

# Visual Exploration of Climate-Related Volunteered Geographic Information

C. Navarra<sup>1</sup>, T. Opach<sup>2</sup>, K. Vrotsou<sup>1</sup>, A. Joling<sup>3</sup>, J. Wilk<sup>1</sup>, T.-S. Neset<sup>1</sup>,

<sup>1</sup> Linköping University, Linköping, Sweden

<sup>2</sup> Norwegian University of Science and Technology, Trondheim, Norway

<sup>3</sup> Deltares, Delft, The Netherlands

---

## Abstract

*This paper describes two approaches for visually exploring climate-related data collected within the citizen science research project, CitizenSensing. The project addresses the need of European cities and their citizens for enhanced knowledge of site-specific conditions regarding climatic risks and adaptation measures. The visual exploration approaches discussed are: (1) a web portal enabling users to gain a low-level overview of the collected data on a map, and (2) a visual analysis tool facilitating in-depth visual data exploration in search of spatio-temporal patterns. The aim of the study is to assess and discuss the potential of such visual exploration approaches in the context of Volunteered Geographic Information.*

## CCS Concepts

• **Human-centered computing** → **Geographic visualization; Information visualization;** • **Applied computing** → **Environmental sciences;**

---

## 1. Introduction

Visual data exploration can be defined as the process of using visualization to gain insight into large sets of data. 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 [Kei01]. Visual data exploration is interactive, intuitive and allows the domain knowledge of a user to be seamlessly incorporated in the analysis. When the data to be visually explored have a geographic reference, these are commonly represented on an interactive map enabling users to search for geographical patterns and spatial relationships and thus to gain additional knowledge about investigated phenomena.

Having this as a starting point and following Shneiderman's information seeking mantra of "overview, zoom and filter, and details-on-demand" [Shn96], this study describes two approaches to visually exploring and analysing data generated by citizens within the CitizenSensing project. The two approaches are respectively concerned with: (1) low-level data overview in a map display context, (2) in-depth visual investigation in search for spatio-temporal patterns. The primary aim of the study is to assess and discuss the potential of such visual exploration approaches in the context of Volunteered Geographic Information (VGI). The data has been collected in one of the CitizenSensing project's campaigns held in the city of Norrköping, Sweden.

The CitizenSensing project develops and evaluates a participatory risk management system that allows urban citizens to voluntarily report, by means of a web application, information about site-specific extreme weather conditions, their impacts, and recommendations for best-practice adaptation measures. The project aims to assess if and how the system increases citizen engagement and their contribution to urban resilience. The targeted weather phenomena, i.e., events with high and low temperature or heavy rainfall, are decided upon with relevant urban stakeholders according to the context of each pilot. The users of the application can view submitted reports and recommendations of best-practice adaptation measures related to specific weather events, based on information from municipal and public authorities. The CitizenSensing web application aims to facilitate a high level of interactivity by engaging citizens, both as providers of locally situated data on weather impacts and adaptation measures, as well as receivers of specific information on local conditions and recommendations on how to respond.

## 2. Related work

Our study is concerned with the visual exploration of spatio-temporal user-generated climate-related data that includes reports on weather conditions, impacts and/or recommendations of appropriate actions. Therefore, it relates to research in the following three areas: (1) the use of visualization for knowledge discovery, (2) vi-

sualization of environmental data, and (3) visualization of volunteered geographic information. These are briefly depicted below.

The importance of interactive visualization and the human-in-the-loop in the analysis of data is underlined by the large body of research on visual exploration and analytics interfaces [WZM\*16]. Several visualization pipelines have been proposed over the past years focusing on different analysis aspects and allowing various levels of user involvement (e.g. [VW05], [Mun09]). A multitude of visual exploration interfaces have been developed that facilitate flexible interactive analyses and knowledge discovery. The common cornerstones of visualization interfaces of any type are:

- the mapping of data values to visual elements,
- the use of coordinated multiple views (CMV) to facilitate multifaceted data exploration [Rob07], [DCCW08],
- the use of interaction techniques that provide flexibility of analysis to the user [YaKS07].

