Visualising hidden spatiotemporal patterns at multiple levels of detail

Ricardo Almeida Silva
ISEL and NOVA LINCS
Instituto Politécnico de Lisboa
Lisboa, Portugal
ricardofcsasilva@gmail.com

João Moura Pires
NOVA LINCS, FCT
Universidade NOVA de Lisboa
Caparica, Portugal
jmp@fct.unl.pt

Nuno Datia
ISEL and NOVA LINCS
Instituto Politécnico de Lisboa
Lisboa, Portugal
datia@isel.ipl.pt

Maribel Yasmina Santos
ALGORITMI Research Centre
University of Minho
Guimarães, Portugal
maribel@dsi.uminho.pt

Bruno Martins
INESC-ID and IST
University of Lisbon
Lisbon, Portugal
bruno.g.martins@tecnico.ulisboa.pt

Fernando Birra
NOVA LINCS, FCT
Universidade NOVA de Lisboa
Caparica, Portugal
fpb@fct.unl.pt

***Abstract*—Crimes, forest fires, accidents, infectious diseases, human interactions with mobile devices (e.g., tweets) are being logged as spatiotemporal events. For each event, its geographic location, time and related attributes are known with high levels of detail (LoDs). The LoD plays a crucial role when analyzing data, enhancing the user's perception of phenomena. From one LoD to another, some patterns can be easily perceived or different patterns may be detected. Modeling phenomena at different LoDs is needed, as there is no exclusive LoD at which data can be analyzed.**

Current practices work mainly on a single LoD, driven by the analysts perception, ignoring the fact that the identification of the suitable LoDs is a key issue for pointing relevant patterns.

This paper presents a Visual Analytics approach called *VAST*, that allows users to simultaneously inspect a phenomenon at different LoDs, helping them to see in what LoDs patterns emerge or in what LoDs the perception of the phenomenon is different. In this way, the analysis of vast amounts of spatiotemporal events is assisted, guiding the user in this process.

The use of several synthetic and real datasets allowed the evaluation of *VAST*, which was able to suggest LoDs with different interesting spatiotemporal patterns and the type of expected patterns.

***Index Terms*—data visualisation, spatiotemporal patterns, multiple levels of detail, visual analytics**

I. Introduction

A spatiotemporal event is a happening occurred in space and time. For example, *homicide((41.8780377, -87.6294422), 09/05/2015 20:00, 2)* represents an homicide that occurred on the latitude and longitude coordinates (41.8780377, -87.6294422), at eight o'clock of a given day, resulting in two victims. Spatiotemporal events can be described as data with the following structure: $event(S, T, A_1, \ldots, A_N)$, where S describes the geographic location of the event, T specifies the time moment, and $A_1, \ldots, A_N$ are attributes detailing what has happened. Spatiotemporal events may follow one or more spatiotemporal patterns that are non-uniform distributions of events across the space and time. Finding such spatiotemporal patterns helps to understand the associated phenomena [1].

Nowadays, Visual Analytics (VA) approaches targeting the analysis of spatiotemporal events have been developed to analyse a single phenomenon (e.g., crimes), focusing on a specific kind of pattern (e.g., spatiotemporal hotspots) [2]–[4]. However, patterns might appear in many different forms [5]. For example, in some phenomena, events form spatiotemporal clusters (e.g., tweets); in others, events form a cloud that moves in space throughout time (e.g., spreading of a disease); or, events occur spread out throughout space, but some regions reveal higher intensity (e.g., robberies). Therefore, focusing a specific type of pattern may leave many patterns undetected.

Also, most VA approaches have been designed to follow a single Level of Detail (LoD) analysis approach [6], [7]. Nevertheless, the LoD matters for the perception of patterns, and often there is no exclusive LoD to study phenomena [8]–[11]. Although the LoD plays a crucial role in the perception of patterns, users have been left with the choice of the LoD to look for patterns.

When users are not familiar with a spatiotemporal phenomenon, usually in early stage of analysis, users likely face difficulties to identify the LoDs in which patterns can be better perceived.

To enhance analyses over spatiotemporal events, we propose to move from a single user-driven LoD to a multiple LoDs analysis approach, providing the user with an understandable high-level overview of the underlying structure of the phenomenon for each LoD. By understandable high-level overview, we mean several hints about the distribution of events in space or/and in time, which can provide a glimpse of the presence or absence of patterns. Following this approach, the user might detect very soon in what LoDs there are potential patterns, and of what kind they are.

Acknowledgments This work has been supported by FCT - Fundação para a Ciência e Tecnologia MCTES, UID/CEC/04516/2013 (NOVA LINCS) and UID/CEC/00319/2013 (ALGORITMI), and COMPETE: POCI010145-FEDER007043 (ALGORITMI).

978-1-5386-7202-0/18/$31.00 ©2018 IEEE
DOI 10.1109/iV.2018.00057

A web-based VA tool anchored on the SUITE framework [12], named *VAST* (**V**isual **A**nalytics for **S**patio**T**emporal events), is here proposed, designed, developed and evaluated. *VAST* allows users to simultaneously view hints about the absence or presence of different kinds of spatiotemporal patterns at multiple LoDs. To the best of our knowledge, there is no other approach that simultaneously supports analyses over spatiotemporal events at multiple LoDs, completely independent from the application domain.

The evaluation of our proposal was conducted with two types of datasets, namely (*i*) synthetic datasets and (*ii*) real datasets. Synthetic datasets with different spatiotemporal patterns at different LoDs were produced. For most cases, *VAST* could provide a correct overview of the phenomenon allowing us to identify the LoDs in which patterns exists and, therefore, the LoDs that should be used to detail the analysis. The real datasets studied were: (*i*) forest fires in Portugal; (*ii*) violent attacks against civilians occurring in Africa; and (*iii*) robberies in Chicago. *VAST* was effective in identifying patterns present in these datasets, at different spatiotemporal LoDs.

The rest of the paper is organized as follows. Section II introduces the background information about granular theory for representing spatiotemporal data at different spatiotemporal LoDs. Section III details the interface of VAST, describing the way phenomena are analyzed at multiple LoDs. Section IV presents the experiments carried out. The relevant related work is summarized at Section V. Finally, we conclude and point out directions for future work in Section VI.

