
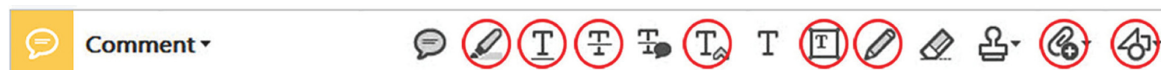


Page Proof Instructions and Queries

Journal Title: IVI
Article Number: 693041

Greetings, and thank you for publishing with SAGE. We have prepared this page proof for your review. Please respond to each of the below queries by digitally marking this PDF using Adobe Reader (free at <https://get.adobe.com/reader>).

Please use *only* the circled tools to indicate your requests and responses, as edits via other tools/methods are not compatible with our software. To ask a question or request a formatting change (such as italics), please click the  tool and then choose “Text Callout.” To access the necessary tools, choose “Comment” from the right-side menu.



Sl. No.	Query
	GQ: Please confirm that all author information, including names, affiliations, sequence, and contact details, is correct.
	GQ: Please review the entire document for typographical errors, mathematical errors, and any other necessary corrections; check headings, tables, and figures.
	GQ: Please ensure that you have obtained and enclosed all necessary permissions for the reproduction of artworks (e.g. illustrations, photographs, charts, maps, other visual material, etc.) not owned by yourself. please refer to your publishing agreement for further information.
	GQ: Please note that this proof represents your final opportunity to review your article prior to publication, so please do send all of your changes now.
	GQ: Please confirm that the funding statement is accurate.
1	AQ: Please check whether the inserted expansion for ‘GAV’ is correct.
2	AQ: Per journal style, page number for Ref. 25 has been moved to reference list. Please check.
3	AQ: Please check the sentence ‘They can be shown as scatterplot points ...’ for clarity.
4	AQ: Please check the sentence ‘However, the question remains ...’ for clarity.
5	AQ: Please check the sentence ‘Four issues are usually mentioned as ...’ for clarity.
6	AQ: Please check the sentence ‘Such feature can be beneficial for ...’ for clarity.
7	AQ: Please check the sentence ‘However, if polyline vertices form ...’ for clarity.
8	AQ: Please check whether variables/terms/functions/Greeks are accurately and consistently used in all instances.
9	AQ: Please check the sentence ‘Then, to check whether the chosen indicators ...’ for clarity.
10	AQ: Please check the sentence ‘The participants report moderate ...’ for clarity.
11	AQ: Please check the sentence ‘Regarding the attitudes toward the ...’ for clarity.
12	AQ: Please check the sentence ‘Regarding the PCP and the polyline glyphs ...’ for clarity.
13	AQ: Please check the sentence ‘This finding confirms the ...’ for clarity.
14	AQ: Per journal style, the references and their citations have been changed from ‘name and date’ to ‘numbered’. Please check and confirm.
15	AQ: Please note that the given URL in Ref. 30 does not lead to the desired web page. Please check.
16	AQ: Please check whether Figures 1 -15 are correct as set.
17	AQ: Please check whether Tables 1 -5 are correct as set.

Augmenting the usability of parallel coordinate plot: The polyline glyphs

Tomasz Opach and Jan Ketil Rød

Information Visualization
1–20
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1473871617693041
journals.sagepub.com/home/ivi


Abstract

Polyline glyphs are minimized thumbnails of polylines from parallel coordinates. Since such glyphs may augment the usability of parallel coordinates, the authors investigate whether there are benefits to be derived from using polyline glyphs that are dynamically linked to parallel coordinates as opposed to the use of the latter visualization technique alone. They also identify user tasks that can be effectively solved if parallel coordinates dynamically linked to polyline glyphs are used. This study adds to the body of previous work a discussion on the features of the polyline glyphs that facilitate the exploration and understanding of multivariate data. Moreover, the authors conduct an empirical study in which parallel coordinates dynamically linked to polyline glyphs are used to solve four tasks. The main finding is that polyline glyphs can facilitate a better insight into the similarities between the multivariate signatures of data items and information acquisition if visual clutter hinders the use of parallel coordinates. The study also reveals that if visual clutter does not occur in parallel coordinates and the polylines from the latter can be differentiated, individuals tend not to use polyline glyphs to study multivariate signatures.

Keywords

Polyline glyphs, glyphs, parallel coordinate plot, geovisualization, user study

Introduction

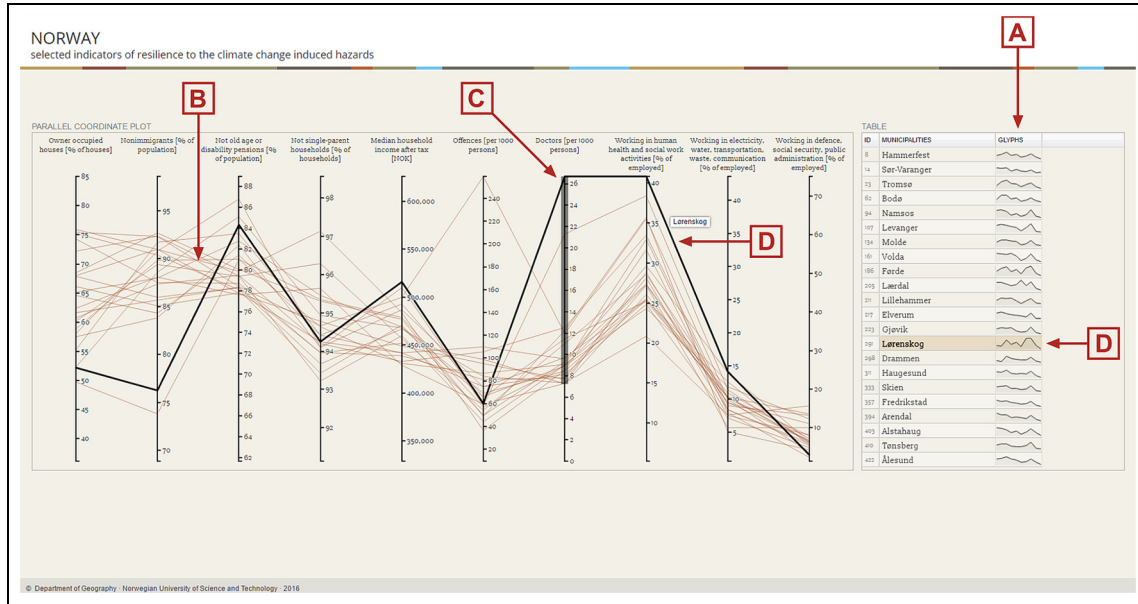
Parallel coordinate plot (PCP)¹ is a technique for visualizing multidimensional data. It shows multivariate signatures as superimposed on each other, which generates advantages and causes inconveniences. Superimposition can facilitate comparison between multivariate signatures and can require little space, even for lots of data items. These are fundamental advantages especially if limited space must be used for showing lots of data items at once and for making comparison feasible. A substantial weakness, however, is that superimposition can lead to visual clutter,² making users reluctant to use PCPs³ as they may perceive PCPs as “too scientific.”⁴ Although there are various clutter-reduction techniques,² they all have certain constraints. Showing multivariate signatures not as superimposed but as juxtaposed (i.e. placing them below or next to each other) can be a remedy. However, juxtaposition does not offer the same

features as superimposition, and therefore, it cannot be treated as a substitute for the latter. Although juxtaposition provides easy access to particular multivariate signatures, it takes up more space and can make comparison difficult, especially if hundreds or thousands of multivariate signatures are shown. Hence, both superimposition and juxtaposition seem viable and useful depending on the extent to which they enable insights. Therefore, showing multivariate signatures as juxtaposed with each other may augment the insight gained when they are viewed as superimposed on each other in parallel coordinates.

Department of Geography, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding author:

Tomasz Opach, Department of Geography, Norwegian University of Science and Technology, Dragvoll, NO-7491 Trondheim, Norway.
Email: tomasz.opach@svt.ntnu.no



AQ16 **Figure 1.** The user interface used in the empirical study. Polyline glyphs (A) dynamically linked to a parallel coordinate plot (B) may facilitate investigations of multivariate signatures of data items presented in the latter plot. Although the signatures are filtered (C), they might still be difficult to differentiate, due to their superimposition. The polyline glyphs (A) could therefore be used to investigate the signatures of selected items: in this case, Lørenskog (D).

Our goal is to investigate whether showing multivariate signatures of data items as both juxtaposed and superimposed in a PCP can augment the gained insight and hence also the usability of the latter visualization technique. Since such a research objective is multifaceted, we narrow our focus to a particular case in which a specific design pattern is used for juxtaposing signatures. Drawing upon inspiration from parallel coordinates and Tufte's sparklines,⁵ we propose the polyline glyphs. These condense what is typically found in parallel coordinates into tiny multivariate glyphs⁶ that can be regularly sized and graphically manipulated (Figure 1). Polyline glyphs are simply minimized and frozen thumbnails of plots from parallel coordinates. Such thumbnails are then grouped into a glyph display and dynamically linked with a PCP, as we assume that using polyline glyphs to show multivariate signatures as juxtaposed can provide advantages over the use of a PCP alone.

Parallel coordinates dynamically linked to polyline glyphs have so far been implemented experimentally twice: in ViewExposed⁴ and in agroEXPLORE.⁷ Although the implementations gave promising results, no evidence was gathered to support the assertion that polyline glyphs add value to a coordinated-view display, thus making the system more useful. With this study, we intend to better investigate this issue in more detail. First, we briefly elaborate on the features of the polyline glyphs in the context of geovisualization, as this area of application is of primary importance to us.

Second, we investigate their effectiveness in the case where they are dynamically linked to a PCP. We therefore conduct a controlled user study to investigate whether the polyline glyphs are well suited to positively augment (not replace) the primary PCP views. In this article, we address the following two research questions:

1. Do polyline glyphs dynamically linked to a PCP augment insights into multivariate geographical data drawn from geovisualization provided through a PCP alone?
2. What user tasks, if any, can be solved if task-solving activities are supported by the dynamic linking of a PCP to polyline glyphs?

This article adds to the body of previous work, both through a discussion on specific features of the polyline glyphs and through a controlled user study in which the effectiveness of their linkage to a PCP is investigated.

Parallel coordinates and multivariate glyphs in geovisualization

Facing parallel coordinates' visual clutter in geovisualization

Parallel coordinates give insights into the multivariate signatures of data entities.⁸ They therefore have many

application domains. In geographic information science (GIScience), parallel coordinates are often used as a “multidimensional detective,”⁹ both independently as a singular data display and as part of composite environments with coordinated and multiple views (CMVs), such as CommonGIS,¹⁰ HealthVisPCP,^{8,9} and ESTAT.¹¹

Although PCPs have advantages, they also feature shortcomings such as limited efficiency in conveying significant patterns due to visual clutter.^{2,3} Although tools such as Mondrian¹² exist, in which dense parallel plots are perfectly usable (e.g. through alpha blending, sampling, or parallel box plots that provide viable alternatives), many data items drawn as superimposed polylines on a limited screen space normally hinder visual analysis. Users, especially those unfamiliar with visual analytics tools, might be unable to reveal trends or perform even simple tasks such as polyline selection.¹³

In the context of CMV geovisualization, parallel coordinates’ visual clutter can be solved through clutter-reduction techniques² such as clustering³ and, as in Mondrian, alpha blending.¹² It can also be done through dynamically linking a PCP to a display that shows multivariate signatures not as superimposed but as geographically or topologically distributed. A good example of the latter is multivariate symbol maps, such as those by Dorling¹⁴ in which both multivariate glyphs and peculiar Chernoff¹⁵ faces are used. Visual clutter in parallel coordinates can also be overcome by dynamically linking a PCP to a technique that shows the signatures not superimposed but juxtaposed, for instance, in a grid plot consisting of multivariate glyphs, as in the GeoViz toolkit.¹⁶

All three clutter-reduction solutions have limitations. For example, alpha blending can make the identification of polylines in a crowded display unfeasible, and showing numerous multivariate symbols on a map causes visual clutter too. Besides, there are only few methods of multivariate symbol mapping.¹⁷ The third approach to solving the problem of clutter requires more screen space. Nevertheless, since it gives easy access to particular data items, it seems promising. A grid plot consisting of multivariate glyphs and in which the same data entities are juxtaposed as those superimposed in a PCP may not only facilitate information acquisition: as the same information is shown twice but differently, it can also augment the insights gained through the PCP and thus augment its usability.