In recent years, the visualization of environmental data is gaining popularity and a number of complex visual exploration interfaces have emerged promoting their analysis. For instance, Li et al. [LZM14] proposed an interactive visual analysis interface dedicated to the exploration of climate change. Jänicke et al. [JS14] illustrated the use of the open source library GeoTemCo with applications on environmental data including storms and natural hazards data. A geographic visualization approach was proposed in [NGB\*16] for addressing 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 proposed by Lukasczyk et al. [LLL\*15] and exemplified with a dataset of water scarcity in Africa. Neset et al. [NWN\*19] designed an interactive visualization tool to support dialogues with water management and ecosystem stakeholders. Finally, Jänicke [Jän19] presented a system based on a series of visual interfaces that enable the flexible exploration of the European Red List of endangered species.

With the rise of platforms encouraging the collection of user-contributed data the number of systems concerned with visualization of VGI [Goo07] has been increasing. A visual analytics system for the exploration of environmental data collected via the citizen science platform enviroCar was proposed by Häubler et al. [HSS\*18]. Seebacher et al. [SMP\*19] have proposed a system allowing the prediction and visual analysis of urban heat islands based on volunteered data from a private weather station network. Tenney et al. [THS19] describe a crowd sensing system that captures geospatial social media topics and visualizes the results in multiple views.

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

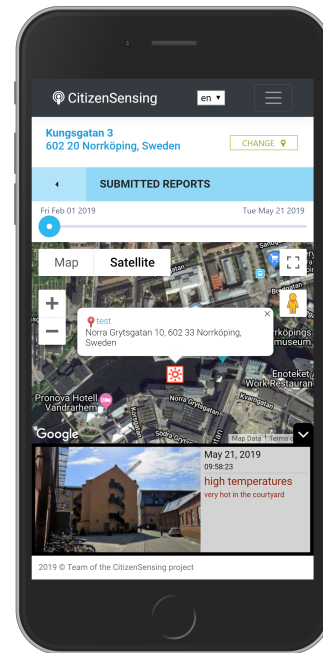


Figure 1: The CitizenSensing web application for data collection.

### 3. Data

The participatory risk management system developed in the CitizenSensing project is composed of a web application (figure 1), a web portal (figure 2), a sensor network, and a database. The system aims to collect and store place-specific information to provide an integrated platform to support local stakeholders, organizations and citizens in increasing their ability to make informed decisions related to climatic risks and increase urban resilience.

The web application enables citizens to submit reports on observed impacts of specific weather events and their impacts as well as seasonal observations, and to assess real-time data obtained from open access weather station networks. The system integrates data from the Netatmo weather station networks [Net] for specific climate parameters, depending on availability (e.g., temperature, air pressure, humidity, and precipitation). As such, citizens act as sensors by submitting place-based information, are enabled to explore the data submitted by other users, which is visualized on the web interface, but are also provided with recommendations for climate adaptation related to specific weather events.

The data, collected through the web application and stored on a database, comprises the following attributes: geographical coordinates, date and time, weather event type, 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. The data is recorded on the server in a MySQL relational database management system. When a report is submitted, the web application does not store any personal information except for the geographical position, provided that the user has enabled the app to retrieve this information.

