman-Freedman, Zalman Kaufman, Eran Kopel, Ravit Bassal, Diana Taran, Lea Valinsky, Vered Agmon, Manor Shpriz, Daniel Cohen, Emilia Anis, and others. 2016. Near real-time space-time cluster analysis for detection of enteric disease outbreaks in a community setting. *J. Infect.* 73, 2 (2016), 99–106.

Janaína Gomide, Adriano Veloso, Wagner Meira Jr., Virgílio Almeida, Fabrício Benevenuto, Fernanda Ferraz, and Mauro Teixeira. 2011. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In *Proceedings of the 3rd International Web Science Conference*. ACM, 3.

Cyril Goutte, Peter Toft, Egill Rostrup, Finn A. Nielsen, and Lars Kai Hansen. 1999. On clustering fMRI time series. *NeuroImage* 9, 3 (1999), 298–310.

Clive W. J. Granger. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 3 (1969), 424–438.

Alex Graves and Jürgen Schmidhuber. 2009. Offline handwriting recognition with multidimensional recurrent neural networks. In *Advances in Neural Information Processing Systems*. 545–552.

Ivo Grosse, Pedro Bernaola-Galván, Pedro Carpena, Ramón Román-Roldán, Jose Oliver, and H. Eugene Stanley. 2002. Analysis of symbolic sequences using the jensen-shannon divergence. *Phys. Rev. E* 65, 4 (2002), 041905.

Matthias Grundmann, Vivek Kwatra, Mei Han, and Irfan Essa. 2010. Efficient hierarchical graph-based video segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'10)*. 2141–2148.

Manish Gupta, Jing Gao, Charu C. Aggarwal, and Jiawei Han. 2014. Outlier detection for temporal data: A survey. *Trans. Knowl. Data Eng.* 26, 9 (2014), 2250–2267.

Ralf Hartmut Güting, Fabio Valdés, and Maria Luisa Damiani. 2015. Symbolic trajectories. *ACM Trans. Spatial Algor. Syst.* 1, 2 (2015), 7.

Daniel A. Handwerker, Vinai Roopchansingh, Javier Gonzalez-Castillo, and Peter A. Bandettini. 2012. Periodic changes in fMRI connectivity. *NeuroImage* 63, 3 (2012), 1712–1719.

A. Hannart, J. Pearl, F. E. L. Otto, P. Naveau, and M. Ghil. 2016. Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Am. Meteorol. Soc.* 97 (2016), 99–110.

Robert M. Haralick and Linda G. Shapiro. 1985. Image segmentation techniques. *Graph. Model Image Process.* 29, 1 (1985), 100–132.

Frank Hardisty and Alexander Klippel. 2010. Analysing spatio-temporal autocorrelation with LISTA-Viz. *Int. J. GIS* 24, 10 (2010), 1515–1526.

Ismail Haritaoglu, David Harwood, and Larry S Davis. 2000. W 4: Real-time surveillance of people and their activities. *Trans. Pattern Anal. Mach. Intell.* 22, 8 (2000), 809–830.

Andrew C. Harvey. 1990. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

John Haslett, Ronan Bradley, Peter Craig, Antony Unwin, and Graham Wills. 1991. Dynamic graphics for exploring spatial data with application to locating global and local anomalies. *Am. Stat.* 45, 3 (1991), 234–242.

Ruth Heller, Damian Stanley, Daniel Yekutieli, Nava Rubin, and Yoav Benjamini. 2006. Cluster-based analysis of FMRI data. *NeuroImage* 33, 2 (2006), 599–608.

Jon Hills, Jason Lines, Edgaras Baranauskas, James Mapp, and Anthony Bagnall. 2014. Classification of time series by shapelet transformation. *Data Min. Knowl. Discov.* 28, 4 (2014), 851–881.

Michael Horton, Joachim Gudmundsson, Sanjay Chawla, and Joël Estephan. 2017. Classification of passes in football matches using spatiotemporal data. *ACM Trans. Spatial Algorithms Syst.* 3, 2 (2017), 6:1–6:30.

Lajos Horváth. 2001. Change-point detection in long-memory processes. *Journal of Multivariate Analysis* 78, 2 (2001), 218–234.

Xiaolin Huang, Marin Matijas, Johan Suykens, and others. 2013. Hinging hyperplanes for time-series segmentation. *IEEE Transactions on Neural Networks and Learning Systems* 24, 8 (2013), 1279–1291.

Yan Huang, Shashi Shekhar, and Hui Xiong. 2004. Discovering colocation patterns from spatial data sets: A general approach. *IEEE Trans. Knowl. Data Eng.* 16, 12 (2004), 1472–1485.

Yan Huang, Liqin Zhang, and Pusheng Zhang. 2008. A framework for mining sequential patterns from spatio-temporal event datasets. *IEEE Trans. Knowl. Data Eng.* 20, 4 (2008), 433–448.

N. Hurlburt, M. Cheung, C. Schrijver, L. Chang, S. Freeland, S. Green, C. Heck, A. Jaffey, A. Kobashi, D. Schiff, and others. 2010. Heliophysics event knowledgebase for the solar dynamics observatory (SDO) and beyond. In *The Solar Dynamics Observatory*. Springer, 67–78.

Alexander Ihler, Jon Hutchins, and Padhraic Smyth. 2006. Adaptive event detection with time-varying poisson processes. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'06)*. 207–216.