Glyphs in geovisualization

Glyphs are commonly defined as geometric shapes scaled by the values of multivariate data where each glyph represents one multidimensional data item.^{18,19}

They have different shapes^{6,20} and are therefore referred to by different names, such as line glyphs and star glyphs. Simplified and small glyphs typically feature neither coordinates nor axes, which makes them similar to Tufte’s⁵ concept of sparklines. The border between these concepts is therefore fuzzy. Glyphs are used in map displays, which are named after the glyph style, such as star-plot maps and “utility symbol” maps.²¹ However, in addition to static or interactive maps, multivariate glyphs can be used in other ways in geovisualization. Glyphs can constitute an autonomous part of CMV environments, as in GeoVISTA Studio²² and the Geovisual Analytics Visualization (GAV) framework,²³ in which multiple visualization techniques are dynamically linked to facilitate information exploration and knowledge construction.¹⁰ Then, to better support analytical purposes, glyphs can be grouped into a data display²² and dynamically linked to map or data displays. They can be shown as scatterplot points^{18,19,24} or juxtaposed in a small multiple, in which “information slices are positioned within the eyespan, so that viewers make comparisons at a glance—uninterrupted visual reasoning”²⁵). As Ward²⁰ and Klippel et al.²⁶ claim, due to such settings, the similarities and differences between glyphs are likely to be identified more efficiently.

As element of CMV tools, juxtaposed multivariate glyphs can serve either as an extra visual component for strengthening tools’ overall quality or as a supportive component for amplifying the communication provided in other displays (i.e. by augmenting their usability). Tools employing PCPs are a good example of when such a dual role of glyphs is potentially of value. Juxtaposed glyphs can, for example, be star-plot glyphs, and since such glyphs can be thought of as a PCP in polar coordinates^{18,19} or as an alternative for representing the data shown in parallel coordinates,²⁶ visual attention can be shifted from parallel coordinates to star-plot glyphs without adjusting visual reasoning. Nevertheless, there is a difference in the way in which the data are presented. In PCPs, all entities are superimposed on each other. Moreover, different intervals for each axis can be shown by labeling and hash marks, and this works for categorical values as well. Star-plot glyphs in turn give an overview of the entire dataset by showing each graphical entity separately in a form of juxtaposed small star plots but without any labels. Hence, despite the limited explanations (labeling), juxtaposition enables access to particulate data items, which often is unfeasible in superposition. The approach can further be simplified. Instead of using polar coordinates, parallel coordinates can be applied so that graphical entities resemble the polylines from a PCP. Such “polyline glyphs” can therefore serve as an extra visualization component and be a

AQ1

AQ2
AQ3

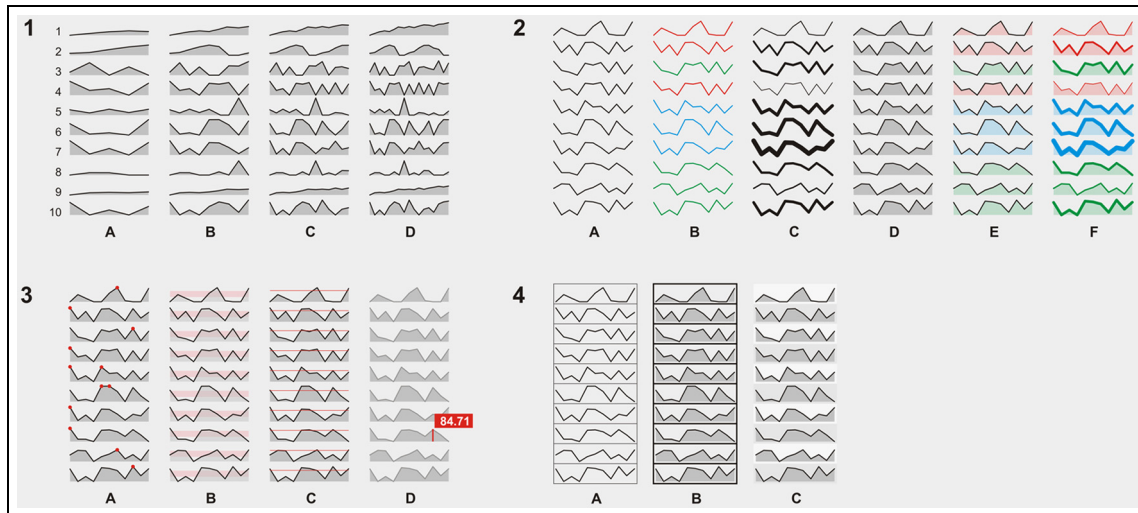


Figure 2. Variations in styles of polyline glyphs: (1) polyline glyphs drawn for 5- (1A), 9- (1B), 13- (1C), and 17-dimensional (1D) data items; (2) variations in styles of glyphs drawn for 11-dimensional data items; (3) examples of techniques that can facilitate the interpretation of polyline glyphs; (4) background style examples (see the main text for details).

supportive visualization technique for parallel coordinates. However, the question remains of what, if any, is the added value of using such glyphs and whether they augment the insights provided by a PCP, which would otherwise be unachievable when using a PCP alone.

AQ4

The polyline glyphs

Four issues are usually mentioned as essential for the effectiveness of glyph-based visualization in general:^{6,18,20,24,27,28} form, multidimensionality (choice of included variables), style and size, and axis ordering within each glyph as well as ordering of glyphs in a display. Although the first issue is not of importance to the polyline glyphs, the remaining three issues should be addressed when designing such glyphs.

AQ5

Multidimensionality

In common with parallel coordinates, polyline glyphs can show data items with either nonrelated or related attributes (e.g. time-series data on one variable). In both cases, at least two variables (attributes) assigned to a data item are required to generate a (poly)line. With growing dimensionality, it might become more difficult to distinguish particular vertices (variables) within a polyline and therefore, only the general shapes of glyphs—their multivariate footprints (signatures)—can be investigated. Such feature can be beneficial for certain user tasks, such as when investigating general

trends within multivariate data records featuring related attributes.

AQ6

Figure 2 shows 10 glyphs drawn for 5- (1A), 9- (1B), 13- (1C), and 17-dimensional (1D) data items. With increasing dimensionality, the glyphs become more complex and less legible. However, if polyline vertices form regular shapes such as curves or waves (glyph #2), align linearly (#1 and #9), or feature something extraordinary such as “peaks” (#5 and #8), patterns or nuances seem to be easily recognizable. If the variation in attributes is substantial and there is nothing distinctive in glyphs’ shapes (#3, #4, #6, #7, and #10), it might be difficult to make comparisons across graphical entities, to identify patterns or nuances, or to memorize certain glyphs. However, if a given system provides some auxiliary data analysis techniques such as a similarity-check mechanism, then the comparison across multivariate glyphs is feasible even if they feature substantial variation in attributes.

AQ7

Styling

Not only vertices but also other features can encode attributes. The simplest glyphs take the form of a unicolor and unistyle polyline (Figure 2, 2A). This basic style works like Turkay et al.’s²⁹ attribute signatures or line-graph sparklines, which are typically used for monitoring stock-market activity⁵ or other time-series data³⁰ where relative differences between general shapes or between succeeding values in a line graph can be read. Similarly, the simplest style can be suitable for showing relative changes (e.g. over time) on

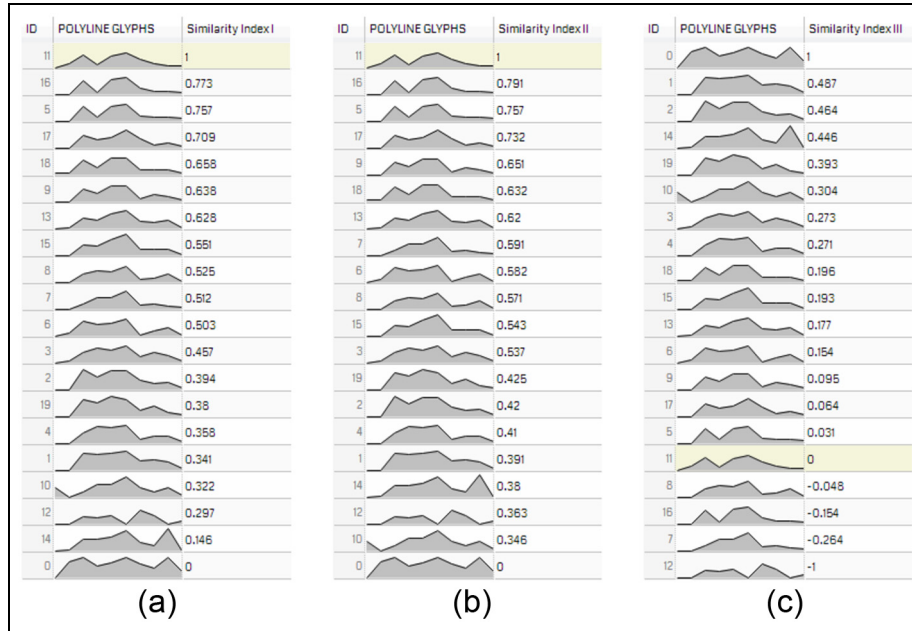


Figure 3. Twenty 10-dimensional glyphs sorted in descending order (from 1 to 0) or divergently (from 0 to either 1 or -1) with respect to their similarity to glyph #11. Three approaches are used to calculate the similarity: (a) the Euclidean distance in the n -dimensional space approach, (b) the city-block distance approach, and (c) the modified city-block distance approach (the equation proposed in this study).

one variable rather than absolute values, proportions, or percentages. There are also other ways to style polyline glyphs. For example, either various colors (Figure 2, 2B) or line widths (2C) can be used to indicate categorical or ordinal data, respectively. Fill can also be employed; for example, unicolor fill (2D) can facilitate reading absolute values, whereas a multicolor fill (2E) can additionally enable the grouping glyphs into categories. Example 2F in Figure 2 shows a combination of various styles. As this example employs most visual variables,³¹ its polyline glyphs feature the highest information density. However, the reader has to expend more effort on interpreting encoded information.

Aside from solutions that—like those mentioned above—facilitate the encoding of extra information in polyline glyphs, there are techniques that can facilitate interpretation, such as the lowest and highest spots (Figure 2, 3A), ranges (B), thresholds (C), and tool-tips (D). These techniques do not necessarily work for nonrelated attributes; however, they can be applied if attributes are transformed into a common range. Although polyline glyphs use aligned axes and, in comparison with star glyphs, it is easier to use them, for example, to detect maxima and minima, mentioned techniques can facilitate interpretation. Indicating the highest spot can facilitate finding the highest values in a multivariate signature if the variables do not differ

very much or if there is more than one variable with the highest value.

Not only polyline glyphs in themselves but also, as Ward⁶ emphasizes, the way they are arranged in a data display is of importance to readers. Size, background, alignment, and space in between glyphs are all features that might influence user satisfaction, and thus communication efficiency. Figure 2 (4) shows three background settings based on a regular arrangement. The main advantage of regular grids is that polyline glyphs fit their cells and the screen space is used efficiently, which is not always feasible for star glyphs, the sizes of which can vary significantly. When hundreds of polyline glyphs are regularly spaced, the background can make it easier to interpret them (Figure 2, 4C). However, for sparklines, Tufte⁵ does not recommend the use of frames that surround data lines because that might generate unintentional optical clutter (Figure 2, 4A, B). A poor visual effect can also occur if polyline glyphs appear too light (4A) or cell frames appear too heavy (4B). If the balance is kept, the layout is legible (4C).