II. Primer on Granular Theory

Granular computing has emerged as a paradigm of knowledge representation [13], where granules are basic ingredients of information. Roughly, a granularity defines a division of a domain in a set of granules disjoint from each other [14]. $Counties$ and $States$ are examples of spatial granularities; $Hours$ and $Days$ are examples of temporal granularities.

Under a general theory of granularities [14], a granular computing approach was devised to model spatiotemporal phenomena at multiple LoDs. This approach was labeled as the granularities-based model [15], where a phenomenon is modeled through a collection of statements. Granules are used in the statements' arguments. For example, we can model a crime event through the statement: $crime(Oakland,\ 03/01/2015\ 18h,\ 1,\ homicide)$ where the granules used come from the granularities $Counties$, $Hours$, $Natural\ Numbers$, and $Crime\ Types$.

Statements are made at some LoD. The set of granularities involved in the statement defines the LoD at which an event is described. For example, the LoD of $crime(Oakland,\ 03/01/2015\ 18h,\ 1,\ homicide)$ is defined by the corresponding granularities: $Counties$, $Hours$, $Natural\ Numbers$, and $Crime\ Types$. Through the granularities-based model, statements can be generalized to coarser LoDs automatically. Using the granularities-based model, we are able to have a phenomenon modeled in multiple LoDs [15].

A. Summarizing Spatiotemporal Events

Let us consider a statement describing an event using a spatial granule $s \in S$ and temporal granule $t \in T$. The pair (s, t) is called a spatiotemporal granule (st-granule) of the spatiotemporal LoD (st-LoD) (S, T).

Each st-granule (s, t) indexes the set of statements spatially located at s and temporally located at t. Typically, at a very detailed st-LoD, events are sparse and mostly non-co-occurring. This means that either st-granules have no events, or have just one event. At coarser st-LoDs, the co-occurrence of events on the same st-granule becomes more likely.

On top of the granularities-based model, Silva et al. [12] developed a SUmmarizIng spatioTemporal Events framework (SUITE) that builds, for each st-LoD, summaries about phenomena represented as spatiotemporal events, called *abstracts*. Abstract values can be a number, a vector, or a matrix. SUITE consider five types of abstracts: *(i)* Global; *(ii)* Spatial; *(iii)* Temporal; *(iv)* Compact Temporal; and *(v)* Compact Spatial.

A Global Abstract summarizes all statements by a single abstract value. Known spatiotemporal statistics [5], [16] (e.g., the Knox or Mantel statistics) can be used to compute Global Abstracts. A Spatial Abstract summarizes, for each $t \in T$, all statements at *t* by a single abstract value, so we get a time series of abstract values, each one summarizing the spatial distribution of the events at granule *t*. Known spatial statistics [5], [17] (e.g., Average Nearest Neighbor or Moran's I) can be used to compute spatial abstracts. A Temporal Abstract summarizes, for each $s \in S$, all statements at *s*, by a single abstract value, so we get a map of abstract values, each one summarizing the temporal distribution of the events at granule *s*. Known temporal statistics [18] can be used to compute temporal abstracts. A Compact Temporal Abstract is just a summarization of a Temporal Abstract, i.e., a summarization of a time series of abstract values into a single abstract value. Similarly, a Compact Spatial Abstract is just a summarization of a Spatial Abstract, i.e., a summarization of a map of abstract values into a single abstract value. ST-Abstracts will refer either the Global Abstracts or the Compact Temporal Abstracts or the Compact Spatial Abstracts.

III. Visual Analytics for SpatioTemporal events

Visual **A**nalytics for **S**patio**T**emporal events (*VAST*) was developed to support analyst in the task of visually inspecting the computed abstracts at many LoDs, simultaneously, allowing users to understand not only the absence or presence of different kinds of spatiotemporal patterns, but also at which LoDs they are visible or at least in what LoDs they are more easy to be found. *VAST* implements the granularities-based model and the SUITE framework, including the abstracts presented in Appendix A.

VAST's design follows the VA Mantra: *"Analyze first, show the important, zoom, filter and analyze further, details on demand*" [8]. First of all, the interface starts by displaying ST-Abstracts at all available st-LoDs. This interactive visualization may provide hints about different patterns within

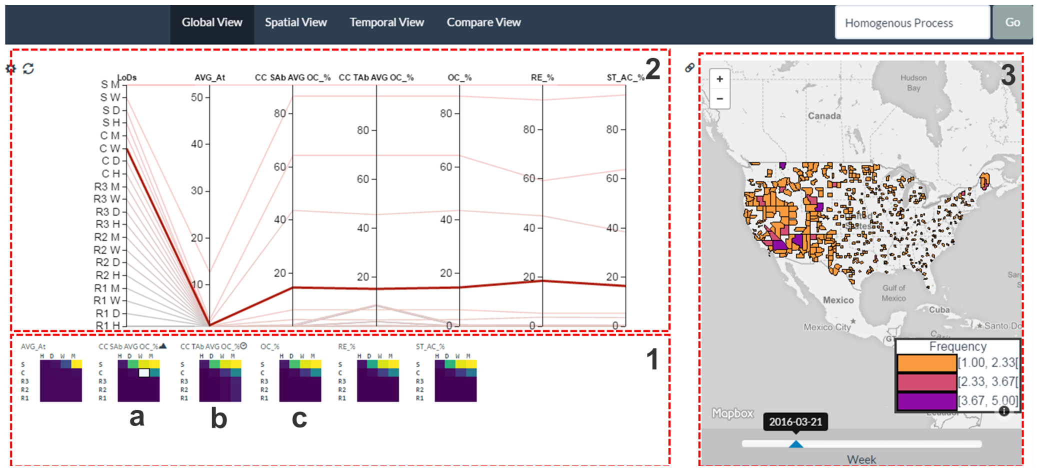


Fig. 1. An overview of the *VAST* interface.

the spatiotemporal events. Then, one can analyze further by looking at Spatial Abstracts (time series of abstract values) or Temporal Abstracts (i.e., "maps" of abstract values). At any moment of the analysis, it is possible to visually inspect the actual spatial distribution of the phenomenon at a specific temporal granule t in a particular st-LoD.