Carla Inclan and George C. Tiao. 1994. Use of cumulative sums of squares for retrospective detection of changes of variance. *J. Am. Stat. Assoc.* 89, 427 (1994), 913–923.

Ashesh Jain, Amir R. Zamir, Silvio Savarese, and Ashutosh Saxena. 2016. Structural-RNN: Deep learning on spatio-temporal graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5308–5317.

Hoyoung Jeung, Man Lung Yiu, Xiaofang Zhou, Christian S. Jensen, and Heng Tao Shen. 2008. Discovery of convoys in trajectory databases. *VLDB* 1, 1 (2008), 1068–1080.

Xiaowei Jia, Ankush Khandelwal, Guruprasad Nayak, James Gerber, Kimberly Carlson, Paul West, and Vipin Kumar. 2017b. Incremental dual-memory LSTM in land cover prediction. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 867–876.

Xiaowei Jia, Ankush Khandelwal, Guruprasad Nayak, James Gerber, Kimberly Carlson, Paul West, and Vipin Kumar. 2017a. Predict land covers with transition modeling and incremental learning. In *Proceedings of the 2017 SIAM International Conference on Data Mining*. SIAM, 171–179.

Zhe Jiang, Shashi Shekhar, Xun Zhou, Joseph Knight, and Jennifer Corcoran. 2015. Focal-test-based spatial decision tree learning. *IEEE Trans. Knowl. Data Eng.* 27, 6 (2015), 1547–1559.

Goo Jun and Joydeep Ghosh. 2011. Spatially adaptive classification of land cover with remote sensing data. *IEEE Trans. Geosci. Remote Sens.* 49, 7 (2011), 2662–2673.

Panos Kalnis, Nikos Mamoulis, and Spiridon Bakiras. 2005. On discovering moving clusters in spatio-temporal data. In *Proceedings of the Conference International Symposium on Spatial and Temporal Databases (ISSTD'05)*. Springer, 364–381.

Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. 2014. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1725–1732.

Anuj Karpatne, Gowtham Atluri, James Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar. 2017. Theory-guided data science: A new paradigm for scientific discovery. *Trans. Knowl. Data Eng.* 29, 10 (2017), 2318–2331.

Anuj Karpatne, James Faghmous, Jaya Kawale, Luke Styles, Mace Blank, Varun Mithal, Xi Chen, Ankush Khandelwal, Shyam Boriah, Karsten Steinhaeuser, and others. 2013. Earth science applications of sensor data. In *Managing and Mining Sensor Data*. Springer, 505–530.

Anuj Karpatne, Zhe Jiang, Ranga Raju Vatsavai, Shashi Shekhar, and Vipin Kumar. 2016a. Monitoring land-cover changes: A machine-learning perspective. *IEEE Geosci. Remote Sens. Mag.* 4, 2 (2016), 8–21.

Anuj Karpatne, Ankush Khandelwal, Xi Chen, Varun Mithal, James Faghmous, and Vipin Kumar. 2016b. Global monitoring of inland water dynamics: State-of-the-art, challenges, and opportunities. In *Computational Sustainability*. Springer, 121–147.

Teerasit Kasetkasem and Pramod Kumar Varshney. 2002. An image change detection algorithm based on markov random field models. *IEEE Trans. Geosci. Remote Sens.* 40, 8 (2002), 1815–1823.

Jaya Kawale, Michael Steinbach, and Vipin Kumar. 2011. Discovering dynamic dipoles in climate data. In *Proceedings of the SIAM International Conference on Data Mining (SDM'11)*. SIAM, 107–118.

Harry H. Kelejian and Ingmar R. Prucha. 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. *Int. Econ. Rev.* 40, 2 (1999), 509–533.

Eamonn Keogh, Selina Chu, David Hart, and Michael Pazzani. 1993. Segmenting time series: A survey and novel approach. In *Data Mining in Time Series Databases*. World Scientific, 1–22.

Eamonn Keogh and Shruti Kasetty. 2003. On the need for time series data mining benchmarks: A survey and empirical demonstration. *Data Min. Knowl. Discov.* 7, 4 (2003), 349–371.

Eamonn Keogh, Jessica Lin, and Ada Fu. 2005. Hot sax: Efficiently finding the most unusual time series subsequence. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'05)*. 8.

Eamonn Keogh, Stefano Lonardi, and Bill'Yuan-chi' Chiu. 2002. Finding surprising patterns in a time series database in linear time and space. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'02)*. ACM, 550–556.

Eamonn Keogh and Chotirat Ann Ratanamahatana. 2005. Exact indexing of dynamic time warping. *Knowl. Inf. Syst.* 7, 3 (2005), 358–386.

Ankush Khandelwal, Anuj Karpatne, Miriam Marlier, Julia Kim, Dennis Lettenmaier, and Vipin Kumar. 2017. An approach for global monitoring of surface water extent variations using MODIS data. In *Remote Sensing of Environment*.

Slava Kisilevich, Florian Mansmann, Mirco Nanni, and Salvatore Rinzivillo. 2009. Spatio-temporal clustering. *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach (Eds.). Springer, Boston, MA, 855–874.

