SIGRAD 2014
M. Obaid, D. Sjölie and M. Fjeld (Editors)

Multi-Scale Trend Visualization of Long-Term Temperature Data Sets

Andreas Kerren[1], Ilir Jusufi[2], and Jiayi Liu[1]

[1]Linnaeus University, Department of Computer Science, ISOVIS Group, Växjö, Sweden
[2]University of California, Department of Computer Science, Davis, CA, USA

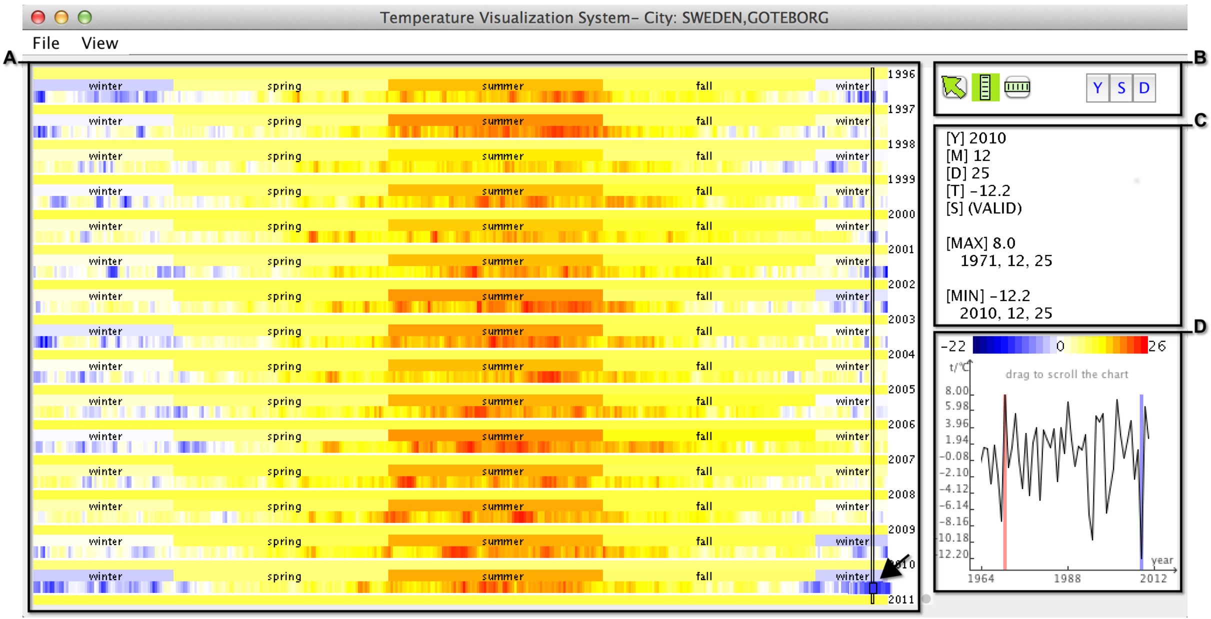

Figure 1: *An overview screenshot of our visualization tool. It shows temperature data for Gothenburg, Sweden, over a time span of 52 years in total. Region A indicates the* main view *of the tool. A* toolbox *is located in the top-right corner (B) and provides quick access to several selection options and view modifications. Information about selected data items is shown in textual format by the* details view *marked with C. The* selection view *(D) shows different selection ranges in the form of a standard timeline plot. A (vertical) range for December 25 over all available years was selected by the user (black arrow).*

Abstract

The analysis and presentation of climate observations is a traditional application of various visualization approaches. The available data sets are usually huge and were typically collected over a long period of time. In this paper, we focus on the visualization of a specific aspect of climate data: our visualization tool was primarily developed for providing an overview of temperature measurements for one location over decades or even centuries. In order to support an efficient overview and visual representation of the data, it is based on a region-oriented metaphor that includes various granularity levels and aggregation features.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical user interfaces (GUI) I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques J.2 [Computer Applications]: Physical Sciences and Engineering—Earth and atmospheric sciences

1. Introduction

Together with the daily discussions on the climate change in climate research and media in general, the need to analyze the permanently increasing amount of climate data is becoming overwhelming. Here, interactive visualizations—possibly in combination with traditional data mining methods—are the key to get meaningful results out of these typically huge time-dependent data sets.