The interface is composed of three main areas, as illustrated in Fig. 1. The first area, **ST-Abstracts** (Fig. 1-1), displays a matrix plot for each ST-Abstract. The symbol ▲ points out a Compact Spatial Abstract (e.g., Fig. 1-1.a) while the symbol ⊙ indicates a Compact Temporal Abstract (e.g., Fig. 1-1.b). When none of these icons is present we have a Global Abstract. Each cell of a matrix plot shows the value of an ST-Abstract at a st-LoD. The skeleton of a matrix plot is displayed in Fig. 2. In the rows, we have the spatial granularities (finer granularities at bottom), and in the columns we have the temporal granularities (finer granularities at left). All used abstract values are numbers and their value is mapped to a color using the color scheme shown in Fig. 2. For example, Fig. 1-1 shows 6 matrices, from left to right: a Global Abstract named *Average Atoms in st-granules*, a Compact Spatial Abstract named *Average of Spatial Occupation Rate*, a Compact Temporal Abstract named *Average of Temporal Occupation Rate* and three more Global Abstracts respectively *Occupation rate*, *Reduction rate* and *Collision Rate*. All matrices are using 5 spatial granularities (State, County, and 3 rasters) and 4 temporal granularities (hours, days, weeks, and months).

The **Dynamic Abstract Area** (Fig. 1-2) is used to present 3 different visualizations: (*i*) a **Global View** that shows a Parallel Coordinates visualization with the same abstracts presented in the ST-Abstracts area; (*ii*) a **Spatial View** that shows the time series corresponding to a few selected Spatial Abstracts, as illustrated in Fig. 3; and, (*iii*) a **Temporal View** that shows the maps corresponding to a few selected Temporal Abstracts, as illustrated in Fig. 4.

In the Parallel Coordinates visualization (Fig. 1-2) each line corresponds to one st-LoD. The most left coordinate represents

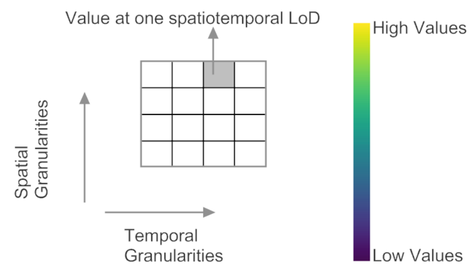


Fig. 2. An overview of the structure of a matrix plot.

the st-LoDs ordered from the more detailed st-LoD (*R1*, *Hour*) to the coarser st-LoD (*State*, *Month*). The other coordinates correspond to the ST-Abstracts presented in ST-Abstract area.

In Fig. 3, there are four cells selected from *Average of Spatial Occupation Rate* (Fig. 3.a) and from *Average of Spatial Collision Rate* (Fig. 3.b). Therefore, eight Spatial Abstracts are visible in the Spatial View, which are organized/grouped by Spatial Abstract and ordered from the more detailed st-LoD to the coarser one.

The Temporal View is illustrated in Fig. 4 and there are two st-LoDs selected from the *Average of Temporal Occupation Rate* (see Fig. 4.b): $(R3,\ Weeks)$ and $(Counties,\ Days)$. As a result, two Temporal Abstracts are displayed. When the st-LoD has a *raster* granularity, the map represents each spatial granule through a point, leading to a dot map (e.g., the map on the right side). Otherwise, the spatial granules are displayed in their original form, which leads to a choropleth map (e.g., the map on the left side).

The last area is the **Phenomena Representation** (Fig. 1-3) used to display spatiotemporal events at a st-LoD using thematic maps. The slider underneath allows the user to scroll temporally through the temporal granules, according to the st-LoD that was chosen. The map displays the number of events for each st-granule.

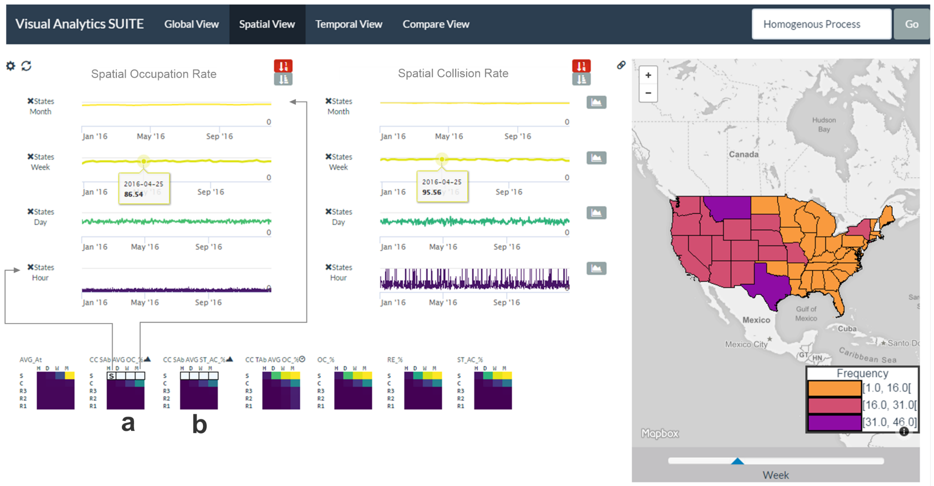


Fig. 3. An overview of the *VAST* interface with Spatial Abstracts.

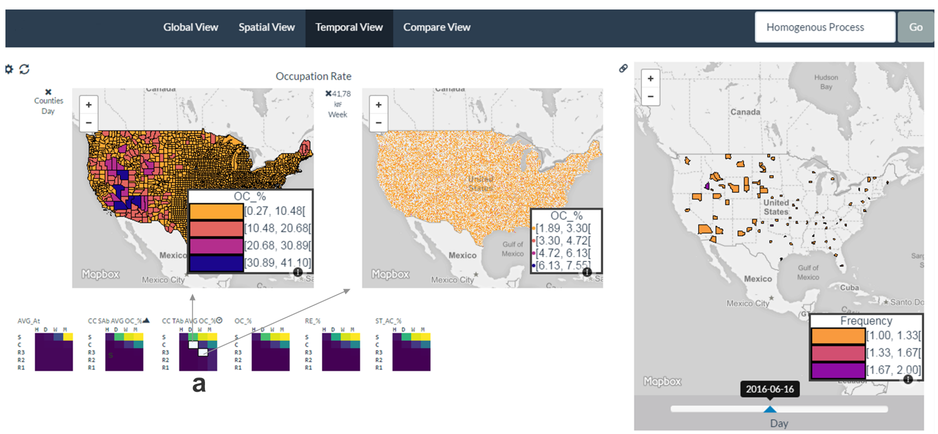


Fig. 4. An overview of the *VAST* interface with Temporal Abstracts.