Robert Kistler, William Collins, Suranjana Saha, Glenn White, John Woollen, Eugenia Kalnay, Muthuvel Chelliah, Wesley Ebisuzaki, Masao Kanamitsu, Vernon Kousky, and others. 2001. The NCEP–NCAR 50–year reanalysis: Monthly means CD–ROM and documentation. *Bull. Am. Meteorol. Soc.* 82, 2 (2001), 247–267.

Edwin M. Knorr and Raymond T. Ng. 1997. A unified notion of outliers: Properties and computation. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'97)*. 219–222.

Edwin M. Knorr, Raymond T. Ng, and Vladimir Tucakov. 2000. Distance-based outliers: Algorithms and applications. *VLDB* 8, 3–4 (2000), 237–253.

Yufeng Kou, Chang-Tien Lu, and Raimundo F. Dos Santos. 2007. Spatial outlier detection: A graph-based approach. In *Proceedings of the International Conference on Tools with Artificial Intelligence (ICTAI'07)*, Vol. 1. IEEE, 281–288.

Louis Kratz and Ko Nishino. 2009. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09)*. IEEE, 1446–1453.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS'12)*. 1097–1105.

Shih-pi Ku, Arthur Gretton, Jakob Macke, and Nikos K. Logothetis. 2008. Comparison of pattern recognition methods in classifying high-resolution BOLD signals obtained at high magnetic field in monkeys. *MRI* 26, 7 (2008), 1007–1014.

Martin Kulldorff. 1997. A spatial scan statistic. *Comm. Stat.-Theory Methods* 26, 6 (1997), 1481–1496.

Martin Kulldorff. 2001. Prospective time periodic geographical disease surveillance using a scan statistic. *J. R. Stat. Soc. A* 164, 1 (2001), 61–72.

Martin Kulldorff, Richard Heffernan, Jessica Hartman, Renato Assunçao, and Farzad Mostashari. 2005. A space–Time permutation scan statistic for disease outbreak detection. *Plos Med* 2, 3 (2005), e59.

Alp Kut and Derya Birant. 2006. Spatio-temporal outlier detection in large databases. *J Comput. Inf. Technol.* 14, 4 (2006), 291–297.

Theodoros Lappas, Marcos R. Vieira, Dimitrios Gunopulos, and Vassilis J. Tsotras. 2012. On the spatiotemporal burstiness of terms. *Proc. VLDB Endow.* 5, 9 (2012), 836–847.

Theodoros Lappas, Marcos R. Vieira, Dimitrios Gunopulos, and Vassilis J. Tsotras. 2013. STEM: A spatio-temporal miner for bursty activity. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. ACM, 1021–1024.

Yann LeCun and Yoshua Bengio. 1995. Convolutional networks for images, speech, and time series. *Handb. Brain Theory Neur. Netw.* 3361, 10 (1995), 1995.

Jae-Gil Lee, Jiawei Han, and Xiaolei Li. 2008. Trajectory outlier detection: A partition-and-detect framework. In *Proceedings of the IEEE International Conference on Data Engineering (ICDE'08)*. 140–149.

Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. 2007. Trajectory clustering: A partition-and-group framework. In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD'07)*. ACM, 593–604.

Jure Leskovec, Kevin J. Lang, and Michael Mahoney. 2010. Empirical comparison of algorithms for network community detection. In *Proceedings of the 19th International Conference on World Wide Web*. ACM, 631–640.

Stefaan Lhermitte, Jan Verbesselt, Inge Jonckheere, Kris Nackaerts, Jan A. N. van Aardt, Willem W. Verstraeten, and Pol Coppin. 2008. Hierarchical image segmentation based on similarity of NDVI time series. *Remote Sensi. Environ.* 112, 2 (2008), 506–521.

Hongfei Li, Catherine A. Calder, and Noel Cressie. 2007. Beyond moran's I: Testing for spatial dependence based on the spatial autoregressive model. *Geogr. Anal.* 39, 4 (2007), 357–375.

Miao Li, Shuying Zang, Bing Zhang, Shanshan Li, Changshan Wu, and others. 2014b. A review of remote sensing image classification techniques: The role of spatio-contextual information. *Eur. J. Remote Sens.* 47, 2014 (2014), 389–411.

Weixin Li, Vijay Mahadevan, and Nuno Vasconcelos. 2014a. Anomaly detection and localization in crowded scenes. *Trans. Pattern Anal. Mach. Intell.* 36, 1 (2014), 18–32.

Wei Li and Jing-Yu Yang. 2009. Comparing networks from a data analysis perspective. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS'09)*. Springer, 1907–1916.

Xiaolei Li, Jiawei Han, and Sangkyum Kim. 2006. Motion-alert: Automatic anomaly detection in massive moving objects. In *International Conference on Intelligence and Security Informatics*. Springer, 166–177.

Xiaolei Li, Jiawei Han, Sangkyum Kim, and Hector Gonzalez. 2007. Roam: Rule-and motif-based anomaly detection in massive moving object datasets. In *Proceedings of the 2007 SIAM International Conference on Data Mining*. SIAM, 273–284.

Xiaolei Li, Zhenhui Li, Jiawei Han, and Jae-Gil Lee. 2009. Temporal outlier detection in vehicle traffic data. In *Proceedings of the IEEE International Conference on Data Engineering (ICDE'09)*. IEEE, 1319–1322.