As part of the CitizenSensing project, several campaigns with

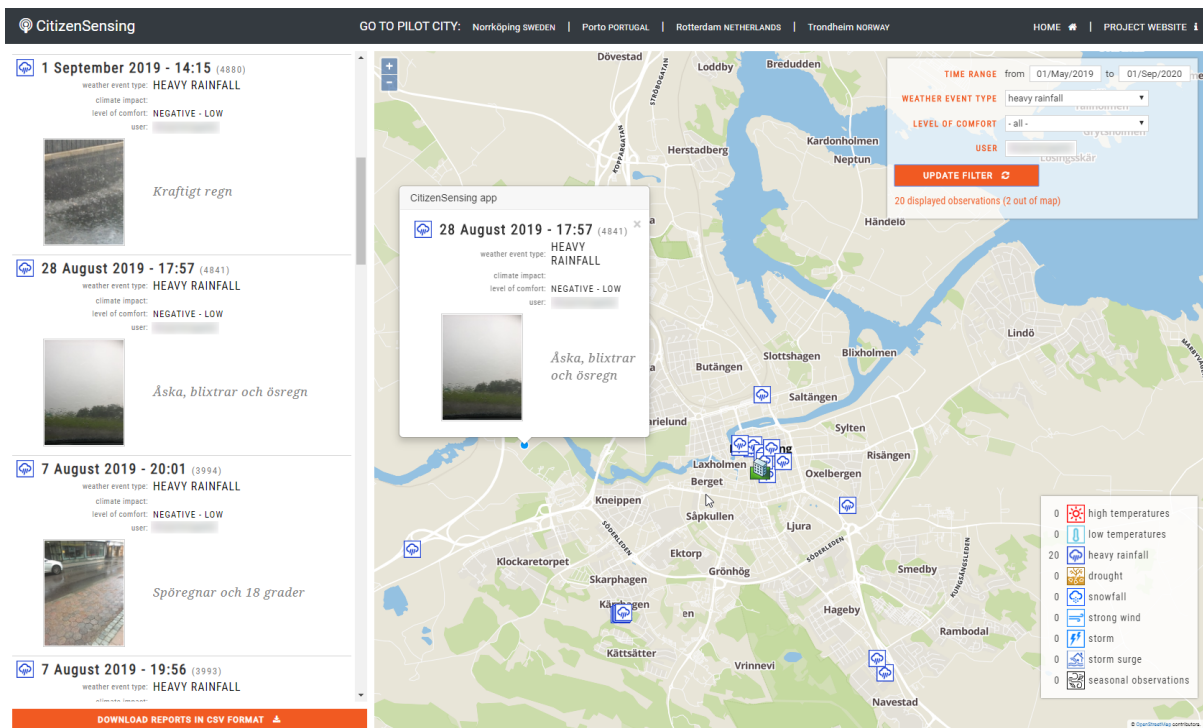


Figure 2: The CitizenSensing web portal.

different user groups were held during 2019 in the project's four pilot cities, Norrköping in Sweden, Porto in Portugal, Rotterdam in the Netherlands, and Trondheim in Norway. Data were collected through user groups instead of single individuals in order to preserve privacy. We assess the potential of the proposed visual analysis approaches by analysing data collected in one of the campaigns held in the city of Norrköping. The data for this campaign consist of 3917 reports submitted by 28 participants grouped into 3 teams. The data were submitted through the CitizenSensing web application during the campaign period May to September 2019.

#### 4. Visual analysis approaches

Two different approaches for visual analysis are employed to facilitate visual exploration of VGI: (1) a web portal, and (2) a visual analysis tool. While the web portal enables users to gain a low-level data overview in a map display context, the visual analysis tool facilitates in-depth visual data investigation to search for spatio-temporal patterns.