Many advanced and powerful tools exist for the visualization of general time-series (cf. the survey [AMM*08], for instance), but most of them are not really used in climate research and unknown to many researchers in this domain. Our aim in this work was the development of an easy to use visualization tool that supports overviews of long-term climate data and facilitates interactive trend analyses. In order to solve such tasks, we had to decide on the visual representation and functionality of our tool: we concentrate on one user-defined location (i.e., city) and one scalar measurement (i.e., temperature), as well as the implementation of a region-oriented visualization metaphor that supports time visualization at multiple scale levels (granularity).

The rest of this paper is structured as follows. Section 2 discusses the most important related work and highlights main differences to our visualization tool. Next, the properties of the input data are given (Section 3). Section 4 presents our visualization approach, discusses interaction features, and provides some implementation aspects. Finally, we conclude with Section 5.

2. Related Work

A large variety of visualization systems have been developed for the interactive analysis of climate data over the past years. Many of them are map-based, i.e., geo-spatial climate data is represented as a map with a gradient color coding. Since map-based tools only present data at one time point, developers either use animation (e.g., trigged by a time slider) to display the whole history of usually one climate attribute, such as done in [VW14]; or they use glyphs for representing (multivariate) climate data over time, like the tool presented in [TSWS05]. Other more advanced visualization tools are composed of several coordinated views (e.g., line plots, maps, or scatterplots) and provide a rich set of interaction and visual analysis techniques, for instance SimVis [LSL*09] or WeatherSpark [Wea14]. Because of the page limit, we do not discuss general temperature and climate data visualizations and refer to the overview paper of Nocke et al. [NSBW08]. In the following paragraphs, we concentrate on our restricted focus area, i.e., on trend visualization of long-term temperature data sets as discussed in the introduction.

According to the previously mentioned overview paper, “other temporal visualization techniques such as pixel-oriented visualizations are rarely used [in climate research]” [NSBW08]. This still appears to be the case also nowadays, but there are some approaches that are clearly related to our design considerations and implementation. As we want to provide an efficient overview of long-term temperature trends, pixel-oriented techniques are particularly well-suited to display this kind of data. Recursive patterns [KKA95] are an early realization of a pixel-based approach for large multivariate data sets. They also support a way to show the hierarchical structure of time-dependent data, i.e., colored pixels may be arranged as groups that represent data for days, for instance, and these groups can be again arranged for other levels of granularity like weeks. However, recursive patterns do not inherently support advanced interaction possibilities. The pixel-based GROOVE tool [LAB*09] supports more interaction and advanced features like the combination of aggregated values with detailed information by using several types of overlays. In contrast to our approach, it is not possible to visualize a trend explicitly by a timeline, e.g., to show the temperature trend of the winter seasons over several years.

A related technique are so-called tile maps [MFSW97] which were designed to represent air quality trends. The attribute values (such as ozone measurements) are arranged on a grid following a calendar division. This approach provides a good overview and also the possibility to compare values of a specific week day (comparison of row elements of the grid) or weeks (comparison of grid columns). But, it is neither interactive, nor does it provide any aggregation features. Shimabukuro et al. [SFdOL04] propose a visualization technique for the analysis of trend patterns at several granular levels. This approach is pretty similar to ours, but they arrange the different levels (days, months, years) into three big separated blocks or regions, i.e., there are two cells for a specific month—one shows the details for each day of this month (by using a pixel-based approach), and the other one shows aggregated data (by using a colored tile). In contrast, our approach hierarchically arranges data of different granular levels side by side similar to an icicle plot.

3. Input Data

We obtain our data sets from the European Climate Assessment & Dataset project (`http://eca.knmi.nl`). Their website offers free data sets collected from different meteorological stations throughout Europe and the Mediterranean. The data can be downloaded via interactive queries as comma separated values (.csv) text documents. Available query parameters are location (country, city), period, elevation, series type (blended or non-blended data of nearby stations), and attribute (average/maximal/minimal temperature, amount of precipitation/sunshine, etc.).

Here, we focus on daily average temperature data. The data comprises temperature measurements that may span more than a hundred years back. Possible faulty readings and errors are marked in the downloaded data files. We tackle this problem of missing/faulty data by using green color in our visualization tool to denote such values.

Figure 2: *Temperature visualization of a single year (2013 in this case). The daily temperatures are drawn from left to right as small color blocks (or nodes) at the bottom of the visual representation. Blueish colors show temperatures below zero degree Celsius, and yellow to orange colors depict temperatures above the freezing point. Color saturation represents the temperature value. Green color indicates missing/faulty data. Seasons are placed on top of the days and are color-coded based on the average temperature of the season. The top-most horizontal strip represents the yearly average temperature.*

4. Visualization and Interaction Approach

In the following, we discuss our visualization and interaction approaches for the visual analysis of temperature time-series together with a brief discussion on implementation aspects.