Yuxuan Li, James Bailey, Lars Kulik, and Jian Pei. 2013. Mining probabilistic frequent spatio-temporal sequential patterns with gap constraints from uncertain databases. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'13)*. IEEE, 448–457.

Yang Li, Yangyan Li, Dimitrios Gunopulos, and Leonidas Guibas. 2016. Knowledge-based trajectory completion from sparse GPS samples. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 33.

Zhenhui Li. 2014. Spatiotemporal pattern mining: Algorithms and applications. In *Frequent Pattern Mining*. Springer, 283–306.

Zhenhui Li. 2017. Semantic understanding of spatial trajectories. In *Proceedings of the International Symposium on Spatial and Temporal Databases*. Springer, Cham, 398–401.

Zhenhui Li, Bolin Ding, Jiawei Han, and Roland Kays. 2010. Swarm: Mining relaxed temporal moving object clusters. *VLDB* 3, 1–2 (2010), 723–734.

Zhenhui Li and Jiawei Han. 2014. Mining periodicity from dynamic and incomplete spatiotemporal data. In *Data Mining and Knowledge Discovery for Big Data*. Springer, 41–81.

L. Liang, Y. Chen, T. J. Hawbaker, Z. Zhu, and P. Gong. 2014. Mapping mountain pine beetle mortality through growth trend analysis of time-series landsat data. *Remote Sensing* 6 (2014), 5696–5716.

T. Warren Liao. 2005. Clustering of time series data a survey. *Pattern Recogn.* 38, 11 (2005), 1857–1874.

Siyuan Liu, Yunhuai Liu, Lionel M. Ni, Jianping Fan, and Minglu Li. 2010. Towards mobility-based clustering. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'10)*. ACM, 919–928.

Wei Liu, Yu Zheng, Sanjay Chawla, Jing Yuan, and Xie Xing. 2011. Discovering spatio-temporal causal interactions in traffic data streams. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1010–1018.

Xiao Liu, Catie Chang, and Jeff H. Duyn. 2013. Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. *Front. Syst. Neurosci.* 7, 4 (2013), 101.

Xiao Liu and Jeff H. Duyn. 2013. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci. USA* 110, 11 (2013), 4392–4397.

Xiaoyan Liu, Zhenjiang Lin, and Huaiqing Wang. 2008. Novel online methods for time series segmentation. *Trans. Knowl. Data Eng.* 20, 12 (2008), 1616–1626.

Aurélie C. Lozano, Naoki Abe, Yan Liu, and Saharon Rosset. 2009a. Grouped graphical Granger modeling for gene expression regulatory networks discovery. *Bioinformatics* 25, 12 (2009), i110–i118.

Aurelie C. Lozano, Hongfei Li, Alexandru Niculescu-Mizil, Yan Liu, Claudia Perlich, Jonathan Hosking, and Naoki Abe. 2009b. Spatial-temporal causal modeling for climate change attribution. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'09)*. 587–596.

C.-T. Lu, Dechang Chen, and Yufeng Kou. 2003a. Algorithms for spatial outlier detection. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'03)*. IEEE, 597–600.

Chang-Tien Lu, Yufeng Kou, Jiang Zhao, and Li Chen. 2007. Detecting and tracking regional outliers in meteorological data. *Inf. Sci.* 177, 7 (2007), 1609–1632.

Mengqian Lu, Upmanu Lall, Jaya Kawale, Stefan Liess, and Vipin Kumar. 2016. Exploring the predictability of 30-day extreme precipitation occurrence using a global SST–SLP correlation network. *J. Clim.* 29, 3 (2016), 1013–1029.

Yingli Lu, Tianzi Jiang, and Yufeng Zang. 2003b. Region growing method for the analysis of functional MRI data. *NeuroImage* 20, 1 (2003), 455–465.

Ross S. Lunetta, Joseph F. Knight, Jayantha Ediriwickrema, John G. Lyon, and L. Dorsey Worthy. 2006. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sens. Environ.* 105, 2 (2006), 142–154.

Qiang Luo, Wenlian Lu, Wei Cheng, Pedro A Valdes-Sosa, Xiaotong Wen, Mingzhou Ding, and Jianfeng Feng. 2013. Spatio-temporal granger causality: A new framework. *NeuroImage* 79 (2013), 241–263.

Mary-Ellen Lynall, Danielle S. Bassett, Robert Kerwin, Peter J. McKenna, Manfred Kitzbichler, Ulrich Muller, and Ed Bullmore. 2010. Functional connectivity and brain networks in schizophrenia. *J. Neurosci.* 30, 28 (2010), 9477–9487.

Heather J. Lynch and Paul R. Moorcroft. 2008. A spatiotemporal Ripleys K-function to analyze interactions between spruce budworm and fire in British Columbia, Canada. *Can. J. Forest Res.* 38, 12 (2008), 3112–3119.

Anne-Katrin Mahlein. 2016. Plant disease detection by imaging sensors–Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 2 (2016), 241–251.

Nikos Mamoulis. 2009. Spatio-temporal data mining. In *Encyclopedia of DB Systems*. Springer, 2725–2730.

Nikos Mamoulis, Huiping Cao, George Kollios, Marios Hadjieleftheriou, Yufei Tao, and David W. Cheung. 2004. Mining, indexing, and querying historical spatiotemporal data. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 236–245.