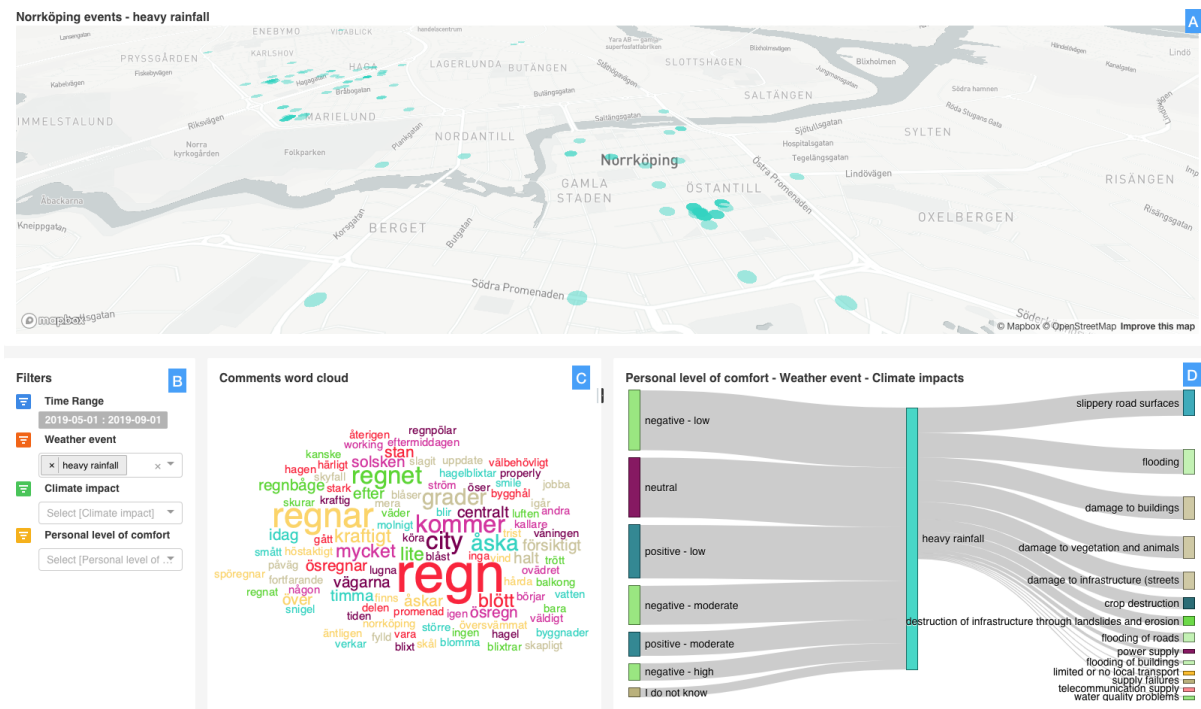
##### 4.1. Web Portal

The CitizenSensing web portal (figure 2) enables flexible exploration of the reported data collected through the different pilots and campaigns. The web portal has been specifically designed for desktop displays and features tailored user interaction to support the filtering and assessment of data. The web portal has been developed using open-source Javascript APIs such as OpenLayers for

web mapping, Bootstrap for responsive layout, and jQuery for user interaction.

In the web portal, submitted reports can be filtered by location, time range (and thus by campaign) or user-id. Furthermore, data can be filtered by a weather event type or by a level of comfort reported by users. Filtered reports are presented on a map display and visually differentiated by map symbols that reflect weather event types under which the reports were submitted. Apart from the map display, filtered reports are listed in a panel to the left-hand side of the map display. To keep consistency, in the web portal, we have used the same symbols as those used in the web application so that users can quickly recognize these, as they have already become familiar with them. If a report includes a picture or textual comment related to the weather event, this information is displayed in the list view. Reports can be selected through the list view or by clicking on an item on the map. The selection of filtering can be changed at any time during the process. The web portal also allows filtered reports to be exported in the Comma Separated Value format (CSV) for additional processing. The data to be exported are based on the filtering criteria set by the user as well as the map range (its geographical bounding box).

The web portal has two versions. The first version, which is publicly available, shows reports without personally identifiable information. Personal data such as comments and pictures are not visible in this version. In turn, the second version of the web portal is password protected and shows comments and photos used only as part of the research project in order to support analysis and evaluation of the reported data.



**Figure 3:** Visual analysis tool dashboard displaying a selection of heavy rainfall events. The two most prominent terms in the word cloud are rain (*regn*) and raining (*regnar*).

## 5. Visual analysis tool

The visual analysis tool was designed to explore climate-related VGI collected in the CitizenSensing project. The tool has been implemented using the Apache Superset software [Apa] that is an open source data visualization and exploration platform typically used for business intelligence. Apache Superset provides an intuitive interface for creating interactive visual dashboards combined with an SQL Interactive Development Environment (IDE). Apache Superset offers a number of visualization charts, based on D3.js [BOH11] or deck.gl [Wan17] that can be dynamically linked for creating interactive visual dashboards.

The proposed visual analysis tool encompasses three visual dashboards, where each is tailored to support specific tasks depending on the use case. The three dashboards make use of six different visualization charts to display various aspects of the data from two separate user-generated data sources. The following charts have been included in the tool:

1. Map view displaying the position of the reported weather events. Colour in this representation represents the type of event (figure 3(A)).
2. Map view with a hexagon heatmap overlay based on the temperature values taken from the Netatmo sensor network. The height and colour of the hexagonal prisms correspond to the average temperature of the bin (figure 4(B)).
3. Sankey diagram displaying the linkages between weather event type, impact type and personal level of comfort collected through the weather event reports. The size of the bands rep-

resents the number of linkages between the reported data (figure 3(D)).

4. Parallel coordinate plot displaying climate parameters, in this case, temperature, air pressure and humidity, from the Netatmo sensor network (figure 4(E)).
5. Line graph representing aggregated minimum, maximum, and average temperature values for a spatio-temporal selection. The data are retrieved from the Netatmo sensor network (figure 4(D)).
6. Word cloud displaying frequent terms submitted in the comments of the reported weather events. The size of the word is proportional to the number of reports (figure 3(C)).

The three dashboards combine these charts in different ways as described in the following section. All views in each dashboard are linked by time and weather events through using a “filter box”(figure 3(B)).