Attribute and glyph reordering

Attribute reordering within a glyph and glyph reordering in a display are both of special importance for polyline glyphs. If these are dynamically linked to a

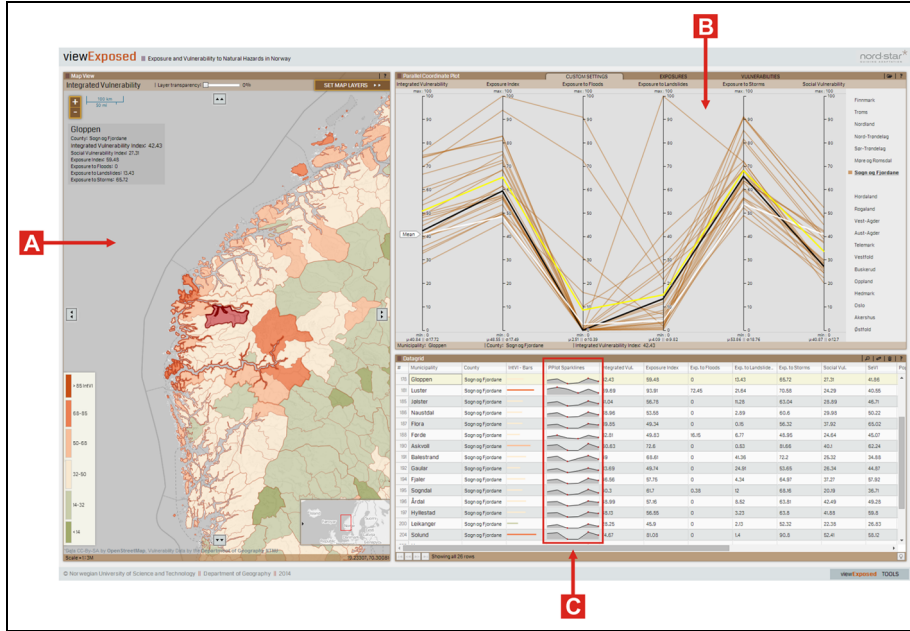


Figure 4. In ViewExposed, polyline glyphs (C) are embedded into the datagrid. Here, the municipality of Gloppen is selected in the PCP (B) and the corresponding row with a polyline glyph and a polygon is highlighted both in the datagrid (C) and on the map (A), respectively.

PCP, axes' reordering in the PCP may result in the attribute reordering within the glyphs. Such reordering can be executed either manually or through optimization algorithms.³ The second fundamental feature of polyline glyphs is reordering in a display. The main benefit of such reordering is that multidimensional visual search is supported, and therefore, comparative analysis of multidimensional attributes between glyphs is feasible.²⁴ Glyph sorting through the use of similarity indices can be helpful if one wants to select a polyline glyph and reorder the remaining glyphs with respect to the similarity to the selected glyph. Many solutions exist,³² such as the commonly used distance-based similarity measures,^{32,33} examples of which are the Euclidean distance in n -dimensional space³⁴ and the city-block metrics.³⁵ For certain reasons, however, these two approaches might not fit the purpose of the visual search in polyline glyphs because the information on whether a glyph is generally “bigger” or “smaller” than the one selected might also be of interest. For such a specific purpose, the following equation normalized to the interval $(-1, 1)$ can be used

$$d(G, B)_{nor} = \begin{cases} \frac{\sum_{i=1}^n (x_{G_i} - x_{B_i})}{d_{max}}, & \text{if } \sum_{i=1}^n (x_{G_i} - x_{B_i}) > 0 \\ \frac{\sum_{i=1}^n (x_{G_i} - x_{B_i})}{-1 \times d_{min}}, & \text{if } \sum_{i=1}^n (x_{G_i} - x_{B_i}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

where d_{nor} is the similarity measure, B is a baseline multidimensional point (selected polyline glyph), G is any multidimensional point (any polyline glyph), n is the number of dimensions, d_{max} is the maximum value observed, and d_{min} is the minimum value (distance) observed. Using this approach, glyphs can be sorted divergently from a selected polyline glyph (a baseline point) with respect to their similarity to that glyph. The approach facilitates the analysis of polyline glyphs with respect to their general size: whether they are smaller (index scores from 0 to -1) or bigger (scores from 0 to 1) than the selected polyline glyph.

Although other approaches exist, the above techniques illustrate how similarity metrics work. Figure 3 shows twenty 10-dimensional glyphs generated for data items randomly selected from real data normalized to the interval $(0, 100)$. The glyphs are ordered with respect to their similarity to one randomly selected glyph #11. In Figure 3, examples A, B, and C show the results for the Euclidean distance, the city-block metrics, and the equation provided above, respectively. At a glance, the glyphs close to the selected one share a similar signature. Moreover, the results from the Euclidean distance and the city-block metrics are fairly similar. For both, the glyphs most similar to glyph #11 are glyphs #16, #5, and #17. The remaining glyphs are located similarly and only glyph #15 has a different position.

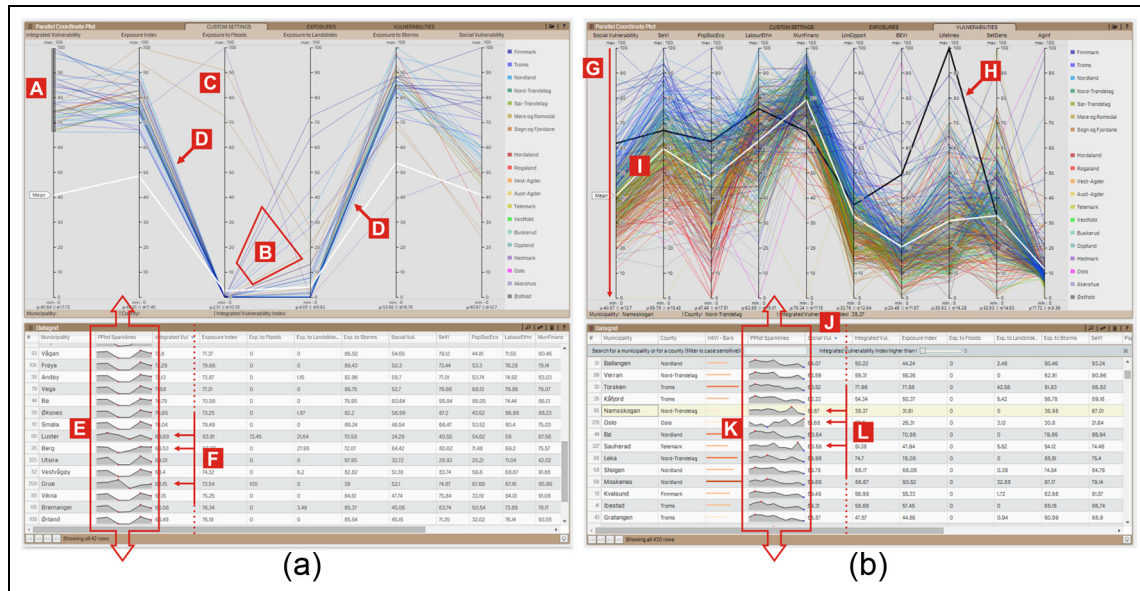


Figure 5. The polyline glyphs in ViewExposed may help (a) identify discords within the multivariate signatures and (b) recognize distinctive multivariate signatures of related attributes (see the text for details).

Although modern software makes it easy to implement polyline glyphs, it is of primary importance to define tasks, solving which might be facilitated by such glyphs. For example, dynamically linking polyline glyphs to parallel coordinates may support tasks in cases when it is important to investigate multivariate signatures presented as both superimposed and juxtaposed. However, first, it is necessary to identify the tasks and whether they are of importance to potential users.

Dynamic linking of parallel coordinates to polyline glyphs: what does it facilitate?

As the two proof-of-concept implementations of polyline glyphs have already been described elsewhere,^{4,7} we only briefly demonstrate how the dynamic linking of parallel coordinates to polyline glyphs may work.

Polyline glyphs (PPlot sparklines) in ViewExposed

ViewExposed is a tool that enables Norwegian decision-makers to identify the location of municipalities that are vulnerable to natural hazards (Figure 4, A) and to understand why they are vulnerable. Its integrated vulnerability index (IntVI) is composed of a number of sub-indices,³⁶ and insight into them is necessary to understand what influences the final IntVI scores. Parallel coordinates (Figure 4, B) are used to augment this understanding, but since they are over-plotted with more than 400 polylines, access to each

of those polylines is difficult. Minimized snapshots (thumbnails) of the polylines, named PPlot sparklines (Figure 4, C), are therefore juxtaposed with each other and dynamically linked to the PCP to allow access to particular data items and to facilitate comparative analysis of their multidimensional signatures.

Two examples demonstrate the possible usage of the glyphs in ViewExposed. First, the glyphs' role is to help identify any discord in the way that the sub-indices influence municipalities' scores on the IntVI. Consider 42 municipalities (data items) that score above 65 on the IntVI (which ranges from 0 to 100). After filtering the polylines in the PCP (Figure 5, A), one can see a general pattern in how the sub-indices influence the IntVI of filtered municipalities as well as how their signatures vary. Several municipalities feature distinctive signatures as they score higher on exposure either to landslides (B) or to floods (C). Because the polylines are superimposed in the PCP (D), it is difficult to examine their shapes and quickly identify municipalities. We assume that the polyline glyphs may facilitate such identification. One can quickly scroll through the 42 glyphs (E) corresponding to the filtered polylines and identify polylines that have shapes that do not match the neighboring polylines (F). One essential inconvenience might result from the number of polylines to be scrolled through, since the more the polylines, the more demanding the visual comparison between juxtaposed glyphs.