Yasuko Matsubara, Yasushi Sakurai, Willem G. Van Panhuis, and Christos Faloutsos. 2014. FUNNEL: Automatic mining of spatially coevolving epidemics. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'14)*. ACM, 105–114.

Jingjing Meng, Junsong Yuan, Mat Hans, and Ying Wu. 2008. Mining motifs from human motion. In *Proceedings of EURO-GRAPHICS*, Vol. 8.

Aviv Mezer, Yossi Yovel, Ofer Pasternak, Tali Gorfine, and Yaniv Assaf. 2009. Cluster analysis of resting-state fMRI time series. *NeuroImage* 45, 4 (2009), 1117–1125.

Tomas Mikolov, Martin Karafiát, Lukas Burget, Jan Cernockỳ, and Sanjeev Khudanpur. 2010. Recurrent neural network based language model. In *Interspeech*, Vol. 2. 3.

Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. 2002. Network motifs: Simple building blocks of complex networks. *Science* 298, 5594 (2002), 824–827.

David Minnen, Charles L. Isbell, Irfan Essa, and Thad Starner. 2007. Discovering multivariate motifs using subsequence density estimation and greedy mixture learning. In *Proceedings of the National Conference on Artificial Intelligence*, Vol. 22. 615.

Diego G. Miralles, Martinus J. van den Berg, John H. Gash, Robert M. Parinussa, Richard A. M. de Jeu, Hylke E. Beck, Thomas R. H. Holmes, Carlos Jiménez, Niko E. C. Verhoest, Wouter A. Dorigo, and others. 2014. El Niño–La Niña cycle and recent trends in continental evaporation. *Nat. Clim. Change* 4, 2 (2014), 122–126.

Varun Mithal, Ashish Garg, Shyam Boriah, Michael Steinbach, Vipin Kumar, Christopher Potter, Steven Klooster, and Juan Carlos Castilla-Rubio. 2011a. Monitoring global forest cover using data mining. *ACM Trans. Intell. Syst. Technol.* 2, 4 (2011), 36.

V. Mithal, A. Garg, I. Brugere, S. Boriah, V. Kumar, M. Steinbach, C. Potter, and S. Klooster. 2011b. Incorporating natural variation into time series-based land cover change identification. In *Proceedings of the 2011 NASA Conference on Intelligent Data Understanding (CIDU'11)*.

Varun Mithal, Zachary O'Connor, Karsten Steinhaeuser, Shyam Boriah, Vipin Kumar, Christopher S. Potter, and Steven A. Klooster. 2012. Time series change detection using segmentation: A case study for land cover monitoring. In *Proceedings of the 2012 Conference on Intelligent Data Understanding (CIDU'12)*. IEEE, 63–70.

Pradeep Mohan, Shashi Shekhar, James A. Shine, and James P. Rogers. 2010. Cascading spatio-temporal pattern discovery: A summary of results. In *Proceedings of the SIAM International Conference on Data Mining (SDM'10)*. 327–338.

Pradeep Mohan, Shashi Shekhar, James A. Shine, and James P. Rogers. 2012. Cascading spatio-temporal pattern discovery. *Trans. Knowl. Data Eng.* 24, 11 (2012), 1977–1992.

Douglas C. Montgomery, Cheryl L. Jennings, and Murat Kulahci. 2015. *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons.

Brendan Morris and Mohan Trivedi. 2009. Learning trajectory patterns by clustering: Experimental studies and comparative evaluation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'09)*. IEEE, 312–319.

Fabrice Moscheni, Sushil Bhattacharjee, and Murat Kunt. 1998. Spatio-temporal segmentation based on region merging. *Trans. Pattern Anal. Mach. Intell.* 20, 9 (1998), 897–915.

Abdullah Mueen. 2014. Time series motif discovery: Dimensions and applications. *Data Min. Knowl. Discov.* 4, 2 (2014), 152–159.

Mark E. J. Newman. 2006. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* 103, 23 (2006), 8577–8582.

Raymond T. Ng and Jiawei Han. 2002. Clarans: A method for clustering objects for spatial data mining. *Trans. Knowl. Data Eng.* 14, 5 (2002), 1003–1016.

Kenneth A. Norman, Sean M. Polyn, Greg J. Detre, and James V. Haxby. 2006. Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 9 (2006), 424–430.

Margaret A. Oliver and Richard Webster. 1990. Kriging: A method of interpolation for geographical information systems. *Int. J. Geogr. Inf. Syst.* 4, 3 (1990), 313–332.

William Pettersson-Yeo, Paul Allen, Stefania Benetti, Philip McGuire, and Andrea Mechelli. 2011. Dysconnectivity in schizophrenia: Where are we now?*Neurosci. Biobehav. Rev.* 35, 5 (2011), 1110–1124.

Karthik Ganesan Pillai, Rafal A. Angryk, and Berkay Aydin. 2013. A filter-and-refine approach to mine spatiotemporal co-occurrences. In *Proceedings of the SIGSPATIAL Interntional Conference on Advances in GIS*. ACM, 104–113.

Karthik Ganesan Pillai, Rafal A. Angryk, Juan M. Banda, Michael A. Schuh, and Tim Wylie. 2012. Spatio-temporal co-occurrence pattern mining in datasets with evolving regions. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'12) Workshops*. IEEE, 805–812.