## 6. Use cases

We present three use cases for the exploration of the collected user-generated data, and discuss opportunities and challenges for the proposed analytical methods. Each case exemplifies one of the dashboards comprising the proposed visual analysis tool. The data used in these cases concern Norrköping, Sweden and correspond to periods with high temperatures, which surpassed the thresholds for heat wave warnings from the Swedish Meteorological and Hydrological Institute (SMHI). Moreover, the data include submitted reports that concern selected weather events such as heavy rainfall,



**Figure 4:** Visual analysis tool dashboard displaying a selection of data for a specific time period of the campaign with high temperature measurements.

high temperature or storms during the entire period to discuss the specific functionality of the presented visual analysis approaches.

### 6.1. Analysis of visual material and text contributions for selected events

The first case concerns the demand for an overview of submitted data in order to determine to what extent reports are linked to specific weather event types. In the CitizenSensing project, the collected information 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, e.g., spatial planning, climate communication efforts, and/or indicates a specific risk in an area.

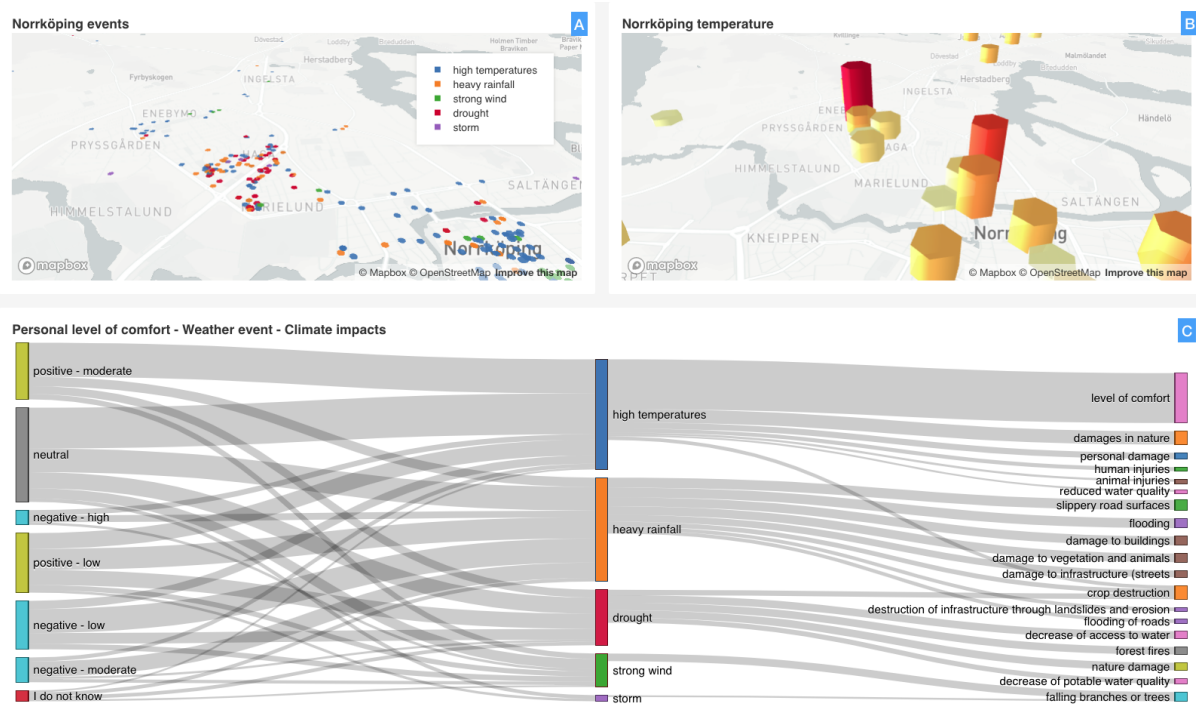
The CitizenSensing web portal (figure 2) fulfils the requirement of a quick comprehensive assessment and allows, in addition, the evaluation of smaller samples such as participating individuals or groups. The web portal facilitates an overview of the types of images and/or comments that are submitted during extreme events, as well as their location. An additional functionality, that enables analysis of the submitted textual contributions, is provided by the proposed visual analysis tool in the form of a word cloud (figure 3(C)). The size of the font shows the frequency of the terms while the words describe weather events (e.g. ‘rain’, ‘thunderstorm’) as well as associated terms (e.g. ‘cautiously’, ‘flooding’), colours are set randomly.

Both approaches allow the user to zoom in on an area that has for instance a high number of reports for a certain weather event or impact, in order to obtain a spatial demarcation of areas that should be further analysed. Reviewing the images submitted together with the reports might, for such a selection, be indicative of whether the high number of reports was due to a certain climatic event, or solely because a large number of users happened to be in the same area.

