



A progressive development of a visual analysis interface of climate-related VGI

Carlo Navarra¹ · Katerina Vrotsou² · Tomasz Opach³ · Almar Joling⁴ · Julie Wilk¹ · Tina-Simone Neset¹

Received: 30 October 2020 / Accepted: 4 September 2021 / Published online: 3 October 2021
© The Author(s) 2021

Abstract

This paper describes the progressive development of three approaches of successively increasing analytic functionality for visually exploring and analysing climate-related volunteered geographic information. The information is collected in the CitizenSensing project within which urban citizens voluntarily submit reports of site-specific extreme weather conditions, their impacts, and recommendations for best-practice adaptation measures. The work has pursued an iterative development process where the limitations of one approach have become the trigger for the subsequent ones. The proposed visual exploration approaches are: an initial map application providing a low-level data overview, a visual analysis prototype comprising three visual dashboards for more in-depth exploration, and a final custom-made visual analysis interface, the CitizenSensing Visual Analysis Interface (CS-VAI), which enables the flexible multifaceted exploration of the climate-related geographic information in focus. The approaches developed in this work are assessed with volunteered data collected in two of the CitizenSensing project's campaigns held in the city of Norrköping, Sweden.

Keywords Geographic visualization · Information visualization · Visual data exploration · Volunteered geographic information

Introduction

A growing interest in citizen sensing implies a growing production of volunteered geographic information (VGI) (Goodchild 2007). Non-professionals use sensors, often low-cost or even self-built, to voluntarily collect data that concern environmental aspects they care about. Although

such voluntarily collected geographic information enables citizens and stakeholders to get insights into issues for which geospatial databases do not exist, the inaccuracy of VGI is often cited as a barrier to its wider use in the decision-making process (Goodchild 2007). Meanwhile, studies have shown that participatory mapping projects can produce more accurate data than official initiatives since participants contribute with unique local knowledge. Efficient approaches are thus needed that enable the comprehensive analysis of VGI with respect to its context and characteristics, but that

This article is part of a Topical Collection in Environmental Earth Sciences on “Visual Data Exploration”, guest edited by Karsten Rink, Roxana Bujack, Stefan Jänicke, and Dirk Zeckzer.

✉ Carlo Navarra
carlo.navarra@liu.se

Katerina Vrotsou
katerina.vrotsou@liu.se

Tomasz Opach
tomasz.opach@ntnu.no

Almar Joling
almar.joling@royalhaskoningdhv.com

Julie Wilk
julie.wilk@liu.se

Tina-Simone Neset
tina.neset@liu.se

¹ Centre for Climate Science and Policy Research, Department for Thematic Studies Environmental Change, Linköping University, Linköping, Sweden

² Division for Media and Information Technology, Department of Science and Technology, Linköping University, Norrköping, Sweden

³ Department of Geography, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

⁴ Deltares, Delft, The Netherlands

also allow for assessments to be made concerning its accuracy, reliability, and relevance to the domain research. Visual data exploration can serve as such an approach providing insight into the data and leading to increased understanding of its advantages and shortcomings.

The primary aim of this study is to investigate the potential of visual exploration, defined as the process of using visualization to gain insight into data, in the context of VGI related to climate change impacts and adaptation. As the term exploration implies, the main focus of such a process is the identification of unknowns and the generation of new hypotheses to be investigated further either through continued visual exploration or by means of automatic techniques (Keim 2001). Visual data exploration is interactive and intuitive and allows the user's domain knowledge to be seamlessly incorporated in the analysis. When the data to be explored have a geographic reference, these are commonly represented on interactive map displays enabling users to reveal unknown geographical patterns and spatial relationships in the data (MacEachren 1994), and thus to gain additional knowledge about investigated phenomena.

Having this as a starting point and following Shneiderman's visual information seeking mantra of "overview, zoom and filter, and details-on-demand" (Shneiderman 1996), this study explores the potential of visual exploration of climate-related VGI generated by the participants of the CitizenSensing¹ project. To achieve this, visual exploration and analysis approaches of progressively increasing sophistication and analytical functionality have been iteratively implemented and tested with real data in the course of the project. This iterative process has resulted in an initial map application providing a low-level data overview, an Apache Superset² prototype comprising three visual dashboards for more in-depth exploration, and a final custom-made visual analysis interface, CitizenSensing Visual Analysis Interface (CS-VAI), which enables the flexible multifaceted exploration of the climate-related geographic data in focus. The data used to exemplify the approaches have been collected in two of the CitizenSensing project's campaigns held in the city of Norrköping, Sweden.

The main contribution of this work is an extensive elaboration of the design and development process through a set of visual exploration tools that facilitate visual analysis of VGI related to extreme weather conditions and climate adaptation. As the final outcome of this process, we propose CS-VAI, a tool for visual data exploration whose main goal is to allow the seamless exploration of the temporal, spatial, and attribute space of VGI.

The remaining of this article is structured as follows. First, a short review of related work is presented followed by a description of the CitizenSensing project and the data collected within it. Subsequently, a description of the three visual exploration approaches developed in this work is provided. The description is divided into three separate subsections corresponding to the iterative process followed in the project. Following this, a section with example use cases of how the proposed tools can be used for exploring different aspects of the data at hand is presented. The article concludes with a section discussing considerations and biases when working with VGI in general as well as particular limitations and considerations specific to our work. Finally, conclusions and future work are outlined.

Related work

This work is concerned with the visual exploration of spatio-temporal user-generated climate-related data that include reports of weather events, their impacts, and/or recommendations of appropriate adaptation measures. Therefore, it relates to research in: (1) the use of visualization for knowledge discovery, (2) visualization of environmental data, and (3) visualization of VGI.

The use of visualization for knowledge discovery

The importance of interactive visualization and the human-in-the-loop in data analysis is underlined by the large body of research on visual exploration and analytics interfaces (Wang et al. 2016). Several visualization pipelines have been proposed focusing on different analysis aspects and allowing various levels of user involvement (e.g. van Wijk (2005), Munzner (2009)). Multitudes of visual exploration interfaces have been developed that facilitate flexible interactive analyses and knowledge discovery. The common cornerstones of these are commonly:

- the mapping of data values to visual elements,
- the use of coordinated multiple views (CMV) to facilitate multifaceted data exploration (e.g. Roberts (2007), Dörk et al. (2008)),
- the use of interaction techniques that provide flexibility of analysis to the user (Yi et al. 2007).

Although visual exploration interfaces differ in terms of content, offered functionality and visual complexity, they have proven of essential value for knowledge discovery and hypothesis formulation in many domains, including environmental studies (Helbig et al. 2017).

¹ <http://citizensensing.eu>

² <https://superset.apache.org>

Visualization of environmental data

A number of visual exploration interfaces have emerged promoting the analysis of environmental data. For instance, Li et al. (2014) integrated three visualization techniques to investigate climate data where a novel technique named “global radial map” served as the main visualization view. In turn, Jänicke and Scheuermann (2014) illustrated the use of the open-source library GeoTemCo for visualizing environmental data referring to three themes: biodiversity, global warming and relationships between weather phenomena, and natural hazards. A geographic visualization approach was proposed by Neset et al. (2016) to address challenges in communication of climate information to lay audiences, particularly related to climate scenarios, impacts, and action alternatives. A web-based framework for collaborative visualization and analysis was presented by Lukaszczuk et al. (2015); they used a dataset of water scarcity in Africa to exemplify the utilization of visual interfaces for collaborative environmental data exploration on distributed desktop and mobile devices. Neset et al. (2019) designed an interactive visualization tool to support dialogues with water management and ecosystem stakeholders. Finally, Jänicke (2019) presented a system to visually explore spatial data on animal and plant species that are threatened with extinction at the European level. In this study, a series of highly interactive visual interfaces has been developed to facilitate dynamic query construction to investigate geospatial, species-related features and developments.

All studies mentioned above employ visualization techniques to facilitate data comprehension by the users. The plethora of implemented solutions and interface configurations makes it challenging to conclude about their commonalities. Even within one dataset, a number of different visual interfaces can be designed to emphasize various data aspects and thus support different user tasks. Therefore, a conceptual stage where adequate visualization and interaction techniques are extensively considered seems to be of essential importance to any development process where a visual interface is to be implemented.