Karthik Ganesan Pillai, Rafal A. Angryk, Juan M. Banda, Tim Wylie, and Michael A. Schuh. 2014. Spatiotemporal co-occurrence rules. In *New Trends in DB & IS*. Springer, 27–35.

Andrea Prati, Ivana Mikic, Mohan M. Trivedi, and Rita Cucchiara. 2003. Detecting moving shadows: Algorithms and evaluation. *Trans. Pattern Anal. Mach. Intell.* 25, 7 (2003), 918–923.

Lawrence Rabiner and B. Juang. 1986. An introduction to hidden markov models. *IEEE Signal Processing Magazine* 3, 1 (1986), 4–16.

Marcus E. Raichle and Abraham Z. Snyder. 2007. A default mode of brain function: A brief history of an evolving idea. *NeuroImage* 37, 4 (2007), 1083–1090.

Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. 2000. The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vis.* 40, 2 (2000), 99–121.

Patrick H. Ryan, Grace K. LeMasters, Pratim Biswas, Linda Levin, Shaohua Hu, Mark Lindsey, David I. Bernstein, James Lockey, Manuel Villareal, Gurjit K. Khurana Hershey, and others. 2007. A comparison of proximity and land use regression traffic exposure models and wheezing in infants. *Environmental Health Perspectives* (2007), 278–284.

Suranjana Saha, Shrinivas Moorthi, Hua-Lu Pan, Xingren Wu, Jiande Wang, Sudhir Nadiga, Patrick Tripp, Robert Kistler, John Woollen, David Behringer, and others. 2010. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* 91, 8 (2010), 1015–1057.

B. P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. van den Bergh, and K. C. Steenkamp. 2011. Unsupervised land cover change detection: Meaningful sequential time series analysis. *J. Select. Top. Appl. Earth Obs. Remote Sens.* 4, 2 (Jun. 2011), 327–335.

Michael Schroder, Hubert Rehrauer, Klaus Seidel, and Mihai Datcu. 1998. Spatial information retrieval from remote-sensing images. II. Gibbs-markov random fields. *IEEE Trans. Geosci. Remote Sens.* 36, 5 (1998), 1446–1455.

Shashi Shekhar and Sanjay Chawla. 2003. Spatial databases: A tour. Vol. 2003. NJ: Prentice hall, Upper Saddle River.

Shashi Shekhar, Michael R. Evans, James M. Kang, and Pradeep Mohan. 2011. Identifying patterns in spatial information: A survey of methods. *Data Min. Knowl. Discov.* 1, 3 (2011), 193–214.

Shashi Shekhar, Zhe Jiang, Reem Y. Ali, Emre Eftelioglu, Xun Tang, Venkata Gunturi, and Xun Zhou. 2015. Spatiotemporal data mining: A computational perspective. *ISPRS Int. J. Geo-Inf.* 4, 4 (2015), 2306–2338.

Shashi Shekhar, Chang-Tien Lu, and Pusheng Zhang. 2001. Detecting graph-based spatial outliers: Algorithms and applications (a summary of results). In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'01)*. ACM, 371–376.

Shashi Shekhar, Ranga Raju Vatsavai, and Mete Celik. 2008. Spatial and spatiotemporal data mining: Recent advances. *Data Mining: Next Generation Challenges and Future Directions* (2008), 1–34.

Shai S. Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. 2002. Network motifs in the transcriptional regulation network of Escherichia coli. *Nat. Genet.* 31, 1 (2002), 64–68.

Stephen M. Smith, Peter T. Fox, Karla L. Miller, David C. Glahn, P. Mickle Fox, Clare E. Mackay, Nicola Filippini, Kate E. Watkins, Roberto Toro, Angela R. Laird, and others. 2009. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. USA* 106, 31 (2009), 13040–13045.

Sucheta Soundarajan, Tina Eliassi-Rad, and Brian Gallagher. 2013. Which network similarity measure should you choose: An empirical study. In *Workshop on Information in Networks*.

Olaf Sporns and Rolf Kötter. 2004. Motifs in brain networks. *PLoS Biol.* 2, 11 (2004), e369.

Michael Steinbach, Pang-Ning Tan, Vipin Kumar, Steven Klooster, and Christopher Potter. 2003. Discovery of climate indices using clustering. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'03)*. ACM, 446–455.

Michael Steinbach, Pang-Ning Tan, Vipin Kumar, Christopher Potter, S. Klooster, and A. Torregrosa. 2002. Data mining for the discovery of ocean climate indices. In *Scientific Data Mining*.

Karsten Steinhaeuser and Anastasios A. Tsonis. 2014. A climate model intercomparison at the dynamics level. *Clim. Dynam.* 42, 5–6 (2014), 1665–1670.

Bryan W. Stiles and Joydeep Ghosh. 1997. Habituation based neural networks for spatio-temporal classification. *Neurocomputing* 15, 3 (1997), 273–307.

Jing Sui, Tülay Adali, Godfrey D. Pearlson, and Vince D. Calhoun. 2009. An ICA-based method for the identification of optimal FMRI features and components using combined group-discriminative techniques. *NeuroImage* 46, 1 (2009), 73–86.

Pei Sun and Sanjay Chawla. 2004. On local spatial outliers. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'04)*. IEEE, 209–216.

Kunihiko Takahashi, Martin Kulldorff, Toshiro Tango, and Katherine Yih. 2008. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *Int. J. Health Geogr.* 7, 1 (2008), 14.