Second, the glyphs may also help reveal distinctive multivariate signatures of related attributes without filtering in the PCP. To investigate, for example, factors

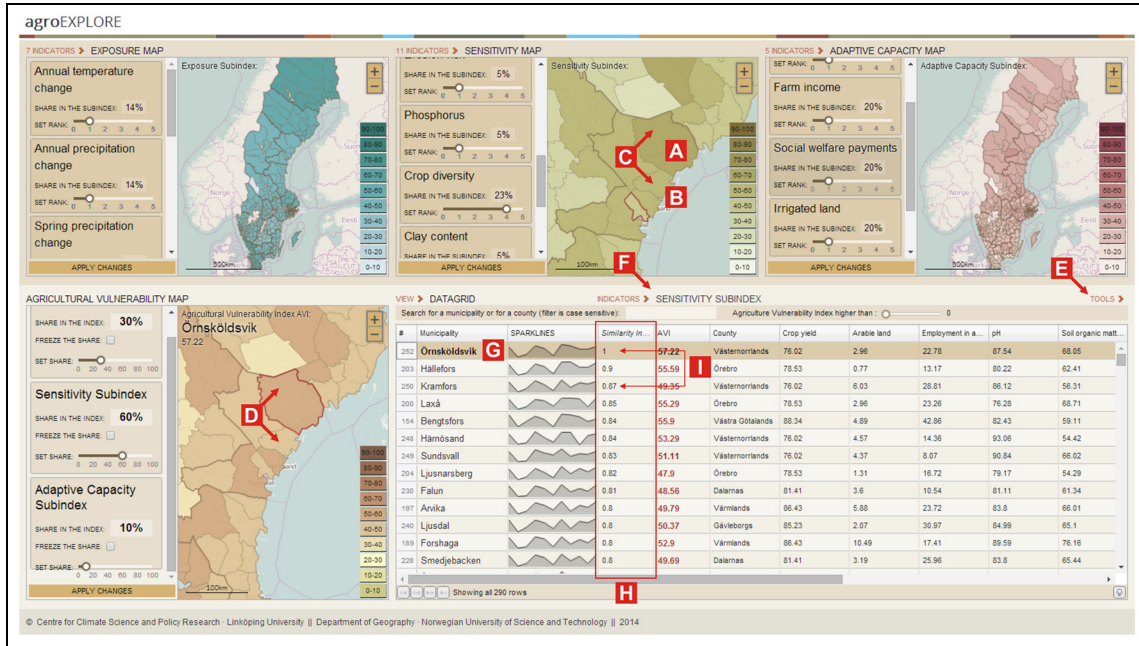


Figure 6. In agroEXPLORE, even if Örnsköldsvik municipality (A) and Kramfors municipality (B) look differently on the choropleth maps (C, D), the glyphs might help reveal similarities between these two municipalities (see the text for details).

that influence the social vulnerability index (SoVI)—one of the components of the IntVI—and identify distinctive signatures within municipalities that score high on the SoVI, one can brush the polylines along the proper axis in the PCP (Figure 5, G) to search for distinctive shapes (H). However, due to the clutter (I), this task might be unfeasible. Therefore, users can use the glyphs instead. The glyphs can be sorted in descending order according to the SoVI scores (J), and then, by scrolling through them (K), it is possible to recognize those that have shapes that do not match the shapes of the neighboring glyphs (L). Hence, municipalities with uncommon signatures can be identified. Such a task resembles a more realistic geospatial analysis context with many data items than the one demonstrated earlier with only 42 items of interests. Nevertheless, whereas in the first example the glyphs augment the utility of the PCP, in this case they serve as an independent component.

Polyline glyphs in agroEXPLORE

The agroEXPLORE tool was designed to enable an agricultural vulnerability index (AVI) to be constructed for Swedish municipalities.⁷ Users can select indicators, adjust their weights, and move between sub-indices on exposure, sensitivity, and adaptive capacity. Additionally, weights can be set for the sub-indices. All the aforementioned actions require on-the-fly insights

into the multivariate signatures of data items. The tool is therefore equipped with a data display in which glyphs can be used either independently or as extension of parallel coordinates.

When assessing agricultural vulnerability, local knowledge is often used to make on-the-fly validations of just calculated vulnerability scores.³⁷ Then, to check whether the chosen indicators and weights give a result that fits local knowledge such as that a municipality should score higher than neighboring places for particular reasons, one needs not only to view the geography of the resulting index but also to compare municipalities' multivariate signatures concerning the indicators used to calculate the AVI. This can be done using polyline glyphs. After assigning indicators to three sub-indices and setting the weights, the AVI scores for municipalities are calculated and shown on the AVI map. Next, in a hypothetical case, one would take a closer look at the municipalities of Örnsköldsvik (Figure 6, A) and Kramfors (B), since they score differently on the sensitivity sub-index (C) and on the AVI (D), even though they border each other and, according to local knowledge, their scores on the AVI should be similar. Therefore, the similarity-check mechanism (E) based on the Euclidean distance is used to check whether these municipalities share similar multivariate signatures regarding the indicators from the sensitivity sub-index (F). After running the similarity-check using Örnsköldsvik as a baseline (G),

AQ9

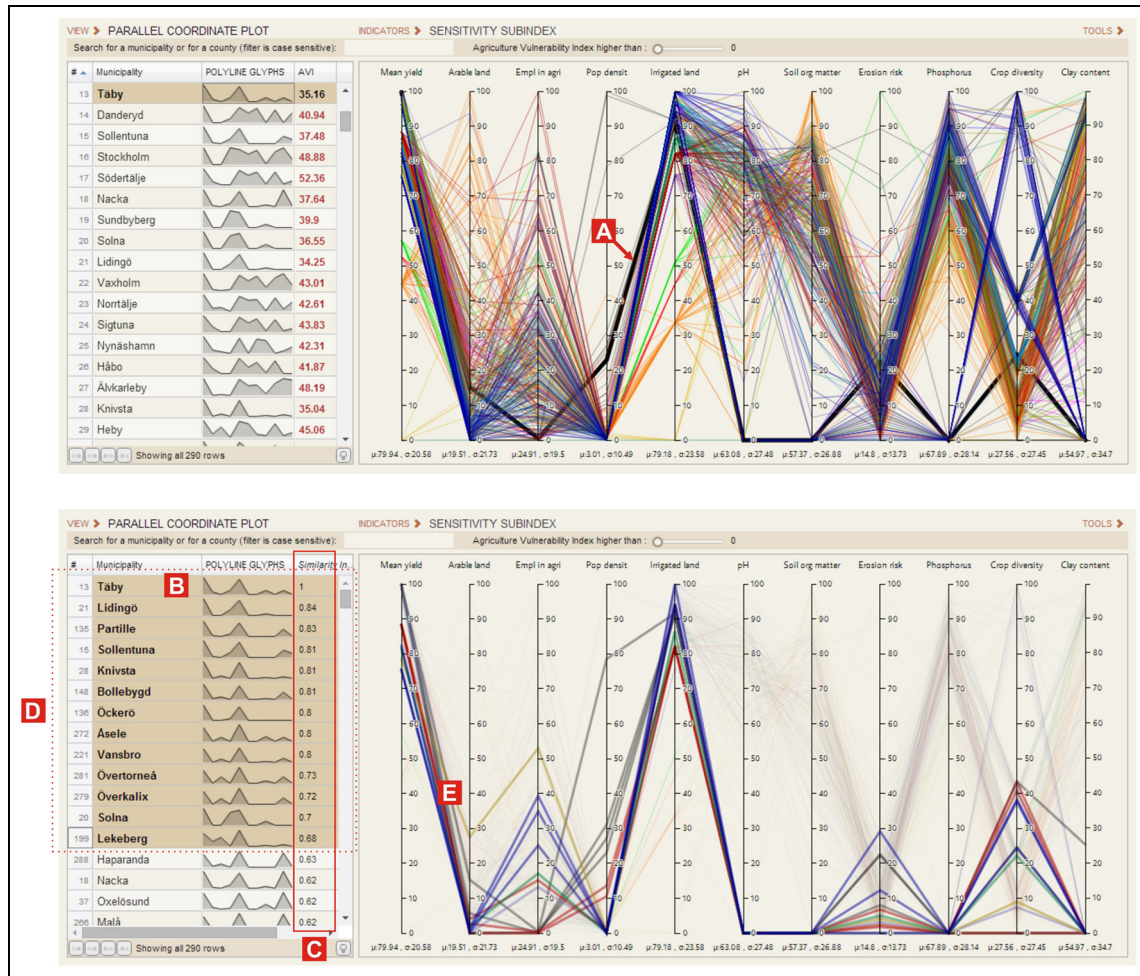


Figure 7. Part of the agroEXPLORE visual interface. After selecting the polyline for the municipality of Tåby (A), it is difficult to identify similar polylines in the parallel coordinates. The polyline glyphs can facilitate the identification. After sorting the glyphs according to their similarity to Tåby municipality (B, C), users can visually identify similar glyphs (D), which can then be highlighted in the parallel coordinates (E).

all municipalities are sorted in descending order with respect to their similarity scores (H). The visual comparison reveals that although Örnköldsvik and Kramfors are categorized differently on the choropleth maps, their signatures are quite similar in terms of the sensitivity sub-index (I). Thereafter, further details can be examined using the PCP.

In the second example, the glyphs may—like in ViewExposed—facilitate the identification of discords within the multivariate signatures that are not clearly legible in a PCP. However, since agroEXPLORE shows items with nonrelated attributes, their signatures can be first sorted according to their similarity to a particular signature, such as that of the municipality of Tåby (Figure 7, A, B), with respect to the similarity scores (C). Congruous glyphs can then be recognized through comparison (D) and the corresponding items

can be left displayed in the PCP (E). Then, the signatures can be compared either as glyphs (D) or in the PCP (E). Both options have advantages and weaknesses. The glyphs enable users to identify municipality names and see the thumbnails of multivariate signatures (D) if they are obviously not too numerous. In turn, the PCP enables users to identify variable names and patterns among selected signatures (E).

Empirical study

Objectives

We conduct an empirical study to acquire evidence of whether the benefits of using a PCP together with polyline glyphs compared to the use of a PCP alone are appreciated by users. Hence, it is of importance to

AQ17 **Table 1.** Four tasks used in the empirical study.

ID	Task	Purpose of task
T#1	In the parallel coordinate plot click on the title of the “Nonimmigrants (% of population)” axis. In the table, municipalities are now sorted descending according to their scores on “Nonimmigrants (% of population).” Investigate the multivariate signatures of municipalities. Briefly comment on how those multivariate signatures vary depending on their decreasing scores on “Nonimmigrants (% of population)”	Examine how users behave when investigating multivariate signatures of data items sorted according to one variable
T#2	Select Sola municipality. Click on the “check similarity” button and use the Euclidean distance method to sort municipalities in descending order according to their similarity to Sola. Click on the “clear selection” button to clear the selection. Investigate the multivariate signatures of municipalities and briefly comment on how they vary depending on their decreasing similarity to Sola	Examine how users behave when investigating multivariate signatures of data items sorted according to their similarity to selected data item
T#3	Use the parallel coordinate plot to select data items that score above 6 on “Doctors (per 1000 persons).” Investigate the multivariate signatures of municipalities. Try to identify distinctive cases. Briefly comment on your findings	Examine how users behave when identifying distinctive cases among multivariate signatures of data items filtered in a parallel coordinate plot
T#4	Select the municipality of Hole. Click on the “check similarity” button and use the Euclidean distance method to sort municipalities in descending order according to their similarity to Hole. Select 20 municipalities that are the most similar to Hole and investigate their multivariate signatures. Try to identify distinctive cases and briefly comment on your findings	Examine how users behave when identifying distinctive cases among multivariate signatures of data items selected in a polyline glyphs display

investigate whether people tend to use a PCP dynamically linked to polyline glyphs in the way outlined in the examples presented above. On the basis of the examples, we formulate two research questions:

1. Are polyline glyphs used as “frozen” thumbnails that help users gain insight into multivariate data besides viewing/investigating them in parallel coordinates?
2. Do polyline glyphs facilitate searches for distinctive multivariate signatures (identifying discords) of data items besides viewing/investigating them in parallel coordinates?

Materials

As study material, we develop a single-page web application in which a PCP is dynamically linked to a table with polyline glyphs (see Figure 1). A selection in one view results in a corresponding selection in the second view. All 428 municipalities (at the time of writing) in Norway are used as multidimensional data items; each is characterized by 10 socioeconomic indicators. Apart from a common functionality, a similarity-check mechanism is implemented because we consider it substantial for polyline glyphs. The three algorithms mentioned previously in the text are used in the mechanism.

Method

Empirical data are gathered through individual user sessions, each of which last up to 45 min. After a 10-min introduction, each participant is given 10 min to become acquainted with the application. Next, each participant is requested to use the application freely to solve four tasks (Table 1) and comment aloud on what he or she does and the conclusions they draw. The tasks are elaborated on the basis of experiences of the proof-of-concept implementations (see section “Dynamic linking of parallel coordinates to polyline glyphs: what does it facilitate?”) and exemplify actions to which the polyline glyphs are tailored. In all tasks, the glyphs are expected to facilitate the identification of distinctive data items by providing insight into their multivariate signatures and by facilitating comparison among those signatures. Nevertheless, the tasks are formulated in a way that does not suggest the use of glyphs. Also, they can be solved by interacting with the PCP so that the participants are able to make their choices freely.