Visualization of volunteered geographic information

With the rise of platforms encouraging the collection of user-generated data, the number of systems concerned with the visualization of VGI (Goodchild 2007) has been increasing. A visual analytics system for the exploration of environmental data collected via the citizen science platform *enviroCar* was proposed by Häußler et al. (2018). In this work, the authors experiment with various visualization techniques to gain insight into the data and, for instance, use a clock metaphor to display changes over time of several data attributes.

In another study, Seebacher et al. (2019) propose an adaptive workspace consisting of five views that facilitate prediction and visual analysis of urban heat islands based on volunteered data from a private weather station network. Tenney et al. (2019) describe a crowd sensing system that captures geospatial social media topics and visualizes the results in multiple views.

Some works that are close to our interests in their scope and focus on the value of the visual exploration of VGI include the SensePlace2 system (MacEachren et al. 2011). SensePlace2 combines visual exploration of social media data (Twitter) with a place–time–concept framework to support geographically grounded situational awareness. This work, similar to ours, puts focus on the specific data attributes of interest in the development and assessment of the potential of their proposed interface. For SensePlace2, the focus is on place, time, and theme (i.e. textual context) to understand and monitor an emerging situation reported through Tweets, while we are focusing on the place, time, image content as well as associated weather parameters to map and better understand weather conditions and their impacts to personal comfort. Also, Dykes et al. (2008) have investigated the potential of visualization for exploring the complexity of VGI and extracting local knowledge. For achieving this, they have experimented with visual exploration approaches on a collection of volunteered geography-related photographs (‘geographs’) for understanding notions of place.

The three areas outlined in this section indicate that visual exploration of citizen-generated data by means of visual interfaces is becoming a common practice for obtaining a better understanding of local environmental conditions. As conditions can be perceived differently by various citizens, it is of primary importance for a sustainable urban development to consider the diverse site-specific information related to urban climate adaptation reported by individuals.

The CitizenSensing project and its climate-related VGI

The work presented in this paper is part of the CitizenSensing project. The project’s main goal is to develop and evaluate a Participatory Risk Management System (PRMS) that allows urban citizens to voluntarily report, by means of a mobile web application, information about site-specific extreme weather conditions, their impacts, and recommendations for best-practice adaptation measures. The project aims to assess whether and how the system can increase citizens’ engagement and contribute to urban resilience. The particular targeted weather phenomena, e.g. events with high and low temperature or heavy rainfall, are decided upon with relevant urban stakeholders and according to the specific

context of different pilot studies. The CitizenSensing mobile web application aims to facilitate a high level of interactivity by engaging citizens, both as providers of place-specific data on weather events, as well as receivers of recommendations on how to respond. The resulting system will provide an integrated platform to support local stakeholders and citizens in increasing their ability to make informed decisions related to climatic risks and increase urban resilience.

The CitizenSensing PRMS is composed of a mobile web application, a sensor network, and a database. The mobile web application enables citizens to submit reports on observed weather events and their impacts as well as seasonal observations. The voluntarily submitted reports are stored in a database comprising: geographical coordinates, date and time, weather event type, climate impact, personal level of comfort, picture (optional), comment (optional), as well as the type of operating system and web browser that was used for the submission. When a report is submitted, no personal information is stored except for the users' geographical position, provided that they have granted access to this information. Depending on availability, the system also integrates real-time weather parameter data, such as temperature, air pressure, humidity, and precipitation, from the Netatmo weather station network.³

The data used in this study for testing the potential of visual exploration, through the proposed approaches, are collected in two of the CitizenSensing project's campaigns in the city of Norrköping, Sweden: (i) a campaign with home-care staff and (ii) a campaign with students of two classes of secondary school. Data were collected from assigned teams instead of single individuals to preserve privacy. To encourage participants to be actively involved in the campaigns, points were given for each uploaded report, image, and comment. Project members held meetings with the homecare staff at the start and end of the campaign, while the students were visited four times during the campaign to follow up on questions and to encourage participation.

Visual exploration approaches of climate-related VGI

The climate-related VGI considered in this project comprises volunteered geo-referenced reports (in the following referred to only as reports) on observed weather events and their impacts collected during the CitizenSensing campaigns. We have investigated the use of visual exploration approaches to successively enhance insight, both with regard to the reports' content and their distribution characteristics and reliability. To this end, three visual interfaces of

progressively increasing analytical functionality have been developed over three iteration phases of the project:

1. The CitizenSensing Web Portal providing a low-level overview of the data in a map display context.
2. An Apache Superset prototype, with three visual dashboards experimenting with various interface configurations, for more in-depth spatio-temporal data exploration.
3. The CitizenSensing Visual Analysis Interface (CS-VAI) enabling comprehensive, multifaceted data exploration in search of spatio-temporal patterns.

In the following, the three visual interfaces are described in detail in the context of their successive advancement during the projects' development process.

The CitizenSensing Web Portal

The CitizenSensing Web Portal (Fig. 1) enables a flexible overview of the reports collected in the CitizenSensing project. The Web Portal displays the reports on a map and features tailored user interaction to support the filtering and initial assessment of the data. A user can perform a low-level data exploration that encompasses basic user tasks from the analytic task taxonomy proposed by Amar et al. (2005), for example, value retrieval or filtering. Hence, users can retrieve values of specific reports and filter reports by location, time range, or user-id. Furthermore, data can be filtered by weather event type or by users' reported personal level of comfort. Filtered reports are presented on a map display and differentiated by map symbols showing their corresponding weather event types. Apart from the map display, filtered reports are listed in a panel to the left of the map. If a report includes a picture or a textual comment related to the reported weather event, this information is displayed in the list view. A report can be selected from the list view or by clicking an item on the map.

The Web Portal has been designed for desktop displays. It has been developed using open-source JavaScript APIs such as OpenLayers for web mapping, Bootstrap for responsive layout, and jQuery for user interaction. Two versions of the Web Portal are available: (1) a publicly available version which shows reports without detailed information such as comments and pictures and (2) a password protected version which shows comments and images and is thus used only within the project for analysis and evaluation of the reported data.

The Web Portal enables the user to get an overview of the spatial distribution and gain an initial understanding of the climate-related VGI considered in this work. This is, however, done in a manual and somewhat cumbersome manner without visual cues to guide the exploration. The

³ <https://dev.netatmo.com/apidocumentation/weather>

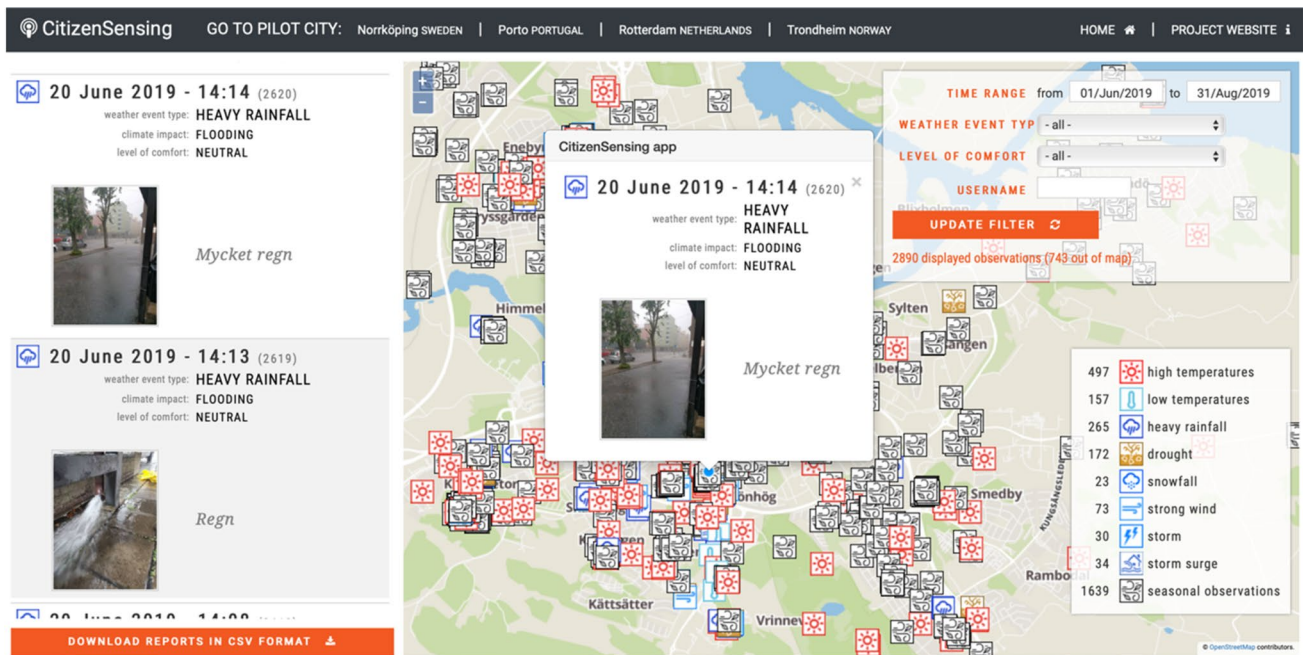


Fig. 1 The CitizenSensing Web Portal for low-level data overview. It comprises a map view displaying the reported events, a side panel displaying the attributes and the associated image for each of the visible events on the map, and a filtering panel

user explores the content of the data by querying individual reports for details on demand and by applying filtering operations to inspect subsets that display specific attributes or occur in specific intervals. This created the need for a more advanced exploration that would allow for in-depth analysis of the data which led to the Apache Superset visual analysis prototype.