Tsubasa Takahashi, Bryan Hooi, and Christos Faloutsos. 2017. AutoCyclone: Automatic mining of cyclic online activities with robust tensor factorization. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 213–221.

Naoya Takeishi and Takehisa Yairi. 2014. Anomaly detection from multivariate time-series with sparse representation. In *Proceeedings of the 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC'14)*. IEEE, 2651–2656.

Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, and Vipin Kumar. 2017. *Introduction to Data Mining*. (2nd ed.) (unpublished).

Jiliang Tang, Yi Chang, and Huan Liu. 2014. Mining social media with social theories: A survey. *KDD Explor.* 15, 2 (2014), 20–29.

Yi-Yuan Tang, Mary K. Rothbart, and Michael I. Posner. 2012. Neural correlates of establishing, maintaining, and switching brain states. *Trends Cogn. Sci.* 16, 6 (2012), 330–337.

Toshiro Tango, Kunihiko Takahashi, and Kazuaki Kohriyama. 2011. A space–Time scan statistic for detecting emerging outbreaks. *Biometrics* 67, 1 (2011), 106–115.

Graham W. Taylor, Rob Fergus, Yann LeCun, and Christoph Bregler. 2010. Convolutional learning of spatio-temporal features. In *Proceedings of the European Conference on Computer Vision*. Springer, 140–153.

Tammy M. Thompson, Sebastian Rausch, Rebecca K. Saari, and Noelle E. Selin. 2014. A systems approach to evaluating the air quality co-benefits of US carbon policies. *Nat. Clim. Change* 4, 10 (2014), 917–923.

Lisa Tompson, Shane Johnson, Matthew Ashby, Chloe Perkins, and Phillip Edwards. 2015. UK open source crime data: Accuracy and possibilities for research. *Cartogr. Geogr. Inf. Sci.* 42, 2 (2015), 97–111.

Kevin Toohey and Matt Duckham. 2015. Trajectory similarity measures. *SIGSPATIAL Spec.* 7, 1 (2015), 43–50.

Sahar Torkamani and Volker Lohweg. 2017. Survey on time series motif discovery. *Data Min. Knowl. Discov.* 7, 2 (2017).

Roberto Trasarti, Fabio Pinelli, Mirco Nanni, and Fosca Giannotti. 2011. Mining mobility user profiles for car pooling. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'11)*. ACM, 1190–1198.

Ilias Tsoukatos and Dimitrios Gunopulos. 2001. Efficient mining of spatiotemporal patterns. In *Proceedings of the Conference International Symposium on Spatial and Temporal Databases (ISSTD'01)*. Springer, 425–442.

Mari R. Tye, Stephen Blenkinsop, Hayley J. Fowler, David B. Stephenson, and Christopher G. Kilsby. 2016. Simulating multimodal seasonality in extreme daily precipitation occurrence. *J. Hydrol.* 537 (2016), 117–129.

Martijn Van Den Heuvel, Rene Mandl, and Hilleke Hulshoff Pol. 2008. Normalized cut group clustering of resting-state FMRI data. *PloS One* 3, 4 (2008), e2001.

Ranga Raju Vatsavai. 2008. *Machine Learning Algorithms for Spatio-Temporal Data Mining*. ProQuest.

Ranga Raju Vatsavai, Auroop Ganguly, Varun Chandola, Anthony Stefanidis, Scott Klasky, and Shashi Shekhar. 2012. Spatiotemporal data mining in the era of big spatial data: Algorithms and applications. In *Proceedings of the SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*. ACM, 1–10.

Florian Verhein and Sanjay Chawla. 2006. Mining spatio-temporal association rules, sources, sinks, stationary regions and thoroughfares in object mobility databases. In *Proceedings of the International Conference on Database Systems for Advanced Applications (DASFAA'06)*, Vol. 3882. Springer, 187–201.

Florian Verhein and Sanjay Chawla. 2008. Mining spatio-temporal patterns in object mobility databases. *Data Min. Knowl. Discov.* 16, 1 (2008), 5–38.

Marcos R. Vieira, Petko Bakalov, and Vassilis J. Tsotras. 2009. On-line discovery of flock patterns in spatio-temporal data. In *Proceedings of the SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (SIGSPATIAL'09)*. ACM, 286–295.

Aurore Voldoire, E. Sanchez-Gomez, D. Salas y Mélia, B. Decharme, Christophe Cassou, S. Sénési, Sophie Valcke, I. Beau, A. Alias, M. Chevallier, and others. 2013. The CNRM-CM5. 1 global climate model: Description and basic evaluation. *Clim. Dynam.* 40, 9-10 (2013), 2091–2121.

Maximilian Walther and Michael Kaisser. 2013. Geo-spatial event detection in the twitter stream. In *Proceedings of the European Conference on Information Retrieval (ECIR'13)*. Springer, 356–367.

Lizhen Wang, Pinping Wu, and Hongmei Chen. 2013. Finding probabilistic prevalent colocations in spatially uncertain datasets. *Trans. Knowl. Data Eng.* 25, 4 (2013), 790–804.

Weiqing Wang, Hongzhi Yin, Ling Chen, Yizhou Sun, Shazia Sadiq, and Xiaofang Zhou. 2017. ST-SAGE: A spatial-temporal sparse additive generative model for spatial item recommendation. *ACM Trans. Intell. Syst. Technol.* 8, 3 (2017), 48.