Oral comments as well as screen and mouse movements are recorded during task execution. Moreover, after completing the tasks, each participant fills in a background questionnaire and is asked to respond to a 10-item Likert scale to measure the satisfaction with the design and functionality of the tool. The first six items correspond directly to the visualization methods

used in the display; the remaining items concern the tool's overall usability.

Participants and data collected

The study is conducted in April 2016 at the Department of Geography in the Norwegian University and Technology, Trondheim. A total of 36 individuals (13 females and 23 males) participate. Their skills and knowledge are considered likely to be representative of target users. All but two participants are either Bachelor or Master's students attending Geographic Information Systems (GIS) courses at our university; the remaining two are PhD candidates. Of the 36 participants, 16 are geography students and the remaining 20 represent either the natural sciences such as biology or are enrolled as Master's students in technology programs.

The results of the background questionnaire reveal that 6 of 36 participants (17%) do not use interactive information visualization of any kind, while 25 participants (69%) do so sometimes, mainly when using learning materials in the form of interactive maps or statistical graphics. Surprisingly, 13 (36%) declare that they never use printed statistical graphics. However, 20 (56%) do use them, mainly when studying; only 3 participants (8%) use printed statistical graphics very frequently. With regard to the level of experience in using interactive information visualization and printed statistical graphics (1 for *no experience* and 5 for *everyday use*), the average scores are 2.8 and 2.5, respectively. The participants report moderate levels of training (1 for *no training* and 5 for *proficient*) in GIS and cartography (3.4), and in computer graphics (2.7), and low levels of training in the use of Internet technologies and user interface design (1.9 for *both*).

The stimulus is displayed on a 24-in screen with a resolution of 1920 × 1080 pixels. The study is conducted in English because most of the participants are international students. We record the screen, the participants' oral comments, and the position of the mouse cursor. Thereafter, the recordings are partitioned into segments featuring "consistent user behavior" (i.e. the behavior in which participants act in a certain way, such as hovering over polylines in the PCP and loudly commenting on the glyphs). In the data analysis, we only include those segments in which participants freely investigate multivariate signatures of municipalities, as this is an integral element of each task instruction (see Table 1). The remaining segments are disregarded because in them the participants act in accordance with the requests. We also disregard parts in which the facilitator clarifies task questions or in which technical problems occur. Ultimately, ca. 4 h of

recordings remain, constituting 62% of the original gathered recordings.

Data analysis

User behavior approaches

Ten user behavior approaches (UBAs; Table 2) are differentiated based on the above-mentioned segments (parts of recordings) featuring consistent user behavior. UBAs from #1 to #7 concern dealing with the table view, whereas those from #8 to #10 deal with the PCP. Approaches #2, #3, and #6 (highlighted in gray in Table 2) concern applying the polyline glyphs.

Not surprisingly, the most basic and intuitive UBA—hovering over either table rows with municipality names (UBA #7) or polylines in the PCP (#8)—is demonstrated by most participants. These two UBAs are demonstrated at least once in any task by 94.4% ($\pm 7.5\%$, 95% confidence interval (CI)) and 88.9% ($\pm 10.3\%$, 95% CI) of the participants, respectively (see Table 2). Regarding the UBAs related to the glyphs, commenting on the glyphs while scrolling through the table (#3) is demonstrated at least once in any task by 69.4% ($\pm 15.1\%$, 95% CI) of the participants. Hovering over the glyphs either without (#6) or with preceding single-row selection (#2) is demonstrated at least once in any task by 61.1% ($\pm 15.9\%$, 95% CI) and 19.4% ($\pm 12.9\%$, 95% CI) of the participants, respectively. Furthermore, in T#1 and T#2, UBA #3 is one of the two approaches demonstrated by most participants, whereas hovering over the polyline glyphs (UBA #6) is one of two approaches demonstrated at least once by most participants in T#2. In T#3 and T#4, fewer participants use the glyphs than in the two previous tasks. The fewest participants use the glyphs in T#4, for which the share of people who demonstrate UBAs related to polyline glyphs is the lowest (see the last row of Table 3); the highest share is observed in T#2.

User behavior analysis: aggregated duration and number of usages

The segments differentiated for particular participants and tasks are aggregated according to the demonstrated UBAs. Absolute and relative durations are then analyzed. In addition, the aggregated number of usages (segments featuring consistent UBAs) and absolute aggregated duration per one usage are examined. All the results are shown in Figure 8.

We use a two-way repeated measure analysis of variance (ANOVA) (RM-ANOVA) to determine whether any change in the absolute aggregated duration (i.e. the dependent variable) is the result of the interaction

Table 2. Ten UBAs demonstrated by participants in the empirical study (n = 36).

#	Code	UBA	% of Individuals who demonstrated a given user behavior at least once during the tasks (95% confidence intervals given in brackets)				
			T#1	T#2	T#3	T#4	Any task
1	TRS-TRH	Clicking on (selecting) a row (municipality name) in the table and hovering over other rows (municipality names) in the table	22.2 (±13.6)	44.4 (±16.2)	5.6 (±7.5)	0	50 (±16.3)
2	TRS-TGH	Clicking on (selecting) a row (municipality name) in the table and hovering over other polyline glyphs in the table	5.6 (±7.5)	11.1 (±10.3)	5.6 (±7.5)	0	19.4 (±12.9)
3	T-CG	Commenting on the polyline glyphs, typically when scrolling through the table	36.1 (±15.7)	58.3 (±16.1)	33.3 (±15.4)	13.9 (±11.3)	69.4 (±15.1)
4	TRS	Clicking on (selecting) a single row (municipality name) in the table	2.8 (±5.4)	2.8 (±5.4)	2.8 (±5.4)	0	2.8 (±5.4)
5	TRS-TRS	Performing a multiple selection by clicking on rows (municipality names) in the table	19.4 (±12.9)	11.1 (±10.3)	2.8 (±5.4)	2.8 (±5.4)	33.3 (±15.4)
6	TGH	Hovering over polyline glyphs in the table	30.6 (±15.1)	52.8 (±16.3)	22.2 (±13.6)	13.9 (±11.3)	61.1 (±15.9)
7	TRH	Hovering over municipality names in the table	44.4 (±16.2)	58.3 (±16.1)	33.3 (±15.4)	36.1 (±15.7)	94.4 (±7.5)
8	PLH	Hovering over polylines in the PCP	36.1 (±15.7)	19.4 (±12.9)	75 (±14.1)	80.6 (±12.9)	88.9 (±10.3)
9	P-CPL	Commenting on the PCP content	16.7 (±12.2)	16.7 (±12.2)	38.9 (±15.9)	50 (±16.3)	69.4 (±15.1)
10	PLF	Filtering polylines by drawing ranges along the PCP axes	16.7 (±12.2)	2.8 (±5.4)	2.8 (±5.4)	2.8 (±5.4)	22.2 (±13.6)

For each task, the two most common approaches are shown in italic.

PCP: parallel coordinate plot; UBA: user behavior approach.

Table 3. The application of UBAs related to polyline glyphs during task execution.

UBAs related to polyline glyphs and their combinations	% of Participants (n = 36) who demonstrated glyph-related approaches or their combinations at least once in the tasks (95% confidence intervals given in brackets)				
	T#1	T#2	T#3	T#4	Sum of all cases (n = 144)
Only #2 (TRS-TGH)	0	0	2.8 (±5.4)	0	0.7 (±0.1)
Only #3 (T-CG)	16.7 (±12.2)	8.3 (±9)	22.2 (±13.6)	8.3 (±9)	13.9 (±0.6)
Only #6 (TGH)	11.1 (±10.3)	11.1 (±10.3)	11.1 (±10.3)	8.3 (±9)	10.4 (±0.5)
#2 and #3	2.8 (±5.4)	8.3 (±9)	0	0	2.8 (±0.3)
#3 and #6	16.7 (±12.2)	38.9 (±15.9)	8.3 (±9)	5.6 (±7.5)	17.4 (±0.7)
#2 and #6	2.8 (±5.4)	0	0	0	0.7 (±0.1)
#2, #3, and #6	0	2.8 (±5.4)	2.8 (±5.4)	0	1.4 (±0.2)
All above	50 (±16.3)	69.4 (±15.1)	47.2 (±16.3)	22.2 (±13.6)	47.2 (±1.1)

UBA: user behavior approach.

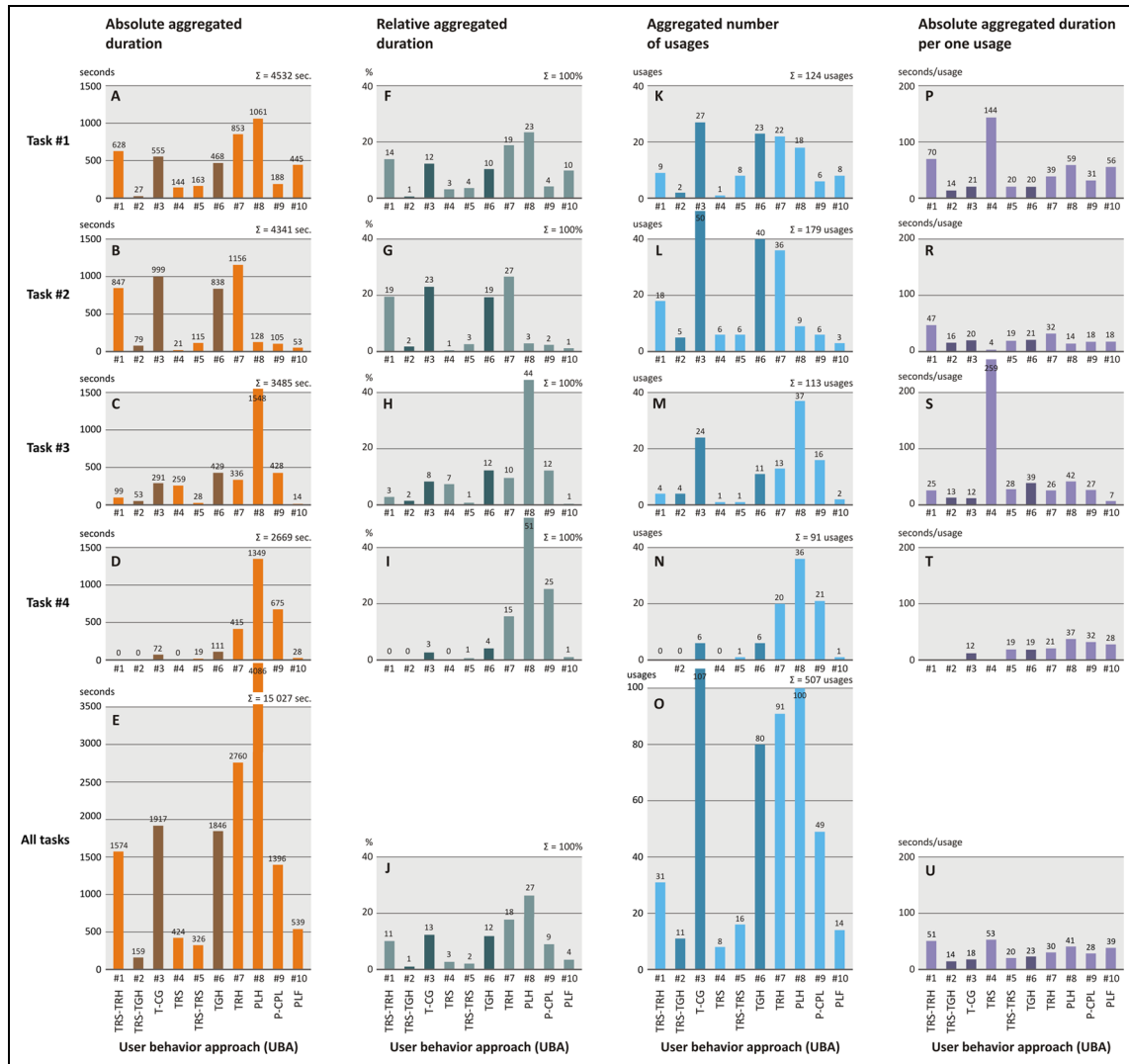


Figure 8. Absolute and relative duration of aggregated UBAs and aggregated number of usages (segments featuring consistent user behavior). The darker bars relate to the UBAs in which polyline glyphs re-used.