Apache Superset prototype

An initial visual analysis prototype was designed to allow more elaborate exploration of the CitizenSensing project's VGI. The prototype was built using the Apache Superset software, an open-source data exploration and visualization platform for business intelligence. Apache Superset provides an intuitive interface for creating visual dashboards combined with an SQL Interactive Development Environment (IDE). The platform offers a number of visualization techniques, based on D3.js (Bostock et al. 2011) or deck.gl (Wang 2017) that can be dynamically linked for creating interactive visual dashboards.

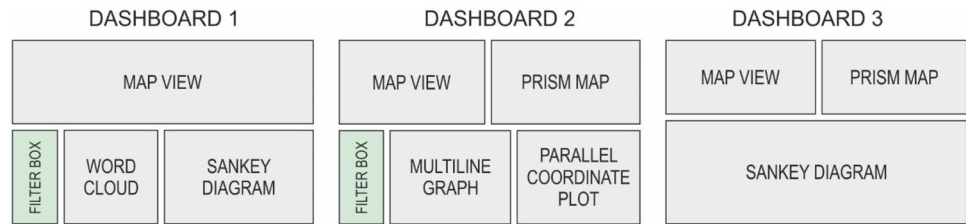
The Apache Superset visual analysis prototype (in the following referred to as Superset prototype) comprises three visual dashboards tailored to support specific user tasks depending on the use case. To this end, six visualization techniques were combined in different configuration designs

to display various aspects of the data from two separate user-generated datasets. The visualization techniques used are as follows:

1. Map view to display positions of reported weather events where the colour of map symbols encodes weather event types.
2. Prism map to display temperature values from the Netatmo sensor network. The heights and the colours of the hexagonal prisms correspond to the average temperature calculated for each hexagonal cell.
3. Sankey diagram that displays the linkages between weather event type, impact type, and personal level of comfort collected through the weather event reports. The size of the bands represents the number of linkages between the reported data.
4. Parallel coordinate plot to display weather parameters from the Netatmo sensor network: temperature, air pressure, and humidity.
5. Multiline graph representing aggregated minimum, maximum, and average temperature values. The data are retrieved from the Netatmo sensor network.
6. Word cloud to display frequent terms submitted in the comments of the reported weather events. The size of the word is proportional to the number of reports.

The three dashboards combine the above visualization techniques in the configurations shown in Fig. 2 and described further in the “Use cases” section. All

Fig. 2 The three visual dashboards designed to experiment with various interface configurations for in-depth exploration of the CitizenSensing climate-related VGI



visualization components in each dashboard are linked by time attribute and weather event type and can be navigated through a filter box. Although the dashboards offer data filtering, no data exploration mechanisms by means of sophisticated interaction techniques are available.

The Superset prototype brought considerable advancements in the available functionality for exploring the climate-related VGI at hand in comparison with the CitizenSensing Web Portal, as described in the following use case examples. However, several substantial limitations were also revealed. The main finding was that the combinations of visualization techniques implemented in the different dashboards allowed a user to get insight into the relations between spatial, temporal, and additional qualitative (i.e. weather event type, climate impacts) attributes of the collected data. However, this insight was restricted by the composition of each dashboard in terms of visualization techniques and data attributes to be used. Moreover, interaction with the data occurred only through the filter box. This limits considerably the exploration flexibility and can potentially impede the analytic flow of thought of the user. Based on these limitations and on the feedback received from the domain experts during the demonstrations of the dashboards, a decision was taken to develop a custom-made visual analysis interface that would allow flexibility of analysis of all aspects of the data in an integrated environment.

The CitizenSensing Visual Analysis Interface

The CitizenSensing Visual Analysis Interface (CS-VAI) was designed to enable more flexible exploration and achieve higher levels of sophistication in the visual analysis of the CitizenSensing project's VGI. CS-VAI is a responsive web-based visual analysis interface with four linked views allowing exploration in geographical, temporal, and attribute data space and details-on-demand functionalities. The interface was developed using the web application framework Angular⁴ in combination with D3.js (Bostock et al. 2011) and deck.gl (Wang 2017) APIs. The four linked views composing CS-VAI are: a map view, a scatter plot, a multiline

graph, and an image clouds view (Fig. 3a–d, respectively), all custom-designed for the interface.

The map view displays reports as map symbols colour-coded by level of comfort or climate impact associated to the reported weather events. Users can pan and zoom the map to filter the reports by a geographical range. Such geographical filtering is then reflected on all other views.

In the scatter plot, each point corresponds to a report and is colour-coded by level of comfort or climate impact. The x-axis of the plot reflects reporting time, whereas the y-axis reflects weather event type. The user can filter the data by level of comfort through clicking on the colour legend of the plot, or by weather event type through clicking on the event-type labels of the y-axis. Furthermore, a focus + context visualization paradigm (Cockburn et al. 2009) is applied on the time axis. An overview plot of the entire time range of the data is provided at the bottom of the plot on which the user can select a specific period to focus on (see Fig. 3b). The scatter plot is then updated to show only the reports within the selected period, while the overview plot continues to show the distribution of reports across the entire time range. The applied temporal filtering is reflected on all other views.

In the multiline graph, the user can enable/disable line graphs that show sensor measurements such as air temperature, humidity, and pressure retrieved from the Netatmo sensor network for the Norrköping municipality. The default view (Fig. 3c) displays the: average (red line), maximum, and minimum values (blue lines) aggregated from the sensor measurements. As a reference, in the air temperature graph, air temperature readings for the city of Norrköping (orange line) retrieved from the Swedish Meteorological and Hydrological Institute were also included.

Lastly, the image clouds view displays the images submitted with the reports. The visualization uses a force-directed layout with multifoci clustering, where each focus point identifies a unique attribute value of the reports. The d3-force module is employed for this that implements a force simulation with a collision force, which uses a buffer radius in combination with a repulsive force to prevent that the nodes overlap, and a modified force cluster function that takes into account the cluster centroids according to the selected attribute and uses them as gravitational points. In the image clouds view, images can be arranged by weather event type, level of comfort, or by their colour content in