4.1. Views

We want to show as much historical temperature data as possible in a single view, while also allowing the user to get insight on selected data ranges. To achieve this, our tool is divided into four distinct parts as shown in Figure 1. The *main view* visualizes the overall temperature at different granular levels (see Figure 1(A)). The *toolbar* holds different buttons that enable various interaction modes of our prototype (Figure 1(B)). The *details view* shows textual information about a selected data point (Figure 1(C)), i.e., the exact date and temperature of the selected/hovered day. When a full range has been selected (see Sect. 4.2), the highest and lowest temperature dates of the selection appear. Finally, the *selection view* shows a timeline visualization of the selected data (Figure 1(D)). In the following, we explain the main and selection views in more detail.

Main View Our initial inspiration came from pixel-based approaches [KKA95]. The advantage of such approaches is that they make maximal use of the available screen space, i.e., we could map the temperature of one day to a single pixel. However, interacting with pixel-based views is sometimes cumbersome due to the small pixel size. Moreover, we wanted to include aggregated data in the form of (meteorological) season and year averages. As each calendar year is divided into four seasons comprised of 365 days, we used a variation of icicle plots to show the daily temperature of a year while providing insight into the averages for the individual seasons and complete years, cf. Figure 2. To build an overview of the entire data set, we stack the visual representations for all years in ascending order as seen in Figure 1(A). We use vertical sliders to navigate the data when dealing with large data sets. One issue of this strategy is that the winter seasons are not mapped perfectly to the years, because they start at the beginning of December and end at the end of February. Therefore, instead of four, we have five segments in the seasons levels as each winter season is split into two parts.

Selection View The selection view (Figure 1(D)) uses a traditional timeline approach and accentuates the highest and lowest temperature dates of the selection as well, while providing context to the other days. The extreme temperatures are marked with blue and red vertical lines respectively. While the textual information of the *details view* gives exact values of such data, the *selection view* provides additional insight into the condition of such extreme temperatures. For example: is the coldest day the result of a gradual temperature drop, or was it a sudden event that happened within a time span of 1-2 days? Those questions can be answered with the help of the timeline plot as shown in Fig. 3.

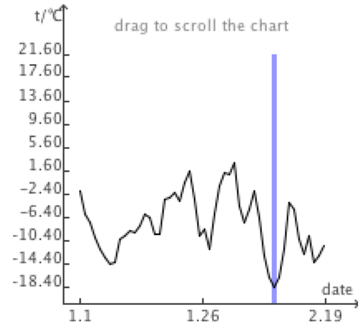

Figure 3: *The* selection view *uses a timeline plot to represent different selection ranges. The blue vertical line represents the lowest temperature in a given range. The timeline can be dragged and scrolled horizontally to view the rest of the data. Zooming is not supported.*

4.2. Interaction

Initially, our tool shows all scale levels at once. Hiding some levels might be useful to reduce clutter. For instance, if users are only interested in season temperatures, then they can hide the day levels—or even the year levels—to fit more data into the view. When a user decides to hide or show a scale level by clicking the toolbox buttons, our tool must redraw the entire view. In order to soften the effect of loosing the mental map in such cases, we use a simple morphing technique to smooth vertical transition changes.

Although we do not represent the days as single pixels in our approach, their width is still small enough to hinder easy mouse selection, especially if the neighboring days share similar temperature values. This makes it even harder to distinguish the boundaries of each day. On the one hand, we overcome this issue by using the *details view* which shows the data of the day node under the mouse cursor. On the other hand, we have implemented an one-dimensional fisheye lens to enable comfortable selection by magnifying the day nodes interactively. In consequence, nodes in the center of the lens will be magnified whereas border nodes will be compressed in the same day level (cf. the conceptional behavior of our fisheye lens in Fig. 4 and the arrow in Fig. 1).

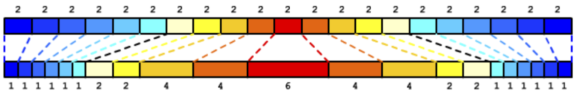

Figure 4: *A conceptual schema of the one-dimensional fish-eye lens for day scale level with a size of 21 node elements. The labels above/below the nodes indicate how many pixels the related nodes encompass without (top) and with (bottom) using the lens.*

Besides single node selection, our tool offers selection possibilities across full horizontal or vertical ranges, i.e., users can select all the days or seasons of a year, or select one particular day, a season or all years of the entire data set. A 'horizontal range selection' can be used to answer questions about extreme temperatures within a year. A 'vertical range selection' can give answers to questions like: what were the hottest and coldest Christmas days in Gothenburg? The answer in this case is sufficiently perceivable in Fig 1. Here, the user has selected the vertical range for December 25 by using the toolbar and mouse selection (see Figure 1(B) and the arrow in the bottom-right corner of the *main view*). The *details view* shows us then the lowest and highest temperature recorded during the last 52 years.