IV. Experiments

VAST was used to conduct experiments over two types of datasets of spatiotemporal events: (*i*) synthetic datasets; (*ii*) real datasets.

A. Synthetic Datasets

To produce synthetic datasets of spatiotemporal events, a configurable generator of spatiotemporal events was used [16] - R package (stpp)[1]. Gabriel et al. [16] expose a set of functions in order to simulate datasets of spatiotemporal events following different patterns, where events are generated within a polygon and a single closed interval.

In a **homogeneous Poisson process**, the events form an independent random sample from the uniform distribution on the spatiotemporal domain in which the events were simulated. This model hardly approaches a pattern in a phenomenon but provides a good basis for comparison as it reflects complete spatiotemporal randomness. In the **Poisson cluster process**, a set of parents are randomly defined, and afterwards, a set of events are generated around each parent (in space and time) where the dispersion of spatiotemporal events is controlled as an input parameter. A **contagious process** can be pictured as a cloud of events moving in space throughout time. The contagion process of a disease, for example, in which the disease is transmitted to other people through direct contact with an infected person. Finally, the **log-Gaussian Cox process** simulates spatiotemporal events such that some regions reveal higher intensity.

All the datasets were generated within the region of the USA and during a period of time of one year. *Dataset 1* was simulated with the Poisson cluster process and includes 30.000 events. The clusters of events are built around the parents simulated within a spatial distance of 110 km and a temporal distance of one day. This dataset was modeled using the granularities-based model. Therefore, the most detailed spatial granularity $Raster\left(0.16km^2\right)$ is based on a grid of 16384 x 16384 cells that cover the analyzed spatial extent

[1] stpp: https://cran.r-project.org/web/packages/stpp/index.html

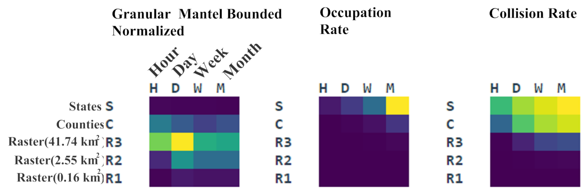


Fig. 5. Global Abstracts: *GMBN*, Occupation rate and Collision rate describing Dataset 1.

of the phenomenon, and each cell has an area of 0.16 km^2. The coarser spatial granularities were obtained by dividing by a factor of 4 the number of cells in the grid. Thus, the considered granularities for space were rasters with cell sizes approximately of 0.16 km^2 - R1, 2.55 km^2 - R2, 41.74 km^2 - R3. The granularities $Counties$ and $States$ were also included. The time granularities used were $Hours$, $Days$, $Weeks$ and $Months$.

The raw data were encoded at the finer st-LoD ($R1$, $Hours$). The granularities-based module was used to automatically produce the data for all st-LoDs and the SUITE framework was used to precompute all the abstracts defined for each LoD. Using *VAST* our analyses started by looking at ST-Abstracts. Fig.5 shows 3 ST-abstracts (i.e., the occupation rate, the collision rate and the Granular Mantel Bounded and Normalized [GMBN] indicator) for all the st-LoDs of Dataset 1.

GMBN is an adaptation of the Mantel test [19] in order to measure the level of spatiotemporal interaction/clustering among st-granules (with events). The value ranges between 0 and 1, where 0 indicates no interaction among st-granules (with events) and 1 means that all st-granules (with events) are near to each other.

The GMBN points to the st-LoD - ($R3$, $Days$) as the one with greatest spatiotemporal interaction. This seems to be compliant with the st-LoD in which the pattern was simulated. Regarding the other global abstracts (i.e., the occupation rate, and the collision rate), their values increase as long as we move to coarser *st-LoDs*. This happens because as long as we move to coarser st-LoDs, the co-occurrence of events in st-granules increases, since the number of st-granules available at coarser st-LoDs decreases.

Given the evidences pointing that there might be a pattern in the st-LoD - ($R3$, $Days$), we use the *phenomenon representation* area to have a grasp of the data at such st-LoD. The data at three different temporal granules, chosen without any particular criterion, are displayed in Fig.6. As you can see, there are clusters of events happening over the USA.

The analysis made so far points out that the Dataset 1 might have a spatiotemporal pattern, and this pattern might be better perceived at $(R3, Days)$. The pattern in question corresponds to clusters of events happening over time.

The Temporal Abstract - Temporal Center Mass's Positioning was observed for three st-LoDs as can be seen in Fig.7. The st-LoDs are: $(R3, Days)$, $(Counties, Days)$ and $(States, Days)$. Orange means that most of the events that occurred in the spatial granule were old while dark blue means that most of the events occurred in the spatial granule were recent in what concerns the extent of the temporal granularity.

Looking at Fig.7, $(R3, Days)$ and $(Counties, Days)$ st-LoDs, the regions with clusters of events can be identified, since spatial granules close to each other have similar values of the Temporal center mass's positioning. In other words, the events occurring near in space seems to occur near in time. Since clusters are happening over time, we use the Compact Spatial Abstract - Spatial average nearest neighbor (Spatial ANN) and its z-score in order to understand when those clusters of events are happening. When the value of the Spatial ANN is less than 1, the trend is towards spatial clustering while if the value is greater than 1, the trend is towards dispersion. Very low or very high z-score values suggest some spatial pattern.

With four st-LoDs $(R3, Hours)$, $(R3, Days)$, $(R3, Weeks)$, $(R3, Months)$, the Spatial Abstracts are displayed in Fig.8. The set of time series for each Temporal Abstract share the same y axis scale, and the color of a time series is given by the color used on the corresponding Compact Spatial Abstract (i.e., matrix plot). Based on Fig.8, the Spatial Abstracts revealed a clustered phenomenon over time, since the average of the Spatial ANN values points to clusters of events throughout time. In the st-LoD - $(R3, Hours)$ we can observe variations between a clustered and a non clustered phenomenon. But in the remaining st-LoDs, the phenomenon seems to be stable and clustered, because the values of the Spatial ANN are constantly close to zero and the corresponding z-scores are negative (i.e., the z-score is not close to zero).