⁴ <https://angular.io>

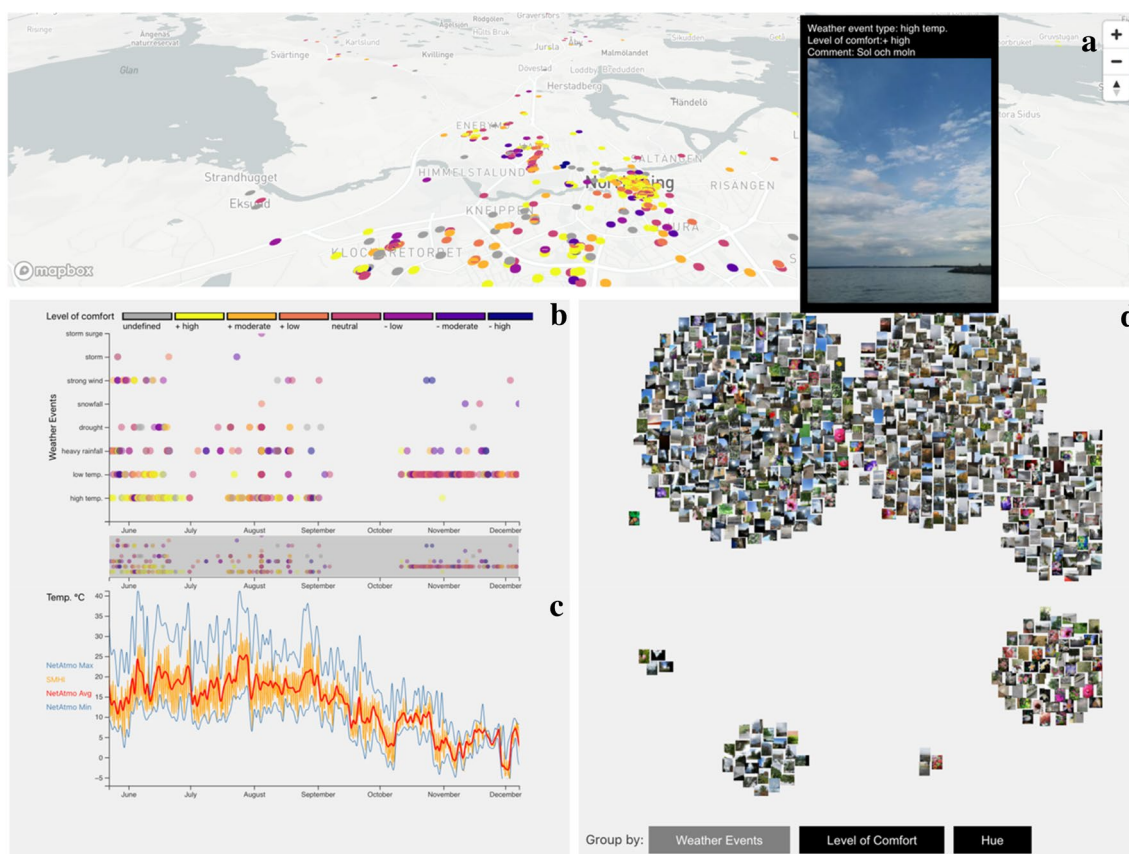


Fig. 3 The CitizenSensing Visual Analysis Interface (CS-VAI) composed of four linked views: **a** a map view displaying the position of reports, **b** a temporal scatter plot displaying the weather events across

the explored time period, **c** a multiline graph displaying collected weather parameters over the same time period, and **d** an image clouds view displaying the reported images using different arrangements

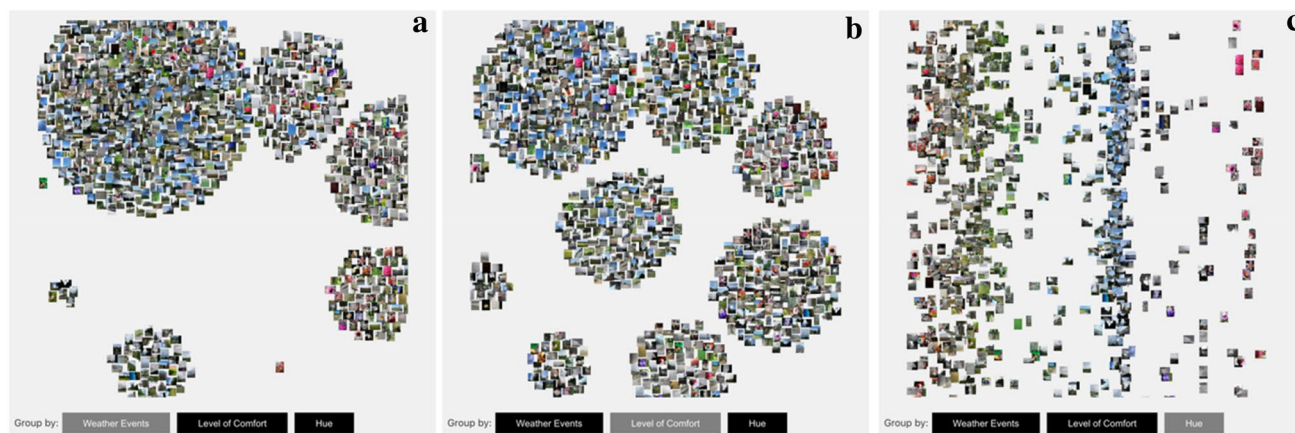


Fig. 4 The three arrangement options of the CS-VAI image clouds view: **a** by weather event type, **b** by level of comfort, and **c** by image colour content

terms of their hue (Fig. 4a–c, respectively). The latter is achieved by calculating the average colour for each image and then using their hue values to sort them along the x-axis.

An essential advantage of CS-VAI's interface configuration is that after the selection of a geographic area (in the map view), and selection of a time range and a considered attribute (in the scatter plot), the user can juxtapose the

observations with the actual weather parameters provided in the multiline graph. Furthermore, filtered reports can be explored in terms of corresponding images. The latter can be easily accessed by clicking on particular image thumbnails. Moreover, grouping images into clouds can enable further insight into their content.

Use cases

Three use cases are presented to exemplify the exploration made possible through the proposed approaches and discuss opportunities and challenges. The data in these use cases were collected during two campaigns of the CitizenSensing project as described in the section “The CitizenSensing project and its climate-related VGI.”

The exploration has been based on the Triad framework for the representation of spatio-temporal dynamics proposed by Peuquet (1994). The framework allows the user to pose questions relating to three basic elements of spatio-temporal data: where (space/location), when (time), and what (objects), and explore aspects of each in terms of the other two. The objects, in the case of this data, correspond to any of the reports’ qualitative attributes, i.e. weather event type, personal level of comfort, climate impact, textual comment, and submitted image. The questions supported in the Triad framework (Peuquet 1994), and thus considered in the following use cases, concern how a user can:

- (1) Explore objects that are present at a given location at a given time (*when + where* → *what*).
- (2) Explore times that a given object occupied a given location (*where + what* → *when*).
- (3) Explore locations occupied by a given object at a given time (*when + what* → *where*).

For each of the questions above, examples of if and how such an exploration can proceed using the proposed approaches are discussed.

when + where → what: Exploration of visual and textual contributions for selected weather event types

The first use case is concerned with exploring the reports’ qualitative attributes (*what*) in terms of their geographical locations (*where*) and their time range (*when*), i.e. under certain spatial and temporal selections. Such an exploration can be of interest for getting an overview and assessing the content of the reports. In the CitizenSensing project, the collected VGI allows the researchers to provide relevant feedback to participants during an ongoing campaign, to identify whether campaign instructions (e.g. observations

specifically directed to a certain impact) are followed, or to provide a synthesis to interested parties, so they could analyse whether the provided information is relevant to spatial planning and climate communication efforts and/or indicates a specific risk in an area. Consequently, it becomes interesting to analyse the visual and textual contributions of the submitted reports (i.e. the images and their corresponding comments) for different weather event types.

The CitizenSensing Web Portal (Fig. 1) enables a quick assessment of the data in relation to this task. The Web Portal offers functionality to filter the reports by time (through selection options) and space (through navigating the map) and displays the reports on a map using weather event type symbols providing an overview of their spatial distribution and density. By selecting individual weather event types, an overview of the images and/or comments corresponding to the reports can be explored in the list panel or by hovering over the map symbols. With this approach, the user can access all available textual and image data; however, as a summarizing overview of the content is missing, data exploration is cumbersome.

Dashboard 1 of the Superset prototype (Fig. 5) provides functionality for filtering reports by time and event types through the filter box. However, there is no support for spatial filtering that is propagated to the data in all views. An overview of the spatial distribution and density of the reports corresponding to a selected weather event type is displayed on the map view (Fig. 5a). The Sankey diagram (Fig. 5d) makes it possible to explore the qualitative attributes related to the selected weather event type. Additional functionality for inspecting the corresponding textual contributions is provided by the word cloud (Fig. 5c), where words describe weather event types (e.g. rain, thunderstorm) and associated terms (e.g. cautiously, flooding), and the words’ font size encodes their frequency. Word colours are set randomly. Dashboard 1 provides an aggregated overview of the textual content of reports and the relations of the weather-type-related attributes. However, as the views are linked only through the filter box, there are no visible relations with their locations.