### 6.2. Selection of specific data categories and their comparison with background data (sensors)

While a pre-selection of data can be made in the web portal, as described in the first case (section 6.1), the visual analysis tool allows an investigation of both the reported data and the data measured by the sensor network. This can be done for a specified spatial or temporal selection. The tool employs a line graph (figure 4(D)) providing an overview of, for instance, peak temperatures during the campaign period. Based on this information, a user can select specific dates for which the spatial distribution of reports (figure 4(A)) and average temperatures for sensor measurements (figure 4(B)) are displayed in the map views. As such, the visual analysis tool provides a number of opportunities to select and view information based on the identification of peaks in the line graphs, spatial and temporal selection, as well as selection and combination of different datasets.

The high flexibility of the different visual representations allows for a tailored analysis, which can be further advanced by the inclusion of a parallel coordinate plot (figure 4(E)), to assess the corre-



**Figure 5:** Visual analysis tool dashboard displaying weather events and sensor data for a selected geographic area, and linkages between reported personal level of comfort, weather event type and impact.

lation between different parameters of the sensor data (air temperature, pressure and humidity). This is of particular importance for the investigation of heat stress that is influenced by high air temperature along with high humidity.

### 6.3. Analysis of spatial or temporal patterns of specific impacts, and assessment of data relationships

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 can be identified for the period of a heat wave, and then examined in terms of impact types that were reported for this spatio-temporal selection.

Figure 5 shows the use of a Sankey diagram to aid the analysis of such linkages. It shows the distribution of the linkages between the reported personal level of comfort, ranked from low to high during the selected period, the different weather events (high temperatures, heavy rainfall, drought, strong wind and storm) and related impacts (figure 5(C)). This type of analysis provides an overview as to whether specific weather events were more frequently linked to low or high levels of comfort and the types of impacts that were reported for specific event types, to provide insight into patterns that could be further assessed by means of qualitative research methods such as surveys or interviews with selected participants, or compared to other time periods, campaigns or pilot regions.

## 7. Conclusions

This paper describes two approaches for visually exploring and analysing climate-related data collected within the citizen science research project, CitizenSensing. The project is concerned with developing a participatory risk management system that allows urban citizens to voluntarily report information about site-specific extreme weather conditions, their impacts, and recommendations for best-practice adaptation measures. Information is volunteered by means of a web application and visually explored within the proposed approaches: (1) a web portal enabling users to gain a low-level overview of the collected data on a map, and (2) a visual analysis tool facilitating in-depth visual data exploration in search for spatio-temporal patterns. The aim of the study has been to assess and discuss the potential of such visual exploration approaches in the context of VGI. The primary focus has been on the proposed visual analysis tool.

The visual analysis tool comprises a set of six visualization charts, selected to display different aspects of the data, and arranged into three visual dashboards, each tailored to support specific tasks depending on the analysis scenario. As such, the tool provides a high level of flexibility for the exploration of spatio-temporal relations in user-generated climate-related data. In the presented work, the analysis scenarios have focused on both data selection and identification of e.g. extreme periods, but also on exploration of linkages between data.

A limitation of the work is concerned with the inherent bias that arises from the data type used in this work; i.e. volunteered ge-

ographic information. In particular for this study, the collection of the volunteered data has occurred through different campaigns across Europe resulting in uneven distributions and hence multiple challenges. For example, the data distribution for specific weather events might be influenced by pre-understanding of the offered categories, which might impact the number, distribution and type of responses for different user groups. Similarly, the impacts that can be selected by users, might not correspond to the perceived experience of the users, and hence might lead to an over-representation of broader categories (e.g. personal level of comfort), rather than providing more specific descriptions of impacts. Finally, bias can also be observed in the crowd-sourced sensor data, contributed through the Netatmo weather stations, considered in this work since the placement of the sensors can considerably affect the accuracy of the collected measurements [MFG\*15].

A future direction of the work is to evaluate the visual analysis tool with stakeholders of the different European cities involved in the CitizenSensing project and further develop both the data collection and visual exploration approaches based on the collected feedback.