Other datasets were simulated using the Poisson cluster and homogeneous processes, and their characteristics are summarized in Table I[2].

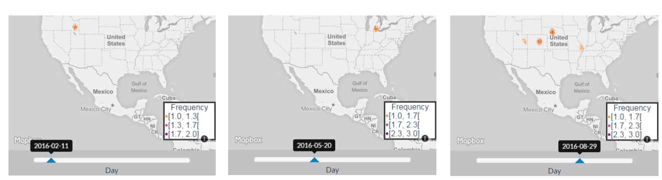
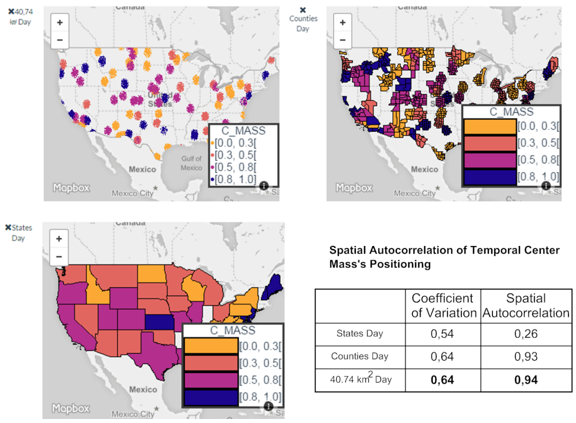


Fig. 6. Dataset 1 at the st-LoD ($R3$, $Days$).

TABLE I
DATASETS OF SPATIOTEMPORAL EVENTS SIMULATED.

Dataset	Model	Number Events	st-LoD
1	Poisson Cluster	30.000	110 Km, Days
2	Homogeneous	30.000	NA
3	Poisson Cluster	30.000	2 Km, Week
4	Poisson Cluster + Homogeneous	33.000	110 Km, Days

[2]The code to simulate the datasets and the actual datasets are available at the repository http://github.com/RFASilva/SimulatedDataSets



	Coefficient of Variation	Spatial Autocorrelation
States Day	0,54	0,26
Counties Day	0,64	0,93
40.74 km^2 Day	**0,64**	**0,94**

Fig. 7. The Temporal Center Mass's Positioning for three *st-LoDs*.

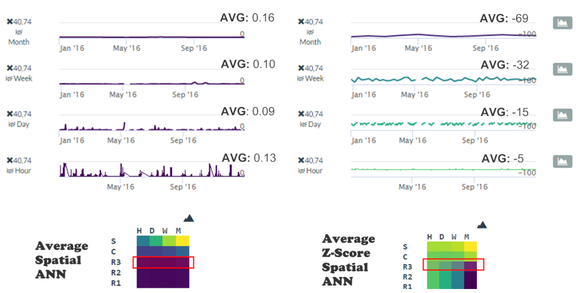


Fig. 8. The Spatial Average nearest neighbor and its z-score in four *st-LoDs*.

Figure 9 shows the ST-abstracts for all st-LoDs of datasets 2, 3 and 4. First of all, the occupation rate follows a similar pattern in all datasets. Dataset 3 stands out from the others regarding the collision rate. This occurs because the clusters in Dataset 3 were simulated within a spatial distance of 2 km, becoming more spatially clustered than in the other datasets. As a result, the co-occurrence of events on the same st-granule starts to occur "sooner", i.e., in finer st-LoDs when compared to the other datasets.

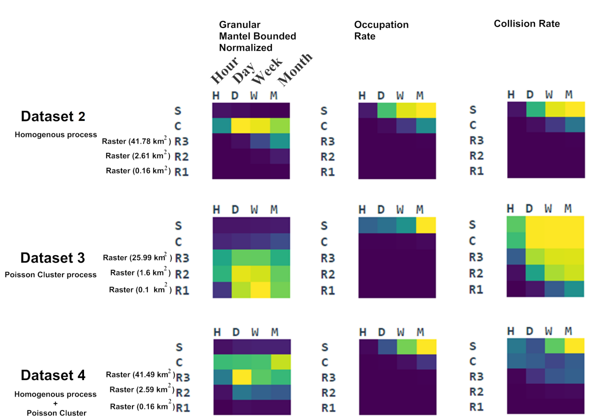


Fig. 9. Global Abstracts regarding Dataset 2, 3, 4.

As to Dataset 3, the GMBN highlights the following st-LoDs: (*i*) $(R1, Days)$; (*ii*) $(R1, Weeks)$; (*iii*) $(R2, Days)$; (*iv*) $(R2, Weeks)$. In this case, the values of spatiotemporal interaction are similar among the four st-LoDs, and therefore, any of the st-LoDs highlighted are good starting st-LoDs to detail our analyzes.

Dataset 4 is similar to Dataset 1, but contains an additional 3.000 events following a homogeneous model. In this case, the GMBN suggest the st-LoD - $(R3, Days)$, which is the st-LoD that better approaches the st-LoD in which the pattern is simulated.

Nevertheless, a single ST-Abstract should not be used in order to immediately guide our analyses for one or more st-LoDs. For example, in Dataset 1, the st-LoDs ($Counties$, $Days$) and ($Counties$, $Weeks$) are pointed as potential st-LoDs in which there might be spatiotemporal interaction. However, this dataset was generated following a homogeneous model. Actually, the values obtained regarding Average of the Spatial ANN for those st-LoDs are 4.2, 4.8, respectively, which discards at least the presence of clusters over time.

Experiments with respect to **contagious process** and **log-Gaussian Cox process** are not reported here due to space restrictions.

B. Real Datasets

Several phenomena were analyzed using the *VAST*. As opposed to synthetic datasets, we are not aware of possible patterns that those phenomena might contain. Here, we report the analysis made about wildfires that occurred in Portugal between 2001 and 2012. This phenomenon is described by a collection of 280.968 spatiotemporal events. The granularities-based model was used in order to model events of wildfires at different LoDs. These events were modeled through statements containing two arguments $wildfires(space, time)$.