Performing this exploration task with CS-VAI (Fig. 6) allows for a considerably larger exploration space and flexibility of investigation compared to the earlier approaches. Data filtering in CS-VAI is applied interactively, i.e. spatial filtering through navigating the map, temporal filtering through selection on the temporal overview of the scatter plot, and filtering by weather event type or climate impact through clicking on the elements of the scatter plot (Fig. 6b). Moreover, filtering is applied in an informed manner since visual cues are provided reflecting the distribution of the data in terms of space, time, and attributes. Furthermore, all views are linked directly, so filtering operations applied in one view are reflected in all others, and thus, a user can

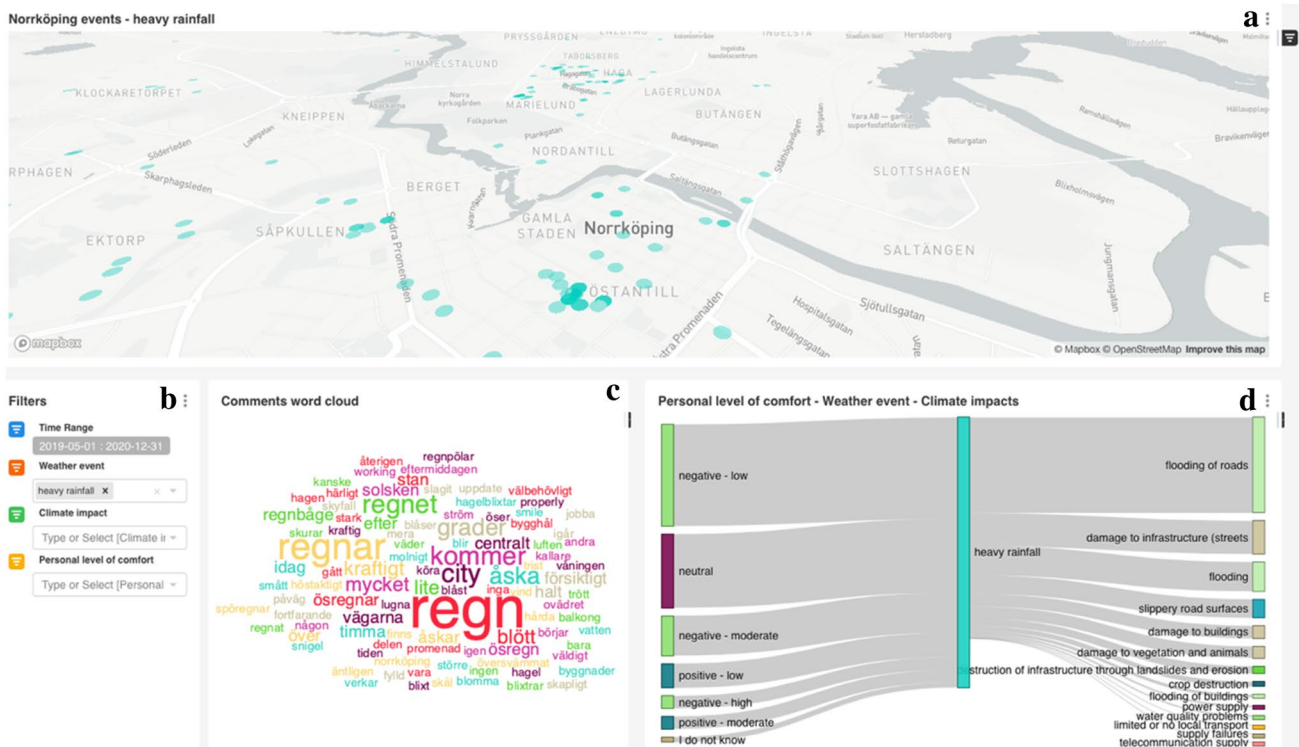


Fig. 5 Superset prototype Dashboard 1 displaying a selection of heavy rainfall events. The dashboard is composed of four views: **a** a map view displaying the position of reports, **b** a filtering panel, **c** a

word cloud displaying the most common terms in the explored report subset, and **d** a Sankey diagram displaying linkages between report attributes

flexibly navigate the search space of the data. Exploration of the images and textual content of the submitted reports is possible in all views. Hovering over the points in the scatter plot and the map reveals corresponding images and textual comments. Further contextual analysis can be performed with the image clouds view. A user can inspect the corresponding images in different arrangements (weather event type, level of comfort, colour content) to get a better understanding of their content and context. Hovering over the images displays their corresponding textual comments.

All three interfaces, i.e. Web Portal, Dashboard 1 of the Superset prototype, and CS-VAI, allow the user to zoom in on areas of spatial and/or temporal interest, for instance, on areas with a high number of reports for a certain weather event type or impact, to obtain a spatial demarcation of areas that should be further analysed. This is achieved with a varying level of flexibility and visual cue support in the different interfaces as previously described. Review of the images and comments submitted together with the reports for such a selection can be indicative of whether a high number of reports are due to a certain climatic event, or solely because a large number of users reported from the same area. Evidence of this can potentially be seen in the image clouds view of CS-VAI.

where + what → when: Selection of specific data categories and their comparison to background weather data (sensors)

The second use case is concerned with exploring the times (*when*) when certain objects (*what*) appear in certain locations (*where*). Of particular interest in this use case is how the submitted reports relate to the weather parameters collected from sensors. This exploration question can give deeper insights to the temporal characteristics of the collected data and environmental conditions. Even though all the proposed tools provide some level of support for exploring temporal aspects, as outlined below, this task, in its strict definition, can only be fully supported by the CS-VAI tool.

The Web Portal allows only a preselection and overview of the reports to be performed, as described in the first case. Dashboard 2 of the Superset prototype (Fig. 7) and CS-VAI (Fig. 6) allow for simultaneous exploration of both the reports and the weather parameters obtained from sensor networks. To achieve this, both tools employ a multiline graph (Fig. 7d and Fig. 6c, respectively) that provides an overview of the weather parameters during the campaign periods. The graph can be used to identify specific time ranges of interest, for which the spatial distribution of reports can be explored in the map view. In Dashboard