between the type of UBA (i.e. UBAs #1–#10) and task type. The test reveals a significant main effect of UBA, $F(5.54, 193.97) = 11.22, p < .001, \eta^2_p = .24$ (Mauchly’s test indicates that the assumption of sphericity is violated, $\chi^2(44) = 177.08, p < .001$; therefore, degrees of freedom are corrected using Greenhouse–Geisser estimates of sphericity ($\epsilon = .62$)), showing that there are significant differences in absolute aggregated duration for the 10 UBAs. The participants spend most time hovering over polylines in the PCP (UBA #8). When it comes to the polyline glyphs, they are used for the longest time to comment on (UBA #3). The test also reveals a significant main effect of task type, $F(3, 105) = 12.11, p < .001, \eta^2_p = .26$, showing that there are significant differences in absolute aggregated duration for the four user

tasks. Moreover, the RM-ANOVA test confirms a significant interaction effect between UBA and task type for absolute aggregated duration, $F(9.04, 316.46) = 5.89, p < .001, \eta^2_p = .14$ (Mauchly’s test indicates that the assumption of sphericity is violated, $\chi^2(377) = 1261.36, p < .001$; therefore, degrees of freedom are corrected using Greenhouse–Geisser estimates of sphericity ($\epsilon = .34$)). It tells us that the profile of absolute aggregated durations across different user tasks is different for the 10 various UBAs.

Regarding the relative duration, the glyphs are used relatively longer in T#1 and T#2, when the summed share of the UBAs relating to glyphs (#2, #3, and #6) is the highest: 23.2% of the participants’ aggregated duration in T#1 (F) and 44.1% in T#2 (G). The role of polyline glyphs is more distinct for all tasks when

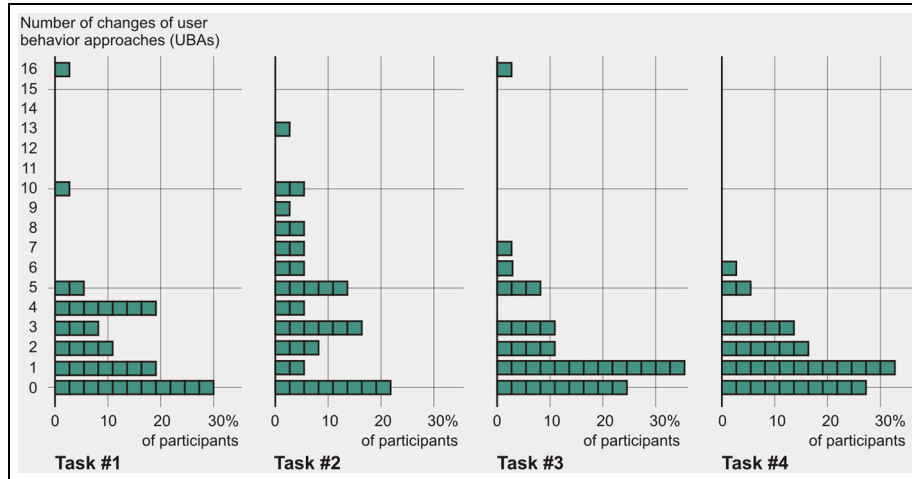


Figure 9. Number of changes in user behavior approaches (UBAs) during task execution, by task.

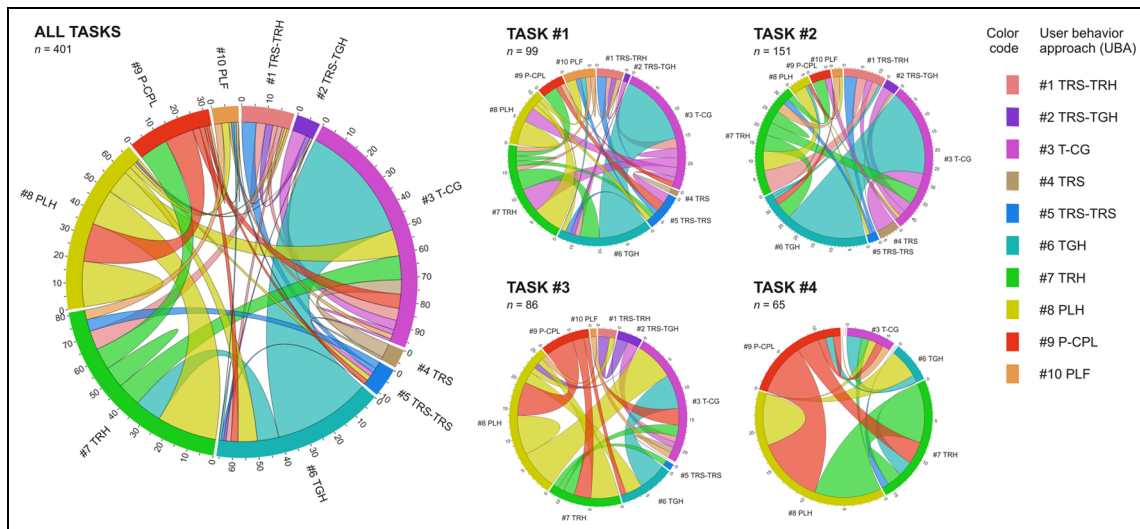


Figure 10. Changes in user behavior approaches (UBAs) summed for all tasks and by task.

the aggregated number of usages is analyzed. Three glyph-related approaches feature high share in the aggregated number of usages, especially in T#1 (K) and T#2 (L) when most segments are assigned to UBA #3. Nevertheless, when normalized using absolute aggregated duration, the absolute aggregated duration per one usage is not distinctive for UBA #3 (U).

Task execution analysis

When executing the tasks, in most cases (36.8%, ±1%, 95% CI) participants demonstrate two different UBAs (Table 4). This strategy is predominant especially in T#3 and T#4. Moreover, maximally six different UBAs are used (only one participant does so).

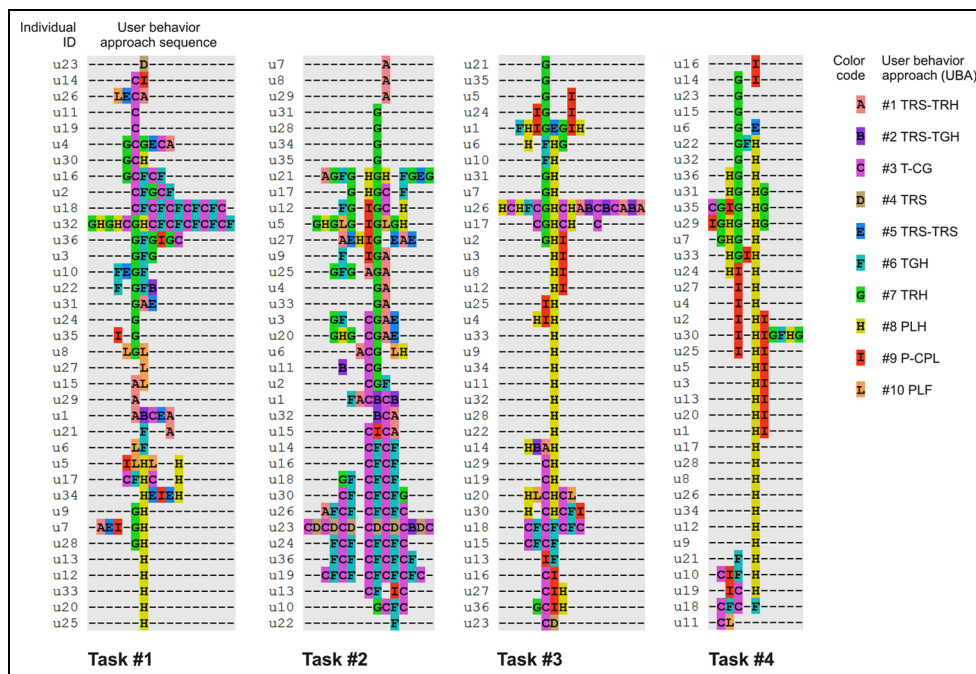
Regarding the total number of UBAs demonstrated during the execution of the tasks (Figure 9), in T#1 and T#2, most participants use only one UBA (no UBA succession); the latter task also features the highest variety in succession numbers. In contrast, in T#3 and T#4, most individuals change UBA once.

A circular layout (Figure 10) is used to gain an insight into the changes in UBAs. In most cases (50 of all 401 successions—12%), after applying UBA #3 (commenting on the polyline glyphs with scrolling through the table), participants hover over polyline glyphs in the table (approach #6); reverse succession is also frequent, as participants behave in this way in 41 of all 401 cases (10%). These two changes (#3→#6 and vice versa) dominate especially in T#1 and T#2. In T#3, the most numerous successions are #3 → #8

Table 4. Number of different UBAs demonstrated by individuals by task.

Number of different UBAs demonstrated by individuals in task execution	% of Participants (n = 36) by task (95% confidence intervals given in brackets)				
	T#1	T#2	T#3	T#4	Sum of all cases (n = 144)
1	30.6 (± 15.1)	22.2 (± 13.6)	25 (± 14.1)	27.8 (± 14.6)	26.4 (± 0.8)
2	27.8 (± 14.6)	19.4 (± 12.9)	47.2 (± 16.3)	52.8 (± 16.3)	36.8 (± 1)
3	25 (± 14.1)	30.6 (± 15.1)	16.7 (± 12.2)	11.1 (± 10.3)	20.8 (± 0.7)
4	13.9 (± 11.3)	13.9 (± 11.3)	5.6 (± 7.5)	8.3 (± 9)	10.4 (± 0.5)
5	2.8 (± 5.4)	13.9 (± 11.3)	2.8 (± 5.4)	0	4.9 (± 0.4)
6	0	0	2.8 (± 5.4)	0	0.7 (± 0.1)

UBA: user behavior approach.

**Figure 11.** Aligned sequences of user behavior approaches (UBAs) by task.

and #8 \rightarrow #3, whereas in T#4 they are #8 \rightarrow #7 and #8 \rightarrow #9.

Furthermore, we employ sequence alignment analysis (SAA)³⁸ to compare the UBA sequences across individuals. We use Clustal X³⁹ to summarize and compare individual UBAs systematically. We focus on the order of UBAs, but not on the durations of their application. Figure 11 shows the aligned UBA sequences grouped by task, where one row represents a UBA sequence for one individual, read from left to right. Although approaches #7 (color code G) or #8 (H) separately or their successions are frequent in all tasks, they are particularly numerous in T#3 and T#4. In the latter two tasks, the successions of UBAs #8 (H) and #9 (I) are frequent too. Regarding the

glyph-related UBAs, certain individuals exhibit similar aligned sequence patterns with the combination of approaches #3 (C) and #6 (F); a series of such successions are visible particularly in T#2 and to a minor extent in T#1.

Comment analysis

Although the participants are requested to comment aloud on the questions posed, not all of them do so, and some of the comments of those who do follow our request are either shallow or inaccurate. Examples of the comments are provided in Table 5. The tasks mainly receive accurate comments (Figure 12) and 20 of 36 individuals comment on all tasks accurately

Table 5. Outlines of the tasks and examples of the related comments.