To start our analysis we chose: (*i*) the GMBN; (*ii*) the Average of Spatial ANN; (*iii*) the Average of the z-score of the Spatial ANN; (*iv*) the Average of Temporal ANN; (*v*) the Average of the z-score of the Temporal ANN.

The Parallel Coordinates was used to simultaneously analyze the global abstracts chosen across all the *st-LoDs* as can be seen in Fig.10. Based on the abstracts chosen, we are interested in understanding *st-LoDs* in which (*i*) the phenomenon seems to be more clustered over time; (*ii*) the phenomenon seems to be more clustered over space; (*iii*) the *st-LoDs* where the spatiotemporal interaction of events seems to be better perceived. To conduct such analysis, we start by looking at the Parallel Coordinates displayed in Fig.10.

There are st-LoDs holding values close to zero with respect to Average Spatial ANN that simultaneously have quite negative values considering its z-score. Therefore, this kind of values have some resemblances with the ones obtained with Poisson cluster simulated datasets. As a result, at this point, we might say that wildfires in Portugal hardly follow a homogeneous model.

Several st-LoDs are holding values close to zero with respect to Average Temporal ANN but their z-scores are close

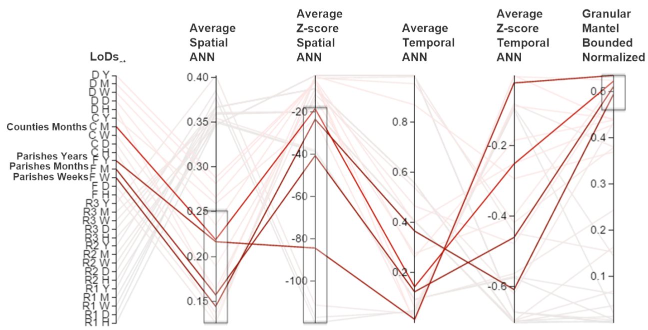


Fig. 10. Overview of wildfires in Portugal using Global Abstracts, Compact Spatial Abstracts and Compact Temporal Abstracts.

to zero, which means wildfires occurring on the same spatial granule are likely not close to each other in time, on average.

Furthermore, this information is telling us that probably we are not dealing with a phenomenon that follows a contagious pattern because events would be close in time as well.

Several st-LoDs have the spatiotemporal interaction among events measured by the GMBN above 0.4, which is similar to the values obtained in Poisson cluster simulated datasets. This reinforces the similarities of the wildfires in Portugal with the Poisson cluster pattern/process.

Based on the preliminary analysis, wildfires in Portugal seem to approach the Poisson cluster pattern. The parallel coordinates visualization was filtered in order to identify the suitable st-LoDs to confirm the previous hypothesis. We just considered st-LoDs with values below 0.25 (approximately) regarding the Average of the Spatial ANN. For the average of its z-score, we just considered values below -20 (approximately). Finally, the top four values of the GMBN were considered, which means values above 0.45 (approximately). The other coordinates (temporal average nearest neighbor and its z-score) were not filtered because there are no domain values that clearly points to clustered or dispersed events in time. From the filtering conducted, four st-LoDs were highlighted: $(Parishes, Weeks)$, $(Parishes, Months)$, $(Parishes, Years)$, $(Counties, Months)$ as can be seen in Fig.10.

To better understand how wildfires occur in space over time, *VAST* was used to display the Spatial ANN and its z-score in a scatter plot for the *st-LoDs* identified previously as can be seen in Fig.11.

The st-LoD - $(Parishes, Weeks)$ (chart on the bottom-right) is the one that better fits the Poisson cluster pattern/process. That is, in general, events occur near one another but there are a few times when events did not occur or occur in a dispersed way. Furthermore, in the st-LoD - $(Parishes, Weeks)$ there is a good tradeoff between the Spatial ANN and its z-score. In other words, there are many temporal granules in which the Spatial ANN's values are around 0.15 (trend toward clustering) and where their z-scores are quite negative (confirmation of clustering). This was further confirmed visually using the phenomenon representation with the st-LoD - $(Parishes, Weeks)$. This can be easily done because when an analyst select a data point (i.e., a temporal granule) from the scatter plot the spatiotemporal events are displayed in the phenomenon representation area.

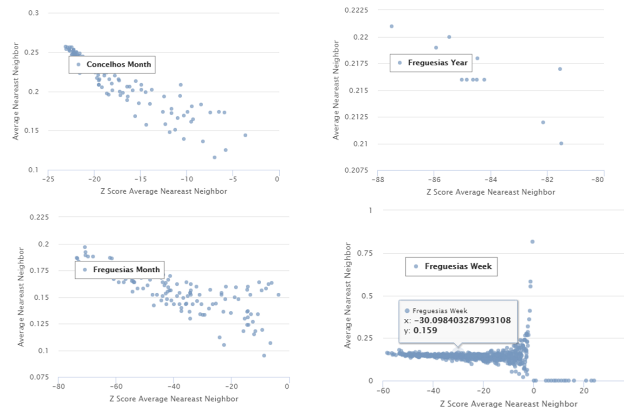


Fig. 11. The Spatial ANN and its Z-score displayed in four st-LoDs.

Similar analyses were conducted with other real data sets. Violent attacks against civilians in Africa[3] have similarities with a Poisson cluster pattern, where the analysis suggests that the pattern is better perceived in the st-LoDs containing the temporal granularities $Months$ or $Years$. Robberies in the City of Chicago[4] have similarities with a log-Gaussian Cox pattern, which is better perceived in the st-LoDs containing the spatial granularity $CommunitiesAreas$.

V. Related Work

Many approaches have been proposed in the literature to make analyses over spatiotemporal events.

From the 115 visualization methods surveyed by [20], just 19 were designed to display spatiotemporal data. A common characteristic in those methods is that they encode spatiotemporal data into visual representations at certain LoDs. In the project carried out by Lahouari et al. [21], 47 applications/visualization methods were assessed. Among the applications studied, 25% (12) were developed to analyze phenomena logged as spatiotemporal events. None of the approaches support data view at multiple spatial and temporal granularities (i.e., st-LoDs).