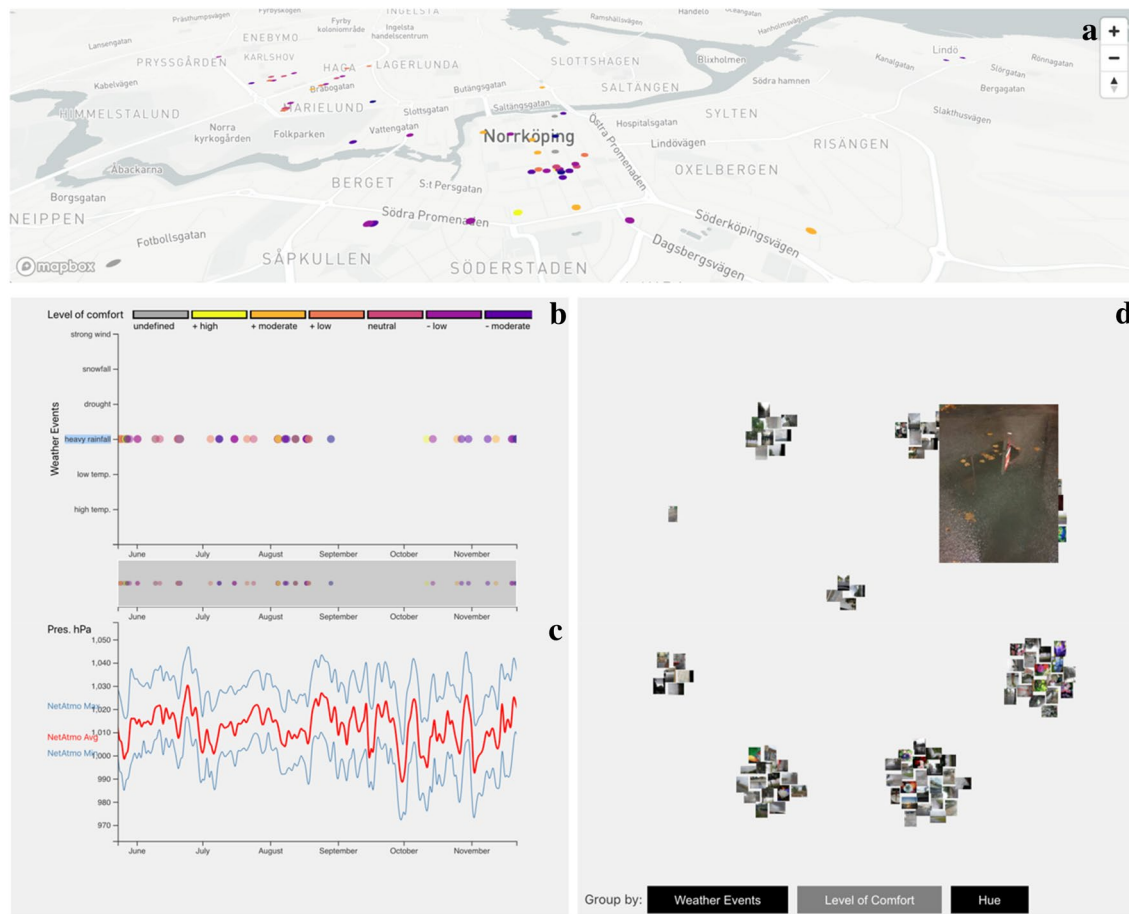


Fig. 6 CS-VAI displaying weather events of heavy rainfall event type. The positions of the events are displayed on the map (a), the scatter plot (b) reveals the temporal distribution of the heavy rainfall events coloured by level of comfort, in the line graph (c), air pressure values

are displayed over the explored time period, and in the image clouds view (d), the corresponding images are displayed, arranged by level of comfort

2, this selection of time ranges is performed through the filter box (Fig. 7c), while in CS-VAI it is done through the focus + context view (Fig. 3b). Additionally, Dashboard 2 allows the exploration of the average temperatures for sensor measurements in its prism map (Fig. 7b) and includes also a parallel coordinates plot (Fig. 7e), so users can examine the correlation between three weather parameters of the sensor data, i.e. air temperature, pressure, and humidity. This is of particular importance to the investigation of heat stress that is influenced by high air temperature along with e.g. high humidity.

Dashboard 2 and CS-VAI partly support the task of this use case by providing a number of opportunities to select, filter, and view information from the different data sources based on the identification of interesting patterns (e.g. peaks) of the weather parameters displayed in the multiline graphs. So, interesting times (*when*) can be identified for which the ‘*what*’ and ‘*where*’ are explored. In addition, CS-VAI provides further functionality that allows users to interactively

select report subsets (objects) with respect to their weather event type, level of comfort, and/or their spatial location and observe the temporal position and distribution of these in the scatter plot (Fig. 6b). Hence, reports (*what*) that appear in certain locations (*where*) can be selected and their temporal characteristics (*when*) explored. Moreover, the placement of the scatter plot above the multiline graph allows users to observe relations between the reports and the temperature variations over the time period (Fig. 6c). This makes it possible to directly juxtapose data from the two sources.

when + what → where: Analysis of spatial patterns of specific impacts and assessment of data relationships

The third use case is concerned with the exploration of spatial areas (*where*) based on the qualitative (*what*) and temporal (*when*) attributes of the reports located within them. This exploration is interesting in terms of the patterns that can be

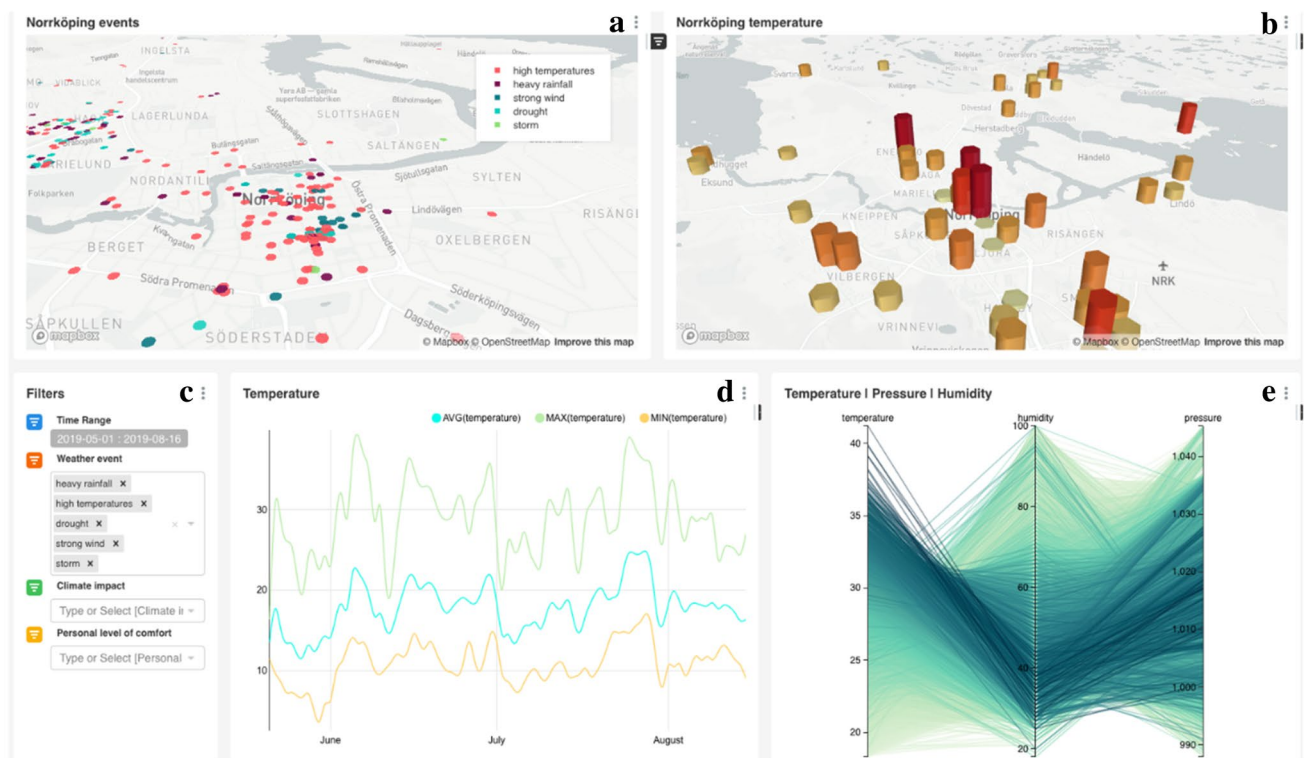


Fig. 7 Superset prototype Dashboard 2 displaying weather parameters from sensor data. The dashboard is composed of: **a** a map view displaying the position of the event reports, **b** a prism map displaying

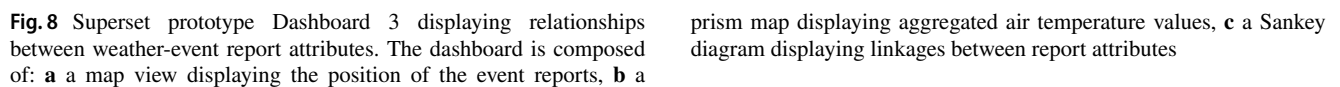
aggregated air temperature values, **c** a filtering panel, **d** a line graph displaying temperature values over time, and **e** a parallel coordinates plot displaying weather parameter values

revealed. The selection of days with, e.g. high temperatures, allows the examination of the spatial extent of sensor measurements and the identification of possible outliers. In this case, specific areas in the city of Norrköping, Sweden, can be identified for the period of a heat wave and then examined in terms of climate impact types that were reported for such a spatio-temporal selection.

The Web Portal's map view enables the visual analysis of the spatial distribution of weather event types given a filtered selection of the reports, but the potential to identify more complex relationships between data records is limited. Therefore, Dashboard 3 of the Superset prototype (Fig. 8) has been configured for this type of analysis. A Sankey diagram is included in the configuration (Fig. 8c) to aid the analysis of linkages between the primary qualitative attributes of the reports: (1) personal level of comfort, ranked from low to high during the selected period; (2) weather event types (high temperatures, heavy rainfall, drought, strong wind, and storm); and (3) corresponding climate impacts. The diagram provides an overview as to whether specific weather event types were frequently linked to low or high levels of comfort and to particular climate impact types that were reported for specific weather event types. The combination of two maps and a Sankey diagram

in Dashboard 3 provides insight into patterns that could be further investigated.