#	Task (outline)	Example comments
1	How do the multivariate signatures of municipalities vary depending on their decreasing rates on the "Non-immigrants (% of population)" variable?	Participant #1: "I can see that is, which are on the top (in the table) are pretty similar in the start, but afterwards it gets a bit different ..." Participant #14: "It is just maybe the top 15 or so [in the table] seem quite similar and it starts changing gradually ..."
2	How do the multivariate signatures of municipalities vary depending on their decreasing similarity to the municipality of Sola?	Participant #1: "As the last time, they are pretty similar in the start. [scrolling down the table] I could tell that there are some differences ..." Participant #10: "... between ones looking most like Sola, they have the same curve basically in the beginning and ... if we see lower down, see with the different curves ..."
3	Try to identify distinctive cases within the multivariate signatures of municipalities that score above 6 regarding "Doctors (per 1000 persons)"	Participant #1: "They are pretty much similar around here [pointing at some polylines in the PCP] but, the biggest variation is at the nonimmigrants So we can see at here they're pretty similar and, but we got some, but, like Lørenskog, Lærdal, and Førde, are, aren't similar to the rest ..." Participant #4: "So, there are some similarities in terms of 'Doctors' and some other parallels; however, there are some <i>kommuner</i> [municipalities] that are being abstract as Oslo, Lærdal, Lørenskog, and Førde at some parallels ..."
4	Try to identify distinctive cases within the multivariate signatures of 20 municipalities the most similar to the municipality of Hole	Individual #13: "... the data pattern is similar except for two outliers in 'Working in electricity and transportation', waste and communication': Fedje and Kvitsøy. They are values relatively higher than the normal range which is between about 5 and 15. ... All the other variables, especially 'Median household income after tax' are likely to have the same values or scores ..." Individual #21: "Well, here in, I guess Fedje, that's interesting, and also, here in Kvitsøy, there are different [pointing at two polylines in the PCP] ... Otherwise, it is almost the same. Here in 'Median household income' they are very similar ..."

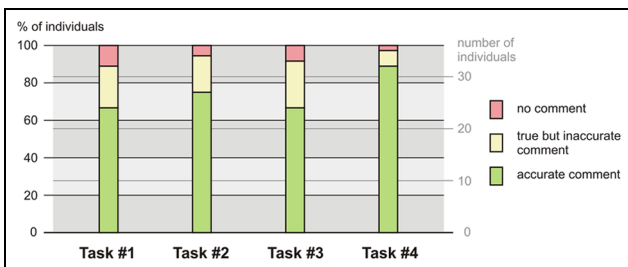


Figure 12. Comment accuracy by task.

(Figure 13). T#1 and T#3 either are not commented on by most participants or result in most inaccurate comments. T#4 results in most accurate comments: 32 of 36 individuals (89%) comment on it in a precise way.

Regarding the attitudes toward the polyline glyphs, only two individuals comment on the glyphs directly, both when solving T#2:

AQ11

Individual #13: OK. So, we are using the glyphs?

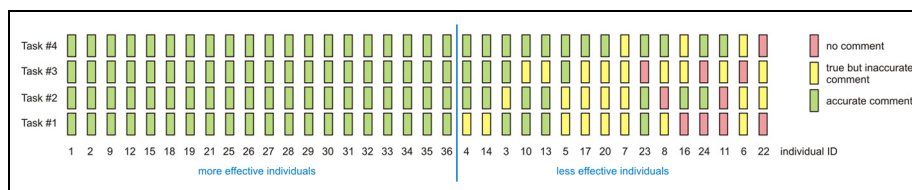


Figure 13. Comment accuracy by individual.

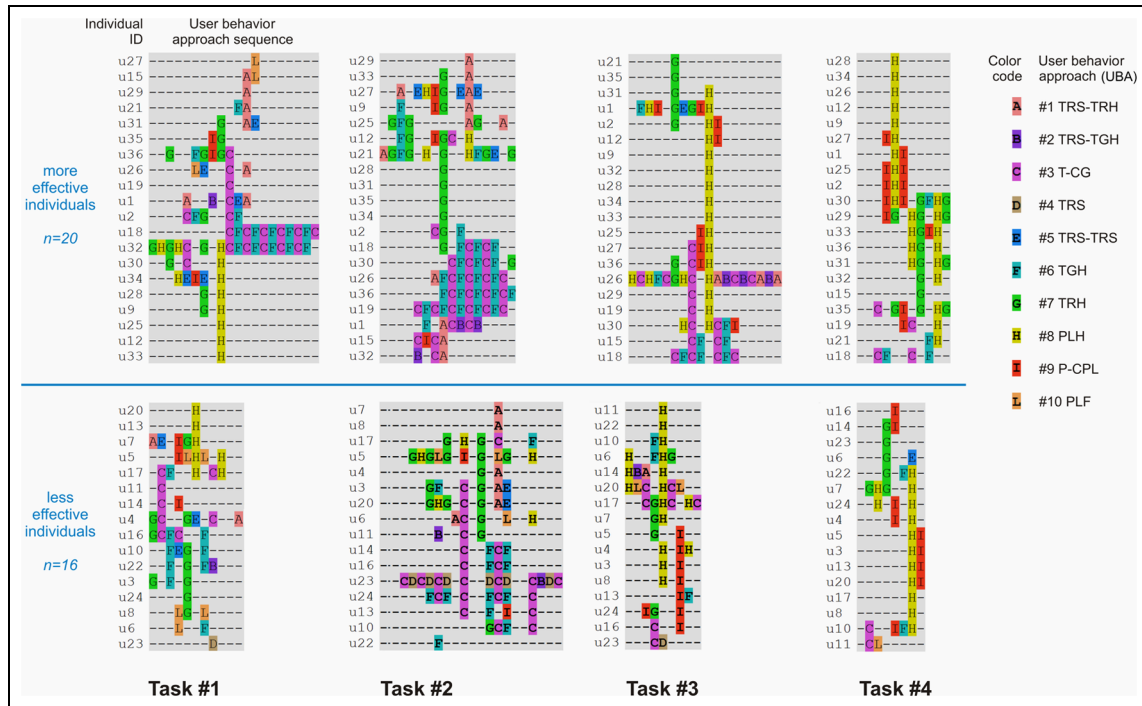


Figure 14. Aligned sequences of user behavior approaches of more and less effective individuals by task.

Facilitator: For instance, it is up to you. You can use these or other one.

Individual #13: I think that glyphs are easier to use.

Individual (#36): I find it easier to ... use the glyphs to see a difference instead of scrolling through them and looking at them that way [pointing to the PCP].

Task execution analysis: more effective versus less effective individuals

More effective individuals are differentiated from less effective ones (Figure 13), with the assumption that those who are more effective are simply those who accurately comment on all tasks. Subsequently, an independent-samples t-test is conducted for each task to compare the total task execution durations for both more and less effective individuals, but the differences are not statistically significant. Also, t-tests are conducted to compare the aggregated relative durations of glyph-related UBAs (approaches #2, #3, and #6) for more and less effective individuals. The differences are insignificant in this case too. A deeper insight into aligned sequences (Figure 14) of UBAs of more and less effective individuals reveals that more UBA successions of UBA #3 and #6 (CF color codes) in T#2, and #8 with #7 (HG) in T#4, are observed for more effective individuals.

Questionnaire outcomes

The layout and the functionality of the visualization tool tested in the empirical study gather mostly positive attitudes. A majority of the participants strongly agree or agree that the design of the tool is successful and the tool is easy to use (see statements S#7 and S#8 in Figure 15, respectively). A majority of the participants also strongly disagree or disagree that it is necessary to learn a lot before starting to use the tool (S#9), and that a tutorial or the support of a technical person is needed to be able to use it (S#10). Nevertheless, the latter statement receives the most diverse responses and the most “I don’t know” responses. Regarding the PCP and the polyline glyphs, most individuals either strongly agree or agree that the use of the PCP (S#1) and the glyphs (S#3) as well as the use of these two techniques in combination (S#5) are easy to understand. They also mostly strongly agree or agree that the use of these three solutions to enable users to investigate data (S#2, S#4, and S#6) is successful. Only a few individuals either strongly disagree or disagree with S#1–6. Although the use of glyphs (S#3, S#4) generates relatively more “I don’t know” responses, the combination of glyphs with the PCP is claimed by most individuals as easy to understand (#5) and successful (#6) when it comes to enabling users to investigate data.

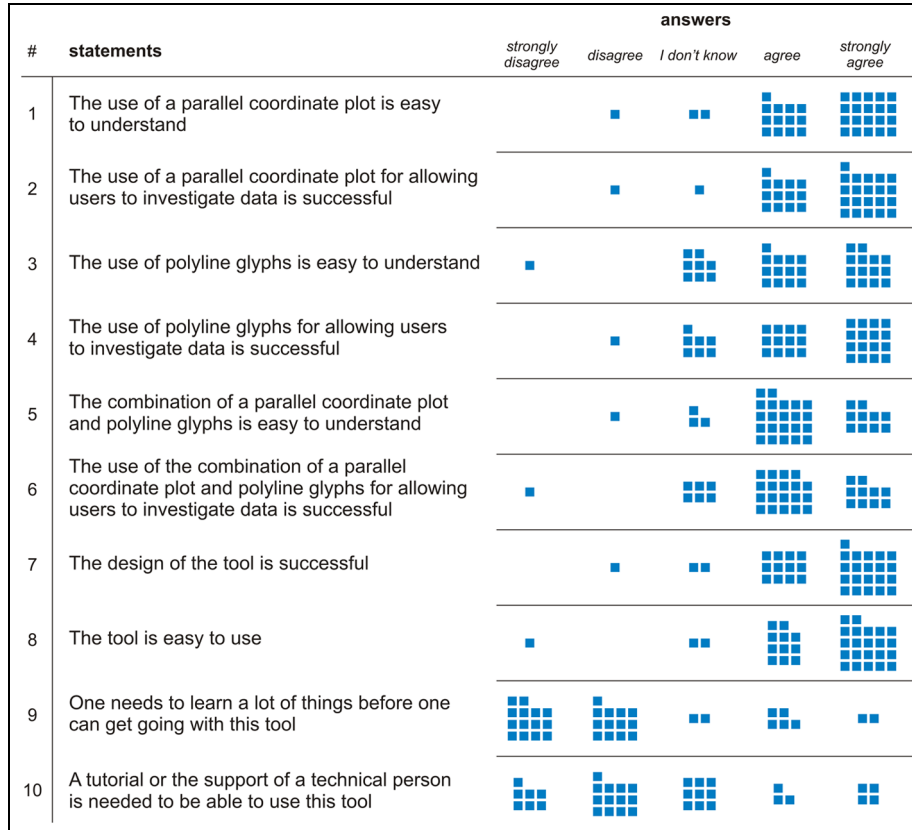


Figure 15. Individuals' attitudes toward the visualization tool used in the empirical study.