Li Wei, Nitin Kumar, Venkata Nishanth Lolla, Eamonn J. Keogh, Stefano Lonardi, and Chotirat (Ann) Ratanamahatana. 2005. Assumption-free anomaly detection in time series. In *Proceedings of the International Conference on Scientific and Statistical Database Management (SSDBM'05)*, Vol. 5. 237–242.

Jianshu Weng and Bu-Sung Lee. 2011. Event detection in twitter. In *Proceedings of the AAAI Conference on Weblogs and Social Media*.

Brandon Whitcher, Peter Guttorp, and Donald B. Percival. 2000. Multiscale detection and location of multiple variance changes in the presence of long memory. *J. Stat. Comput. Simul.* 68, 1 (2000), 65–87.

Elizabeth Wu, Wei Liu, and Sanjay Chawla. 2010. Spatio-temporal outlier detection in precipitation data. In *Knowledge Discovery from Sensor Data*. Springer, 115–133.

Xiangye Xiao, Xing Xie, Qiong Luo, and Wei-Ying Ma. 2008. Density based co-location pattern discovery. In *Proceedings of the SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (SIGSPATIAL'08)*. ACM, 29.

Xiaosong Yang and Timothy DelSole. 2012. Systematic comparison of ENSO teleconnection patterns between models and observations. *J. Clim.* 25, 2 (2012), 425–446.

Lexiang Ye and Eamonn Keogh. 2009. Time series shapelets: A new primitive for data mining. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'09)*. ACM, 947–956.

Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn Keogh. 2016. Matrix profile I: All pairs similarity joins for time series: A unifying

view that includes motifs, discords and shapelets. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'16)*.

Qingbao Yu, Erik B. Erhardt, Jing Sui, Yuhui Du, Hao He, Devon Hjelm, Mustafa S. Cetin, Srinivas Rachakonda, Robyn L. Miller, Godfrey Pearlson, and others. 2015b. Assessing dynamic brain graphs of time-varying connectivity in fMRI data: Application to healthy controls and patients with schizophrenia. *NeuroImage* 107 (2015), 345–355.

Rose Yu, Dehua Cheng, and Yan Liu. 2015a. Accelerated online low rank tensor learning for multivariate spatiotemporal streams. In *Proceedings of the International Conference on Machine Learning*. 238–247.

Demetrios Zeinalipour-Yazti, Song Lin, and Dimitrios Gunopulos. 2006. Distributed spatio-temporal similarity search. In *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*. ACM, 14–23.

Chao Zhang, Keyang Zhang, Quan Yuan, Luming Zhang, Tim Hanratty, and Jiawei Han. 2016. GMove: Group-level mobility modeling using geo-tagged social media. In *Proceedings of the Knowledge Discovery and Data Mining Conference (KDD'16)*. 1305–1314.

Liang Zhao, Qian Sun, Jieping Ye, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. 2015. Multi-task learning for spatio-temporal event forecasting. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1503–1512.

Yindi Zhao, Liangpei Zhang, Pingxiang Li, and Bo Huang. 2007. Classification of high spatial resolution imagery using improved Gaussian Markov random-field-based texture features. *Trans. GeoSci. Remote Sens.* 45, 5 (2007), 1458–1468.

Yu Zheng. 2015. Trajectory data mining: An overview. *ACM Trans. Intell. Syst. Technol.* 6, 3 (2015), 29.

Yan-Tao Zheng, Zheng-Jun Zha, and Tat-Seng Chua. 2012. Mining travel patterns from geotagged photos. *ACM Trans. Intell. Syst. Technol.* 3, 3 (2012), 56.

Hua Zhou, Lexin Li, and Hongtu Zhu. 2013. Tensor regression with applications in neuroimaging data analysis. *J. Acoust. Soc. Am.* 108, 502 (2013), 540–552.

Xun Zhou, Shashi Shekhar, and Reem Y. Ali. 2014. Spatiotemporal change footprint pattern discovery: An inter-disciplinary survey. *Data Min. Knowl. Discov.* 4, 1 (2014), 1–23.

Xun Zhou, Shashi Shekhar, Pradeep Mohan, Stefan Liess, and Peter K. Snyder. 2011. Discovering interesting sub-paths in spatiotemporal datasets: A summary of results. In *Proceedings of the SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 44–53.

Xun Zhou, Shashi Shekhar, and Dev Oliver. 2013. Discovering persistent change windows in spatiotemporal datasets: A summary of results. In *Proceedings of the SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*. ACM, 37–46.

Zhengyi Zhou and David S. Matteson. 2015. Predicting ambulance demand: A spatio-temporal kernel approach. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'15)*. ACM, 2297–2303.

Yan Zhu, Zachary Zimmerman, Nader Shakibay Senobari, Chin-Chia Michael Yeh, Gareth Funning, Abdullah Mueen, Philip Brisk, and Eamonn Keogh. 2016. Matrix profile II: Exploiting a novel algorithm and GPUs to break the one hundred million barrier for time series motifs and joins. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'16)*.

Zhe Zhu, Curtis E. Woodcock, and Pontus Olofsson. 2012. Continuous monitoring of forest disturbance using all available landsat imagery. *Remote Sens. Environ.* 122 (2012), 75–91.