In CS-VAI (Fig. 3), the exploration of relationships between the reports' attributes and the potential spatial patterns they exhibit can be performed in the different views in a more integrated manner compared to the other tools. The scatter plot (Fig. 3b) displays weather event types (y-axis) over time (x-axis) and relates these to the reported level of comfort or the climate impact (colour). Users can perform temporal or attribute selections in the plot and explore the spatial distribution of the selected subsets in the map. The image clouds view (Fig. 3d) reveals groupings and similarities of the corresponding submitted images with respect to different attributes (weather event type, level of comfort, colour content of the image), so users can more deeply explore such relationships. A selection in one view highlights corresponding record representations in the remaining views, and so the spatial distribution of observed image groups can be inspected, and spatial patterns potentially revealed. With the functionality of CS-VAI, users can systematically explore different aspects of the data and flexibly examine relationships and patterns within and between them.



Visualization and interaction	Presence in tools				
	CitizenSensing web portal	Superset dashboards			CS-VAI
		1	2	3	
List view	Fig. 2a				
Map view	Fig. 2b	Fig. 5a	Fig. 7a	Fig. 8a	Figs. 3a and 6a
Word cloud		Fig. 5c			
Sankey diagram		Fig. 5d		Fig. 8c	
Prism map			Fig. 7b	Fig. 8b	
Parallel coordinate plot			Fig. 7e		
Multiline graph			Fig. 7d		Figs. 3c and 6c
Image clouds view					Figs. 3d and 6d
Scatter plot					Figs. 3b and 6b
Data attribute filtering (what?)	Available	Available	Available	–	Available
Spatial filtering (where?)	Available				Available
Temporal filtering (when?)	Available (discrete time range)	Available (discrete time range)	Available (discrete time range)		Available (continuous)

Discussion

The exploration of VGI facilitated by interactive visual interfaces involves several aspects that need to be considered.

First, our study indicates that the use of visual exploration shows potential for the multifaceted analysis of VGI since it enables the analysis of such data with respect to both their content and spatio-temporal characteristics. In this work, the spatio-temporal potential was explored through the three visual exploration and analysis approaches of successively increasing sophistication designed and developed during the course of the CitizenSensing project.

Second, although the proposed interfaces differ in terms of layout and functionality, they were all tested with the same climate-related VGI data content. The three central elements of the Triad framework (Peuquet 1994): time, space, and attributes, have been used as a theoretical basis for this testing and have been reflected in various layout configurations implemented in the interfaces. Table 1 summarizes the visualization techniques employed in the three approaches along with the information on the interactive functions the visual interfaces offer. The use of the Triad theoretical framework has been beneficial for two reasons. First, it has provided structure to the development process by driving the successive refinement of the approaches with respect to how effectively they answer questions relating to these elements. Second, it has ensured that the proposed approaches are generalizable to a certain level, since even though the approaches have been developed around a specific set of climate-related VGI, the exploration has been built around the generic core characteristics of the data.

Third, there are limiting characteristics and biases that are inherent to the nature of VGI and the way it is collected in general, which are independent of the visual interfaces used to explore them. These aspects include unequal spatial and temporal distribution of VGI, lack of quality verification, degree of participation, and motivations for participation and have been the focus of numerous articles relating to citizen science (e.g. Goodchild 2007; Goodchild and Li 2012; Haklay 2010, 2013; Meier et al. 2015). Such issues are naturally also present in the datasets used in this study. Spatio-temporal visual exploration as a general approach and consequently also the visual interfaces proposed here can prove valuable for exploring this aspect of VGI.

The optimal coverage of climate-related data would be a high number of reports distributed around the clock across a city from diverse end-users. The type, number, and spatio-temporal distribution of the data in this study was, however, dependent on the set-up and context of the campaigns during which the data were collected, similar to the study by Hung et al. (2016). Students uploaded a high number of reports near their school, and usually during school hours.

The homecare staff showed a higher degree of spatial and temporal variability in their reporting as one of the teams with one of the highest number of reports had work duties spread across the municipality. Such characteristics of the reports can be explored to various levels of detail in the proposed approaches by making temporal selections that match the campaign periods and inspecting their distribution and attributes in the views.

Motivation for volunteer participation is an issue closely linked to the amount, type, and quality of collected data (Baruch et al. 2016). In both campaigns, participants collected data in teams, a conscious choice to protect individual privacy which, however, may have decreased individual motivation. Moreover, a few teams uploaded most reports which illustrates the uneven motivation of the teams. Participation often decreases over time, as did happen in the homecare staff campaign, which can be seen in Fig. 3b as a reduced density of reports over time in the scatter plot of CS-VAI. With this in mind, extra visits were made to the students in attempt to keep an even level of motivation throughout the entire campaign. To encourage motivation, gamification was used in both campaigns, whereby points were awarded for each report. This proved to both discourage and encourage participation. Some participants reported diminished motivation when their teammates were inactive, while others explained that their team inspired them to participate. Of the teams with one of the highest number of reports, competition was pointed out as the highest motivational factor. Gamification also influenced the spatial variability of the data, as participants made multiple reports at the same location to gain additional points. These patterns could be revealed by identifying a high number of reports during a single day in the time range overview with the CS-VAI.

Another aspect that influences the content and quality of the collected data relates to the specification level of the instructions given to contributing participants as well as the participants' preconceptions of the task. For example, the data distribution for specific weather events might be influenced by pre-understanding of the included categories, which might impact the number, distribution, and type of responses for different user groups. Similarly, the impacts that users can select, might not correspond to their perceived observations or experiences, which may lead to an over-representation of broader categories (e.g. personal level of comfort or seasonal observations). Exploration of these aspects of the data can be effectively performed by studying the image content of the reports in relation to corresponding data attributes (e.g. weather event type) as enabled through the image clouds view of CS-VAI. Homecare staff were given a general introduction about reporting and a majority of uploaded images lacked comments on their perception of the event or adaptation recommendations. A more detailed introduction to secondary students in a subsequent campaign

where they were encouraged and reminded to upload such information resulted in more diverse and comprehensive reports.

The climate-related VGI was collected as part of a project to assess whether and how a participatory risk management system might be useful to urban citizens and relevant organizations. If a municipality or other relevant organization would operate a similar system or part of it, considerable effort would be needed to ensure adequate spatial and temporal distribution and quality assurance. End-user fatigue, a common barrier in participatory studies, and gamification both challenge data quality. To ensure an even spatial and temporal distribution of relevant data, organizations could engage municipal employees or participants of local organizations where collaboration could be sustained over time, a process that would demand significant amounts of human resources. To gain information on extreme events, night-time occurrences, or unrepresented spatial areas, targeted campaigns could be arranged with specific individuals and groups.

Conclusions

The aim of the presented work has been to explore the potential of visual exploration in the context of VGI related to climate change impacts and adaptation. This has been performed through the progressive design and development of three approaches of successively increasing analytic functionality for visually exploring and analysing climate-related VGI collected within the CitizenSensing project.

The work has pursued an iterative development process where the limitations of one approach have become the trigger for the subsequent ones. The CitizenSensing Web Portal was designed for providing an overview of the collected VGI to a wider audience. For this reason, only elementary exploration functionality is supported through simple geographic visualization and interaction techniques. In turn, the Superset prototype with its three dashboards was the initial attempt for achieving more in-depth exploration, where each of the dashboards developed was tailored towards one of the Triad framework elements. While the three dashboards allowed greater insight into the data, the exploration was inhibited by the fragmented interaction possibilities and the limited opportunities for cross-comparisons across many attributes. The importance of being able to freely juxtapose different attributes of the data and easily follow hypotheses that emerge during the visual exploration process became obvious. This brought about the design of the final interface proposed, CS-VAI, whose main goal was to allow the user to seamlessly explore the temporal, spatial, and attribute space of the data at hand in an integrated manner.

The proposed visual interfaces make it possible to explore the collected climate-related VGI with respect to their spatial, temporal, and qualitative attributes (such as weather event type and level of comfort) and gain insight concerning the content and quality of the data as well as the data collection process. However, there are still several interesting future directions to this work. One is to perform a systematic user-based evaluation of CS-VAI with stakeholders involved in the CitizenSensing project where we measure task performance and also collect eye-tracking data and subsequently to further develop both the data collection and visual exploration approaches based on the collected feedback. Another direction we aim to pursue is to integrate AI-based image processing and text mining into CS-VAI to enable the identification of impacts of extreme weather events.