Results and discussion

Individuals who participate in our empirical study use a PCP dynamically linked to polyline glyphs (Figure 1) and solve four tasks (Table 1). In the tasks, they are requested to investigate 10-dimensional signatures that show selected socioeconomic indicators. The general aim of the tasks is either to examine how the signatures vary (T#1, T#2) or to identify municipalities featuring distinctive signatures (T#3, T#4). The individuals are expected to investigate the signatures of all 428 Norwegian municipalities in T#1 and T#2, or of their subset in T#3 (32 municipalities) and T#4 (21). Although the participants are neither suggested nor encouraged to use the glyphs to investigate the signatures, most of them freely chose to do so. The participants use the glyphs in all four tasks and only 7 of 36 participants do not use the glyphs at all in any task. Three of the 10 differentiated UBAs (Table 2) concern the glyphs, and of those 3, the most commonly demonstrated is commenting on glyphs, typically along with scrolling through the glyph display (UBA #3) or hovering over the glyphs (#6): 69.4% ($\pm 15.1\%$, 95% CI) and 61.1% ($\pm 15.9\%$, 95% CI), respectively, of the individuals do this at least once in any task, mainly in

T#1 and T#2 when investigating all multivariate signatures is to be done with visual clutter occurring in the PCP. The combination of UBAs #3 and #6 is also common, as is evident from the circular layout (Figure 10). Such behavior is frequent especially in T#2, when the participants investigate all 428 signatures that are sorted in descending order according to their similarity to the one for Sola municipality. In T#2, 38.9% ($\pm 15.9\%$, 95% CI) of the participants demonstrate the combination of UBA #3 and UBA #6 at least once (Table 3). Moreover, T#2 is also the task in which the glyphs are used at least once by most users (69.4%, $\pm 15.1\%$, 95% CI). If one considers the absolute and relative durations of aggregated UBAs and the aggregated number of usages (Figure 8), the tasks in which glyphs are most commonly used are again T#1 and T#2, in particular. Regarding the latter task, the participants demonstrate UBA #3 in 23% of the total aggregated durations (Figure 8, G) and this is the second highest score after the most common UBA #7 (27%). In turn, UBA #3 has the highest score in T#2, if one considers the total aggregated number of usages: 50 of 179 (i.e. 28%; Figure 8, L).

The task execution analysis confirms the above findings. The UBA successions for which people

comment on the glyphs and hover over them (UBAs #3 and #6) are the most frequent in T#2. Then, 12 of 36 individuals do so at least once (see Figure 11). In T#3 and T#4, the polyline glyphs are most seldom applied. Furthermore, in these two tasks, the executions are the most consistent regarding both the number of different UBAs demonstrated by individuals (mostly one or two, see Table 4) and the number of changes in UBAs (mostly one or no change, see Figure 9). Typical for T#3 and T#4 is also UBA #8—hovering over polylines in the PCP. With regard to more and less effective individuals (Figure 13), there are no distinguishable characteristics in the differences in their aligned sequences of UBAs (Figure 14). The only difference is that more successions of UBA #3 with #6 in T#2 and UBA #8 with #7 in T#4 are observed for more effective individuals. In other words, in T#2 more accurate comments are given by those who scroll through or hover over the glyph display, and in T#4 by those who hover over polylines in the PCP and rows in the table.

Polyline glyphs encode data in a way similar to parallel coordinates, except that they show multivariate signatures as juxtaposed with each other instead of as superimposed on each other. Moreover, these tiny glyphs encode data in a simplified way (without axes and labels) and can be easily integrated with parallel coordinates. Although in such combination the same data are shown twice, our investigation reveals that for specific tasks this has advantages over the use of parallel coordinates alone. We discover that when a PCP is dynamically linked to polyline glyphs and individuals are expected to use such a tool to gain insight into multivariate signatures of data items, many decide to use polyline glyphs instead of parallel coordinates. Using polyline glyphs are especially favored when individuals are requested to investigate multivariate signatures to see how they vary depending on their similarity to a selected signature (T#2 in our empirical study). Then, individuals tend to scroll through a glyph display with sorted polyline glyphs and investigate these “frozen” thumbnails instead of hovering over the polylines in parallel coordinates. Apart from using the glyphs to investigate how the signatures vary, the participants also report using them to identify distinctive cases, as they can easily see them within the glyphs. This finding confirms the common claim that if multivariate signatures are positioned within the eyespan and it is feasible to make comparisons at a glance-uninterrupted visual reasoning, the similarities and differences among glyphs are likely to be identified more efficiently.^{20,25,26} This can be backed up by the statement given by one of our individuals: “I find it easier ... to use the glyphs to see a difference.”

AQ13

Surprisingly, the individuals do not tend to use the polyline glyphs if the PCP to which the glyphs are dynamically linked is not visually cluttered. If only a dozen polylines are shown in a PCP, as when solving T#4, individuals avoid using polyline glyphs to identify distinctive polylines and only use the PCP.

Conclusion

The main benefit to be gained from using a PCP dynamically linked to polyline glyphs over the use of a PCP alone is that if a PCP and polyline glyphs show hundreds of multivariate data entities and users try to gain insights into the multivariate signatures of those entities, polyline glyphs are especially favorable if investigating the signatures according to their similarity to selected one is to be done. If the number of data entities is reduced and visual clutter does not impede the differentiation of particular signatures in the PCP, users might disregard polyline glyphs.

Linking a PCP to polyline glyphs augments insights into multivariate data provided through a PCP alone by facilitating the investigation of multivariate data items without their preliminary filtering and thus reducing their number. Visual analysis of data is not easily achievable in a PCP alone if the data consist of hundreds of multivariate items. We therefore claim that a PCP dynamically linked to polyline glyphs supports the solving of user tasks in which initially all data items are to be investigated briefly, even if visual clutter occurs in the PCP, and thereafter, only selected items are to be analyzed in detail.

Apart from the use of polyline glyphs together with PCPs, they can also serve as an independent component. However, in such case, the effectiveness of polyline glyphs needs to be further tested.

Funding

The research presented in this article contributes to the ClimRes project (Climate change and natural hazards: the geography of community resilience in Norway), which is funded by the Research Council of Norway.

References

1. Inselberg A. The plane with parallel coordinates. *Visual Comput* 1985; 1(4): 69–91.
2. Ellis G and Dix A. A taxonomy of clutter reduction for information visualization. *IEEE T Vis Comput Gr* 2007; 13(6): 1216–1223.
3. Heinrich J and Weiskopf D. State of the art of parallel coordinates. In: Sbert M and Szirmay-Kalos L (eds) *Eurographics—state of the art reports*. Geneva: The Eurographics Association, 2013, pp. 95–116.

AQ14

4. Opach T and Rød JK. Cartographic visualization of vulnerability to natural disasters. *Cartographica* 2013; 48(2): 113–125.
5. Tufte ER. *Beautiful evidence*. 1st ed. Cheshire, CT: Graphics Press, 2006.
6. Ward MO. A taxonomy of glyph placement strategies for multidimensional data visualization. *Inform Visual* 2002; 1(3–4): 194–210.
7. Wiréhn L, Opach T and Neset TS. Assessing agricultural vulnerability to climate change in the Nordic countries—an interactive geovisualization approach. *J Environ Plann Manag* 2016. 60: 115–134.
8. Edsall RM. Design and usability of an enhanced geographic information system for exploration of multivariate health statistics. *Prof Geogr* 2003; 55(2): 146–160.
9. Edsall RM. The parallel coordinate plot in action: design and use for geographic visualization. *Comput Stat Data An* 2003; 43(4): 605–619.
10. Andrienko N, Andrienko G, Voss H, et al. Testing the usability of interactive maps in CommonGIS. *Cartogr Geogr Inf Sc* 2002; 29(4): 325–342.
11. Robinson AC, Chen J, Lengerich EJ, et al. Combining usability techniques to design geovisualization tools for epidemiology. *Cartogr Geogr Inf Sc* 2005; 32(4): 243–255.
12. Theus M and Urbanek S. *Interactive graphics for data analysis: principles and examples*. Boca Raton, FL: Chapman & Hall/CRC Press, 2008.
13. Fotheringham AS, Brunsdon C and Charlton M. *Quantitative geography: perspectives on spatial data analysis*. 2nd ed. London: SAGE, 2004.
14. Dorling D. The visualization of local urban change across Britain. *Environ Plann B* 1995; 22(3): 269–290.
15. Chernoff H. The use of faces to represent points in k-dimensional space graphically. *J Am Stat Assoc* 1973; 68(342): 361–368.
16. Hardisty F and Robinson AC. The GeoViz Toolkit: using component-oriented coordination methods for geographic visualization and analysis. *Int J Geogr Inf Sci* 2011; 25(2): 191–210.
17. Slocum TA, McMaster RB, Kessler FC, et al. *Thematic cartography and geographic visualization*. 3rd ed. Upper Saddle River, NJ: Pearson Education, 2010.
18. Gribov A, Unwin A and Hofmann H. About glyphs and small multiples: Gauguin and the Expo. *Stat Comp Stat Graphics Newsl* 2006; 17(2): 18–22.
19. Ünlü A and Malik WA. Interactive glyph graphics of multivariate data in psychometrics. The software package Gauguin. *Methodology: Eur* 2011; 7(4): 134–144.
20. Ward MO. Multivariate data glyphs: principles and practice. In: Chen C-H, Härdle W and Unwin AR (eds) *Handbook of data visualization*. Berlin: Springer, 2008, pp. 179–198.
21. Andrienko G, Andrienko N and Jankowski P. Building spatial decision support tools for individuals and groups. *J Decis Syst* 2003; 12(2): 193–208.
22. Takatsuka M and Gahegan M. GeoVISTA studio: a codeless visual programming environment for geoscientific data analysis and visualization. *Comput Geosci* 2002; 28(10): 1131–1144.
23. Ho QV, Lundblad P, Åström T, et al. A web-enabled visualization toolkit for geovisual analytics. *Inform Visual* 2012; 11(1): 22–42.
24. Chung DHS, Legg PA, Parry ML, et al. Glyph sorting: interactive visualization for multi-dimensional data. *Inform Visual* 2015; 14(1): 76–90.
25. Tufte ER. *Envisioning information*. 1st ed. Cheshire, CT: Graphics Press, 1990, p. 67.
26. Klippel A, Hardisty F, Li R, et al. Colour-enhanced star plot glyphs: can salient shape characteristics be overcome? *Cartographica* 2009; 44(3): 217–231.
27. Ropinski T, Oeltze S and Preim B. Survey of glyph-based visualization techniques for spatial multivariate medical data. *Comput Graph* 2011; 35(2): 392–401.
28. Borgo R, Kehrler J, Chung DHS, et al. Glyph-based visualization: foundations, design guidelines, techniques and applications. In: Sbert M and Szirmay-Kalos L (eds) *Eurographics—state of the art reports*. Geneva: The Eurographics Association, 2013, pp.39–63.
29. Turkay C, Slingsby A, Hauser H, et al. Attribute signatures: dynamic visual summaries for analyzing multivariate geographical data. *IEEE T Vis Comput Gr* 2014; 20(12): 2033–2042.
30. Frumination. Spark it up, http://www.frumin.net/ation/2009/05/spark_it_up.html (2009, accessed 28 August 2016).
31. Bertin J. *Sémiologie graphique: les diagrammes, les réseaux, les cartes*. Paris: Gauthier-Villars, 1967.
32. Baikousi E, Rogkakos G and Vassiliadis P. Similarity measures for multidimensional data. In: *Proceedings of the IEEE 27th international conference on data engineering (ICDE)*, Hannover, 11–16 April 2011, pp. 171–182. New York: IEEE.
33. Ashby FG and Ennis DM. Similarity measures. *Scholarpedia* 2014; 2(12): 4116.
34. Boyd JP. Information distance for discrete structures. In: Shepard RN, Romney AK and Nerlove SB (eds) *Multidimensional scaling: theory and applications in the behavioral sciences*, vol. 1. New York: Seminar Press, 1972, pp. 213–223.
35. Arabie P. Was euclid an unnecessarily sophisticated psychologist? *Psychometrika* 1991; 56: 567–587.
36. Rød JK, Opach T and Neset T-S. Three core activities toward a relevant integrated vulnerability assessment: validate, visualize, and negotiate. *J Risk Res* 2015; 18(7): 877–895.
37. Mitter H, Kirchner M, Schmid E, et al. The participation of agricultural stakeholders in assessing regional vulnerability of cropland to soil water erosion in Austria. *Reg Environ Change* 2014; 14(1): 385–400.
38. Wilson C. Activity pattern analysis by means of sequence alignment methods. *Environ Plann A* 1998; 30: 1017–1038.
39. Jeanmougin F, Thompson JD, Gouy M, et al. Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 1998; 23(10): 403–405.