VA approaches also have been proposed in the literature. Some of these preview VA approaches support separate analyses of space and time anchored on descriptive statistics, most commonly considering one st-LoD at a time like the works [3], [4], [7], [22]–[25]. Other approaches support analyses that look for spatiotemporal patterns [2], [26]. However, these kinds of approaches follow analyses based on a single LoD, and in some cases, they are developed for the detection

[3]Data are available at https://www.acleddata.com/

[4]Data are available at https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data

and exploration of a particular spatiotemporal pattern in a particular domain application [27]–[29].

Nevertheless, VA approaches working across several LoDs are starting to be developed. Goodwin et al. [10] propose a framework for analyzing multiple variables across spatial LoDs and geographical locations; Robinson et al. [11] developed a VA approach, called STempo, which computes temporal patterns in multiple temporal LoDs in spatiotemporal events. Finally, Swedberg et al. [30] developed a VA approach to help users in the detection of calendar related periodicity in spatiotemporal events. This work allows the analysis at multiple spatial LoDs and temporal LoDs. However, the number of the spatial LoDs that we can analyze, simultaneously, are limited to two, and the analyses supported are only based on descriptive statistic $\mathbb{COUNT}$.

To the best of our knowledge, there are no approaches that work across several spatial and temporal LoDs, working with space and time together, and therefore, looking for spatiotemporal patterns at different spatiotemporal LoDs.

VI. Conclusions and Future Work

To enhance the analyses over spatiotemporal events, we first propose to move from a single user-driven LoD to a multiple LoDs analysis approach, providing the user with an understandable high-level overview of the underlying structure of the phenomenon for each LoD. This approach can provide several hints about the different facets of spatiotemporal events and a first insight on the presence or absence of patterns at particular LoDs. According to his analytical goal and domain knowledge, the user may be able to better guide his analysis thus avoiding an information overload.

To conduct analyzes in this new mindset, *VAST* was developed. The tool allows to visually inspect hints about the absence or presence of different kinds of spatiotemporal patterns at multiple LoDs, simultaneously, following a coordinated strategy among the visualizations provided. To the best of our knowledge, there is no other prototype or application that supports analyses over spatiotemporal events at multiple LoDs, simultaneously, following the VA Mantra.

Experiments were conducted with two types of datasets describing spatiotemporal events: (i) synthetic datasets; (ii) real datasets. In them, *VAST* was able not only to provide an overview of the presence or absence of different spatiotemporal patterns but also suggest appropriate spatiotemporal LoDs that allow us to better perceive the corresponding patterns.

Future work can be directed for the development of heuristics to suggest automatically LoDs to analyze the data are needed and should be a priority because if the number of abstracts grows considerably it might be overwhelming to the user. This issue relates to the learning curve. Each abstract looks for a feature or pattern which frequently is expressed in terms of a range of values. According to the value, it means one thing or the other. Thus, a user needs to get familiar with the abstracts and their interpretation. Requiring a user to memorize all the abstracts and their interpretation might be overwhelming, specially if we consider the joint interpretation of abstracts. So again, heuristics to suggest automatically LoDs are needed.

Appendix A
Implementation Details

VAST is a web-based application, built on top of a RESTful Web services and a relational database (PostgreSQL + PostGis). The user interface (UI) is implemented using HTML5+CSS+JavaScript, and uses WebGL to display efficiently thematic maps [31]. The abstracts implemented are summarized in the next tables.

TABLE II
Abstracts usable as global, temporal or spatial.

Measure	Description
Collision Rate (%)	Percentage of spatiotemporal granules with events, where atom collisions exists
Occupation Rate (%)	Percentage of granules with events

TABLE III
Global only Abstracts.

Measure	Description
Granular Mantel Bounded and Normalized	Measures the spatiotemporal interaction
Reduction Rate (%)	Measures the reduction of atoms used
Average Atoms in spatiotemporal granules (%)	Measures the average of atoms indexed by spatiotemporal granules

References

[1] A. S. Hering, C. L. Bell, and M. G. Genton, "Modeling spatio-temporal wildfire ignition point patterns," *Environmental and Ecological Statistics*, vol. 16, no. 2, pp. 225–250, 2009.

[2] R. Maciejewski, S. Rudolph, R. Hafen, A. Abusalah, M. Yakout, M. Ouzzani, W. S. Cleveland, S. J. Grannis, and D. S. Ebert, "A visual analytics approach to understanding spatiotemporal hotspots," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 16, no. 2, pp. 205–220, 2010.

[3] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, "Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 19, no. 12, pp. 2149–2158, 2013.

[4] I. Cho, W. Dou, D. X. Wang, E. Sauda, and W. Ribarsky, "Vairoma: A visual analytics system for making sense of places, times, and events in roman history," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 1, pp. 210–219, 2016.

[5] S. Shekhar, Z. Jiang, R. Y. Ali, E. Eftelioglu, X. Tang, V. Gunturi, and X. Zhou, "Spatiotemporal data mining: A computational perspective," *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2306–2338, 2015.

[6] L. Zhang, A. Stoffel, M. Behrisch, S. Mittelstadt, T. Schreck, R. Pompl, S. Weber, H. Last, and D. Keim, "Visual analytics for the big data era—A comparative review of state-of-the-art commercial systems," in *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on*. IEEE, 2012, pp. 173–182.

[7] L. Lins, J. T. Klosowski, and C. Scheidegger, "Nanocubes for real-time exploration of spatiotemporal datasets," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 19, no. 12, pp. 2456–2465, 2013.

TABLE IV
SPATIAL ONLY ABSTRACTS.

Measure	Description
BRAY-CURTIS SIMILARITY FOR ATOMS	Calculates the similarity based on the counts of atoms, between consecutive temporal grains
BRAY-CURTIS SIMILARITY FOR SYNTHESIS	Calculates the similarity based on the number of granular synthesis, between consecutive temporal grains
CORRELATION INDEX FOR ATOMS	Correlation between the number of atoms of consecutive temporal grains
CORRELATION INDEX FOR SYNTHESIS	Correlation between the number of granular synthesis of consecutive temporal grains
DICE SIMILARITY (BINARY)	Dice index (event / no event) between consecutive temporal grains
JACCARD SIMILARITY (BINARY)	Jaccard index (event / no event) between consecutive temporal grains
GOWER SIMILARITY (BINARY)	Similarity (event / no event) between consecutive temporal grains
MORAN'S I	Calculates the spatial autocorrelation among nearby locations, given a domain specific variable
SPATIAL SCOPE	Measures the spatial extent
SPATIAL CONSECUTIVE DISTANCE BETWEEN CENTERS OF MASS	Measures the distance between consecutive centers of mass

TABLE V
SPATIAL OR TEMPORAL ABSTRACTS.