4.3. Implementation Details

The prototype is implemented in Java in order to support multiple platforms [Liu12]. No third-party visualization libraries were used, i.e., the visualizations were programmed in pure Java Graphics2D without hardware acceleration. Our visualization approach mainly involves the drawing of rectangles. The number of rectangles to be drawn depends on the size of the data set and is calculated as $n = d + (1+5)y$, where d is the total number of data records (i.e., all days of the time range), y is number of years, and n is the total number of rectangles. For each year $y = d/365$, we need six rectangles to represent the year itself and the season level (leap year excluded). Thus, 7,420 rectangles have to be drawn if the data set includes records for 20 years, for instance. Rendering these objects once is no problem. However, the system would become too slow to handle the smooth interactions discussed in the interaction part of Section 4. Therefore, we have incorporated an image buffering mechanism that allows real time interaction even with data sets of $n \approx 37,000$.

5. Conclusions and Future Work

We presented a new visualization approach for the analysis of temperature trends of one specific location over years or even decades. An important aspect of this work are the different aggregation possibilities that are smoothly integrated within an interactive view. Even if our tool was originally designed for temperature data, it can also be used to show any type of quantitative data, such as atmosphere pressure, humidity, wind condition, or rainfall. However, we have not implemented efficient comparison techniques for such tasks by now. Those have to be added to the tool. Another aspect of future improvements are multiple selections that are needed to perform more complex analysis tasks and more support for zooming operations. In the light of those useful extensions, our current tool can be considered as an initial step towards analyzing climate data and should also be validated by user studies in the future.

References

[AMM*08] Aigner W., Miksch S., Muller W., Schumann H., Tominski C.: Visual methods for analyzing time-oriented data. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 14*, 1 (2008), 47–60. 2

[KKA95] Keim D., Kriegel H.-P., Ankerst M.: Recursive pattern: A technique for visualizing very large amounts of data. In *Proceedings of the IEEE Conference on Visualization (Vis '95)* (Oct 1995), pp. 279–286. 2, 3

[LAB*09] Lammarsch T., Aigner W., Bertone A., Gartner J., Mayr E., Miksch S., Smuc M.: Hierarchical temporal patterns and interactive aggregated views for pixel-based visualizations. In *Proceedings of the 13th International Conference on Information Visualisation (IV '09)* (2009), pp. 44–50. 2

[Liu12] Liu J.: Visualization of weather data: Temperature trend visualization. Bachelor's thesis, Linnaeus University, School of Computer Science, Physics and Mathematics, Växjö, Sweden, 2012. 4

[LSL*09] Ladstädter F., Steiner A. K., Lackner B. C., Pirscher B., Kirchengast G., Kehrer J., Hauser H., Muigg P., Doleisch H.: Exploration of climate data using interactive visualization. *Journal of Atmospheric and Oceanic Technology 27*, 4 (2009), 667–679. 2

[MFSW97] Mintz D., Fitz-Simons T., Wayland M.: Tracking air quality trends with SAS/GRAPH. In *Proceedings of the 22nd Annual SAS User Group International Conference (SUGI '97)* (1997), SAS, pp. 807–812. 2

[NSBW08] Nocke T., Sterzel T., Böttinger M., Wrobel M.: Visualization of climate and climate change data: An overview. In *Digital Earth Summit on Geoinformatics 2008: Tools for Globale Change Research* (2008), Wichmann, Heidelberg, pp. 226–232. 2

[SFdOL04] Shimabukuro M., Flores E., de Oliveira M., Levkowitz H.: Coordinated views to assist exploration of spatio-temporal data: A case study. In *Proceedings of the 2nd International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV '04)* (2004), pp. 107–117. 2

[TSWS05] Tominski C., Schulze-Wollgast P., Schumann H.: 3D information visualization for time dependent data on maps. In *Proceedings of International Conference on Information Visualisation (IV '05)* (2005), pp. 6–8. 2

[VW14] Viégas F., Wattenberg M.: Wind Map, last accessed: 20-04-2014. http://hint.fm/wind/. 2

[Wea14] WeatherSpark: Beautiful Weather Graphs and Maps, last accessed: 20-04-2014. http://weatherspark.com. 2