Funding Open access funding provided by Linköping University. This work was part of the Citizen Sensing Project, which is part of ERA4CS, an ERA-NET initiated by JPI Climate, and funded by FORMAS (SE), RCN (Norway), NWO (The Netherlands), and FCT (Portugal) with co-funding by the European Union (Grant 690462).

Availability of data and material (data transparency) The collected data are not publicly available at the moment of submission.

Code availability (software application or custom code) The code is not publicly available at the moment of submission.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amar R, Eagan J, Stasko J (2005) Low-level components of analytic activity in information visualization. In: IEEE Symposium on Information Visualization. Minneapolis, Minnesota, USA, pp 111–117. <https://doi.org/10.1109/INFVIS.2005.1532136>
- Baruch A, May A, Yu D (2016) The motivations, enablers and barriers for voluntary participation in an online crowdsourcing platform. *Comput Hum Behav* 64:923–931. <https://doi.org/10.1016/j.chb.2016.07.039>

- Bostock M, Ogievetsky V, Heer J (2011) D3: Data-Driven Documents. *IEEE Trans vis Comput Graph* 17(12):2301–2309. <https://doi.org/10.1109/TVCG.2011.185>
- Cockburn A, Karlson A, Bederson BB (2009) A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput Surv (CSUR)* 41(1):1–31. <https://doi.org/10.1145/1456650.1456652>
- Dörk M, Carpendale S, Collins C, Williamson C (2008) VisGets: coordinated visualizations for web-based information exploration and discovery. *IEEE Trans vis Comput Graph* 14(6):1205–1212. <https://doi.org/10.1109/TVCG.2008.175>
- Dykes J, Purves R, Edwardes A, Wood J (2008) Exploring volunteered geographic information to describe place: visualization of the ‘Geograph British Isles’ collection. In: *Proceedings of the GIS research UK 16th annual conference GISRUK*, pp 256–267
- Goodchild MF (2007) Citizens as sensors: the world of volunteered geography. *GeoJournal* 69:211–221. <https://doi.org/10.1007/s10708-007-9111-y>
- Goodchild MF, Li L (2012) Assuring the quality of volunteered geographic information. *Spat Stat* 1:110–120. <https://doi.org/10.1016/j.spasta.2012.03.002>
- Haklay M (2010) How good is volunteered geographical information? a comparative study of OpenStreetMap and ordnance survey datasets. *Environ Plann B Plan Des* 37(4):682–703. <https://doi.org/10.1068/b35097>
- Haklay M (2013) Citizen science and volunteered geographic information: overview and typology of participation. *Crowdsourcing geographic knowledge*. Springer, Berlin, pp 105–122
- Häußler J, Stein M, Seebacher D, Janetzko H, Schreck T, Keim DA (2018) Visual analysis of urban traffic data based on high-resolution and high-dimensional environmental sensor data. In: *Envir-Vis 2018: Workshop on Visualisation in Environmental Sciences*. <https://doi.org/10.2312/envirvis.20181138>
- Helbig C, Dransch D, Böttinger M, Devey C, Haas A, Hlawitschka M, Kuenzer C, Rink K, Schäfer-Neth C, Scheuermann G, Kwasnitschka T, Unger A (2017) Challenges and strategies for the visual exploration of complex environmental data. *Int J Digit Earth* 10(10):1070–1076. <https://doi.org/10.1080/17538947.2017.1327618>
- Hung KC, Kalantari M, Rajabifard A (2016) Methods for assessing the credibility of volunteered geographic information in flood response: a case study in Brisbane, Australia. *Appl Geogr* 68:37–47. <https://doi.org/10.1016/j.apgeog.2016.01.005>
- Jänicke S (2019) Visual exploration of the European red list. *Worksh vis Environ Sci (EnvirVis)*. <https://doi.org/10.2312/envirvis.20191103>
- Jänicke S, Scheuermann G (2014) Utilizing GeoTemCo for visualizing environmental data. *Worksh vis Environ Sci (EnvirVis)*. <https://doi.org/10.2312/envirvis.20141107>
- Keim DA (2001) Visual exploration of large data sets. *Commun ACM* 44(8):38–44. <https://doi.org/10.1145/381641.381656>
- Li J, Zhang K, Meng Z-P (2014) Vismate: interactive visual analysis of station-based observation data on climate changes. In: *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp 133–142. <https://doi.org/10.1109/VAST.2014.7042489>
- Lukasczyk J, Liang X, Luo W, Ragan ED, Middel A, Bliss N, White D, Hagen H, Maciejewski R (2015) A collaborative web-based environmental data visualization and analysis framework. In: *EnvirVis@ EuroVis*, pp 25–29. <https://doi.org/10.2312/envirvis.20151087>
- MacEachren AM (1994) Visualization in modern cartography: Setting the agenda. In: MacEachren AM, Taylor D (eds) *Visualization in modern cartography*. Pergamon, Oxford, pp 1–12
- MacEachren AM, Jaiswal A, Robinson AC, Pezanowski S, Savelyev A, Mitra P, Zhang X, Blanford J (2011) Senseplace2: geotwitter analytics support for situational awareness. In: *2011 IEEE conference on visual analytics science and technology (VAST)*, IEEE, pp 181–190. <https://doi.org/10.1109/VAST.2011.6102456>
- Meier F, Fenner D, Grassmann T, Jänicke B, Otto M, Scherer D (2015) Challenges and benefits from crowd sourced atmospheric data for urban climate research using Berlin, Germany, as testbed. In: *ICUC9-9th International Conference on Urban Climate Jointly with 12th Symposium on the Urban Environment*
- Munzner T (2009) A nested model for visualization design and validation. *IEEE Trans vis Comput Graph* 15(6):921–928. <https://doi.org/10.1109/TVCG.2009.111>
- Neset TS, Glaas E, Ballantyne AG, Linnér BO, Opach T, Navarra C, Johansson J, Bohman A, Rød JK, Goodsite M (2016) Climate change effects at your doorstep: geographic visualization to support Nordic homeowners in adapting to climate change. *Appl Geogr* 74:65–72. <https://doi.org/10.1016/j.apgeog.2016.07.003>
- Neset TS, Wilk J, Navarra C, Capell R, Bartosova A (2019) Visualization-supported dialogues in the Baltic Sea Region. *Ambio* 48(11):1314–1324. <https://doi.org/10.1007/s13280-019-01250-6>
- Peuquet DJ (1994) It’s about time: a conceptual framework for the representation of temporal dynamics in geographic information systems. *Ann Assoc Am Geogr* 84(3):441–461. <https://doi.org/10.1111/j.1467-8306.1994.tb01869.x>
- Roberts JC (2007) State of the art: coordinated and multiple views in exploratory visualization. In: *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, pp 61–71. <https://doi.org/10.1109/CMV.2007.20>
- Seebacher D, Miller M, Polk T, Fuchs J, Keim DA (2019) Visual analytics of volunteered geographic information: detection and investigation of urban heat islands. *IEEE Comput Graph Appl* 39(5):83–95. <https://doi.org/10.1109/MCG.2019.2926242>
- Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualizations. *Proc IEEE Symp vis Lang*. <https://doi.org/10.1109/VL.1996.545307>
- Tenney M, Hall GB, Sieber RE (2019) A crowd sensing system identifying geotopics and community interests from user-generated content. *Int J Geogr Inf Sci* 33(8):1497–1519. <https://doi.org/10.1080/13658816.2019.1591413>
- Van Wijk JJ (2005) The value of visualization. *VIS 05. IEEE vis* 2005:79–86. <https://doi.org/10.1109/VISUAL.2005.1532781>
- Wang Y (2017) Deck.gl: large-scale web-based visual analytics made easy. In: *IEEE Workshop on Visualization in Practice*, 1–4. Retrieved from <http://arxiv.org/abs/1910.08865>
- Wang X-M, Zhang T-Y, Ma Y-X, Xia J, Chen W (2016) A survey of visual analytic pipelines. *J Comput Sci Technol* 31(4):787–804. <https://doi.org/10.1007/s11390-016-1663-1>
- Yi JS, ah Kang Y, Stasko J (2007) Toward a deeper understanding of the role of interaction in information visualization. *IEEE Trans vis Comput Graph* 13(6):1224–1231. <https://doi.org/10.1109/TVCG.2007.70515>