Measure	Description
NEAREST NEIGHBOR (NN)	Measures the level of clustering
Z-SCORE NEAREST NEIGHBOR (Z-NN)	Measures the z-score of the level of NN
CENTER'S MASS POSITIONING	Measures the position of the centers of mass
FREQUENCY RATE (%)	Percentage of events happened in granules

[8] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, "Visual Analytics: Definition, Process, and Challenges," in *Information Visualization*, ser. Lecture Notes in Computer Science, A. Kerren, J. Stasko, J.-D. Fekete, and C. North, Eds. Springer Berlin Heidelberg, 2008, vol. 4950, pp. 154–175.

[9] M. Sips, P. Köthur, A. Unger, H.-C. Hege, and D. Dransch, "A visual analytics approach to multiscale exploration of environmental time series," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 18, no. 12, pp. 2899–2907, 2012.

[10] S. Goodwin, J. Dykes, A. Slingsby, and C. Turkay, "Visualizing Multiple Variables Across Scale and Geography," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 22, no. 1, pp. 599 – 608, 2016.

[11] A. C. Robinson, D. J. Peuquet, S. Pezanowski, F. A. Hardisty, and B. Swedberg, "Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data," *Cartography and Geographic Information Science*, pp. 1–13, 2016.

[12] R. Silva, J. M. Pires, M. Y. Santos, and N. Datia, "Enhancing exploratory analysis by summarizing spatiotemporal events across multiple levels of detail," in *Geospatial Data in a Changing World, Selected papers of the 19th AGILE Conference on Geographic Information Science*, ser. Lecture Notes in Geoinformation and Cartography, T. Sarjakoski, M. Y. Santos, and T. L. Sarjakoski, Eds. Springer, 2016. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-33783-8_13

[13] J. T. Yao, A. V. Vasilakos, and W. Pedrycz, "Granular computing: perspectives and challenges," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1977–1989, 2013.

[14] R. A. Silva, J. M. Pires, and M. Y. Santos, "A granularity theory for modelling spatio-temporal phenomena at multiple levels of detail," *International Journal of Business Intelligence and Data Mining*, vol. 10, no. 1, p. 33, 2015.

[15] R. A. Silva, J. M. Pires, M. Y. Santos, and R. Leal, "Aggregating Spatio-temporal Phenomena at Multiple Levels of Detail," in *AGILE 2015*. Springer Science \mathplus Business Media, 2015, pp. 291–308.

[16] E. Gabriel, B. Rowlingson, and P. Diggle, "stpp: an r package for plotting, simulating and analyzing spatio-temporal point patterns," *Journal of Statistical Software*, vol. 53, no. 2, pp. 1–29, 2013.

[17] D. Ebdon, *Statistics in geography*, 1985.

[18] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[19] N. Mantel, "The detection of disease clustering and a generalized regression approach," *Cancer research*, vol. 27, no. 2 Part 1, pp. 209–220, 1967.

[20] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of time-oriented data*. Springer Science & Business Media, 2011.

[21] K. Lahouari, B. Jean-Yves, D. Paule-Annick, M. Hélène, and S.-M. Cécile, "Représenter les dynamiques des territoires : un état des lieux, de nouveaux enjeux," 2014. [Online]. Available: http://www.map.cnrs.fr/jyb/puca/

[22] S. Kisilevich, M. Krstajic, D. Keim, N. Andrienko, and G. Andrienko, "Event-based analysis of people's activities and behavior using flickr and panoramio geotagged photo collections," in *Information visualisation (IV), 2010 14th international conference*. IEEE, 2010, pp. 289–296.

[23] A. Malik, R. Maciejewski, T. F. Collins, and D. S. Ebert, "Visual analytics law enforcement toolkit," in *Technologies for Homeland Security (HST), 2010 IEEE International Conference on*. IEEE, 2010, pp. 222–228.

[24] A. M. MacEachren, A. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. Blanford, "Senseplace2: Geotwitter analytics support for situational awareness," in *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*. IEEE, 2011, pp. 181–190.

[25] G. Andrienko, N. Andrienko, H. Bosch, T. Ertl, G. Fuchs, P. Jankowski, and D. Thom, "Thematic patterns in georeferenced tweets through space-time visual analytics," *Computing in Science & Engineering*, vol. 15, no. 3, pp. 72–82, 2013.

[26] D. Guo, J. Chen, A. M. MacEachren, and K. Liao, "A visualization system for space-time and multivariate patterns (vis-stamp)," *IEEE transactions on visualization and computer graphics*, vol. 12, no. 6, pp. 1461–1474, 2006.

[27] D. Wang, W. Ding, H. Lo, M. Morabito, P. Chen, J. Salazar, and T. Stepinski, "Understanding the spatial distribution of crime based on its related variables using geospatial discriminative patterns," *Computers, Environment and Urban Systems*, vol. 39, pp. 93–106, 2013.

[28] J. Chae, D. Thom, H. Bosch, Y. Jang, R. Maciejewski, D. S. Ebert, and T. Ertl, "Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition," in *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on*. IEEE, 2012, pp. 143–152.

[29] D. Thom, H. Bosch, S. Koch, M. Wörner, and T. Ertl, "Spatiotemporal anomaly detection through visual analysis of geolocated twitter messages," in *Pacific visualization symposium (PacificVis), 2012 IEEE*. IEEE, 2012, pp. 41–48.

[30] B. Swedberg and D. Peuquet, "Perse visual analytics for calendar related spatiotemporal periodicity detection and analysis," *GeoInformatica*, pp. 1–21, 2016.

[31] D. Cardoso, R. Alves, J. M. Pires, F. Birra, and R. A. Silva, "Gisplay - Extensible Web API for Thematic Maps with WebGL," in *Computational Science and Its Applications–ICCSA 2017*. Springer, 2017.