## References

- [Apa] Apache Superset (incubating), <https://superset.incubator.apache.org/>. Accessed 5 March 2020. 4
- [BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D3: Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309. 4
- [DCCW08] DÖRK M., CARPENDALE S., COLLINS C., WILLIAMSON C.: Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1205–1212. 2
- [Goo07] GOODCHILD M. F.: Citizens as sensors: the world of volunteered geography. *GeoJournal* 69, 4 (2007), 211–221. 2
- [HSS\*18] HÄUSSLER J., STEIN M., SEEBACHER D., JANETZKO H., SCHRECK T., KEIM D. A.: Visual analysis of urban traffic data based on high-resolution and high-dimensional environmental sensor data. In *EnvirVis 2018: Workshop on Visualisation in Environmental Sciences* (2018). 2
- [Jän19] JÄNICKE S.: Visual Exploration of the European Red List. *Workshop on Visualisation in Environmental Sciences (EnvirVis) (2019) 1500* (2019). doi:10.2312/envirvis.20191103. 2
- [JS14] JÄNICKE S., SCHEUERMANN G.: Utilizing geotemco for visualizing environmental data. 2
- [Kei01] KEIM D. A.: Visual Exploration of Large Data Sets. *Communications of the ACM* 44, 8 (2001), 38–44. 1
- [LLL\*15] LUKASCZYK J., LIANG X., LUO W., RAGAN E. D., MIDDEL A., BLISS N., WHITE D., HAGEN H., MACIEJEWSKI R.: A collaborative web-based environmental data visualization and analysis framework. In *EnvirVis@ EuroVis* (2015), pp. 25–29. 2
- [LZM14] LI J., ZHANG K., MENG Z.-P.: Vismate: Interactive visual analysis of station-based observation data on climate changes. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)* (2014), IEEE, pp. 133–142. 2
- [MFG\*15] MEIER F., FENNER D., GRASSMANN T., JÄNICKE B., OTTO M., SCHERER D.: 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* (2015). 7
- [Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics* 15, 6 (2009), 921–928. 2
- [Net] Weather API Documentation, <https://dev.netatmo.com/apidocumentation/weather>. Accessed 5 March 2020. 2
- [NGB\*16] NESET T. S., GLAAS E., BALLANTYNE A. G., LINNÉR B. O., OPACH T., NAVARRA C., JOHANSSON J., BOHMAN A., RØD J. K., GOODSITE M.: Climate change effects at your doorstep: Geographic visualization to support Nordic homeowners in adapting to climate change. *Applied Geography* 74 (2016), 65–72. URL: <http://dx.doi.org/10.1016/j.apgeog.2016.07.003>, doi:10.1016/j.apgeog.2016.07.003. 2
- [NWN\*19] NESET T. S., WILK J., NAVARRA C., CAPELL R., BARTOSOVA A.: Visualization-supported dialogues in the Baltic Sea Region. *Ambio* 48, 11 (2019), 1314–1324. doi:10.1007/s13280-019-01250-6. 2
- [Rob07] ROBERTS J. C.: State of the art: Coordinated & multiple views in exploratory visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)* (2007), IEEE, pp. 61–71. 2
- [Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. *Proceedings 1996 IEEE Symposium on Visual Languages* (1996), 336–343. doi:10.1109/VL.1996.545307. 1
- [SMP\*19] SEEBACHER D., MILLER M., POLK T., FUCHS J., KEIM D. A.: Visual analytics of volunteered geographic information: Detection and investigation of urban heat islands. *IEEE computer graphics and applications* 39, 5 (2019), 83–95. 2
- [THS19] TENNEY M., HALL G. B., SIEBER R.: A crowd sensing system identifying geotopics and community interests from user-generated content. *International Journal of Geographical Information Science* 33, 8 (2019), 1497–1519. 2
- [VW05] VAN WIJK J. J.: The value of visualization. In *VIS 05. IEEE Visualization, 2005.* (2005), IEEE, pp. 79–86. 2
- [Wan17] WANG Y.: Deck.gl: Large-scale Web-based Visual Analytics Made Easy. In *IEEE Workshop on Visualization in Practice* (2017), pp. 1–4. URL: <http://arxiv.org/abs/1910.08865>, arXiv:1910.08865. 4
- [WZM\*16] WANG X.-M., ZHANG T.-Y., MA Y.-X., XIA J., CHEN W.: A survey of visual analytic pipelines. *Journal of Computer Science and Technology* 31, 4 (2016), 787–804. 2
- [YaKS07] YI J. S., AH KANG Y., STASKO J.: Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics* 13, 6 (2007), 1224–1231. 2