A Study on Visual Representations for Active Plant Wall Data Analysis

Kahin Akram Hassan *, Yu Liu, Lonni Besançon, Jimmy Johansson and Niklas Rönnberg

Department of Science and Technology, Linköping University, Campus Norrköping, 602 21 Norrköping, Sweden; yu.a.liu@liu.se (Y.L.); lonni.besancon@liu.se (L.B.); jimmy.johansson@liu.se (J.J.); niklas.ronnberg@liu.se (N.R.)
* Correspondence: kahin.akram.hassan@gmail.com

Received: 15 April 2019; Accepted: 15 May 2019; Published: 21 May 2019

Abstract: The indoor climate is closely related to human health, well-being, and comfort. Thus, an understanding of the indoor climate is vital. One way to improve the indoor climates is to place an aesthetically pleasing active plant wall in the environment. By collecting data using sensors placed in and around the plant wall both the indoor climate and the status of the plant wall can be monitored and analyzed. This manuscript presents a user study with domain experts in this field with a focus on the representation of such data. The experts explored this data with a Line graph, a Horizon graph, and a Stacked area graph to better understand the status of the active plant wall and the indoor climate. Qualitative measures were collected with Think-aloud protocol and semi-structured interviews. The study resulted in four categories of analysis tasks: Overview, Detail, Perception, and Complexity. The Line graph was found to be preferred for use in providing an overview, and the Horizon graph for detailed analysis, revealing patterns and showing discernible trends, while the Stacked area graph was generally not preferred. Based on these findings, directions for future research are discussed and formulated. The results and future directions of this research can facilitate the analysis of multivariate temporal data, both for domain users and visualization researchers.

Keywords: visualization; qualitative evaluation; temporal multivariate data; active plant walls

1. Introduction

An indoor climate is comprised of the air quality, temperature, humidity, carbon dioxide levels, and many more parameters. The quality of the indoor climate can be of great significance to human health, well-being and comfort. One way to improve such climates is by placing an aesthetically pleasing active plant wall (Figure 1) in the environment, which can contribute to the climate via evaporation, air purification and water retention [1]. An active plant wall is integrated with vegetation, soil and irrigation, a fan to accelerate ventilation, and drainage systems. These systems may be vertically supported and can be internally or externally attached to a wall but can also be designed as a standalone product. Active plant walls can also effectively reduce main pollutants, such as particulate matter and volatile organic compounds (VOC), and stabilize carbon dioxide concentrations to a healthy level [1]. Since the quality of the indoor climate is difficult to notice, Internet of Things (IoT) solutions could be used to discover changes. By placing sensors in and around the plant wall, multivariate data can be collected over time. Such collected data is large as it contains many variables (one for each sensor). The data typically contains 10–20 variables and is usually measured on a time interval between 5–15 min over several weeks, months or even years.

This collected data comes with many challenges and analysis is one of many examples. Understanding the effectiveness of these plant walls is of importance to the domain experts, and Vertical Plant System AB (VPS) is one example company in need for such knowledge. With no effective way to monitor the indoor climate nor the systems, experts have to visit each and every one for control
and maintenance. This is costly for the experts both in terms of time and workload. Therefore, being able to effectively monitor one or several systems located at different locations will result in increased work productivity and lead to the experts being more time-efficient. To achieve this, a quick overview of the data is required. This data can then be visualized and remotely analyzed by the domain experts to better understand and optimize the systems as well as the environment climate. To the best of the authors’ knowledge, this is a first attempt to deal with the visualization of such data for improving the health of an active plant wall as well as the indoor climate.

Figure 1. An active plant wall is aesthetic and visually appealing to the eye. They make a great impact on decorating urban environments made out of concrete and bricks. To keep them healthy and functioning, monitoring and maintenance is necessary. This green wall is installed in the kitchen of an IT consulting company. (Photo courtesy of Vertical Plant System AB).

The aim of the work presented in this paper is to facilitate domain experts’ understanding of the visual representations to make informed decisions and optimize the system efficiency, but also introduce common visual representations to the domain. This was done through a qualitative study with domain experts. The approach taken was to present experts with several visual representations and evaluate the representations’ efficiency in helping them to understand the data, and ultimately use this knowledge to better understand the behaviour of the system.

From the visualization literature, the following three visual representations were chosen: a Line graph [2], a Stacked area graph [3,4] and an Horizon graph [5,6]. The reason for choosing these representations was to present the users with a variety of representations, designed for different purposes, while all having the ability to communicate temporal multivariate data and provide an overview. The Line graph is commonly used in social media, web pages, books, newspapers etc. and is familiar to a majority of laypeople [7]. The Stacked area graph is generally aesthetically pleasing [8,9] and therefore often considered by the visualization community. It is also a space-efficient representation. The Horizon graph is not commonly known by laypeople and is sparsely used in the visualization community [7,10–15], however, this representation is also space-efficient for presenting large data. This gives a range from “familiar” to “unfamiliar” visual representation to be evaluated. Besides evaluating users’ insight by using these representations, information on the relationship between familiarity, complexity and insight is also provided. This can give the reader additional information regarding how to design and think around data visualization within the area of the indoor climate environment.

This paper contributes first a qualitative evaluation with domain experts instead of laypeople as is often done in previous studies (see for instance [7]). Based on this, the paper offers a categorization of the identified user responses linked to analysis tasks and discuss them in comparison to previous findings. Finally, directions for future research within similar application domains are discussed.

2. System Setup

A Vertical plant wall (Figure 1), consists of two parts: a physical plant wall with local units, and a cloud server part (Figure 2). In the following subsections, both parts are quickly described.

2.1. Cloud Server

The core concept of the IoT Hub service is device twins which are JSON objects that sit in the centre of the IoT Hub and record device properties. Local device properties are synchronized to the
properties recorded in a device twin to realize remote management through the cloud infrastructure. By updating device properties, administrators can remotely modify and control the watering, lighting and ventilation functions of a plant wall system. The sensor readings in the local unit are compressed into messages and periodically transmitted to the IoT Hub from the microprocessor through a WiFi protocol. A function application is created to record sensor data into an SQL database (Figure 2). Upon reception of a new message in the IoT Hub, the function application executes the insertion operation.

2.2. Local Hardware

The local hardware system periodically measures, collects and stores data in the cloud via local units including a microcontroller, a microprocessor with a WiFi module, multiple sensors and actuators (Figure 2). An ultrasonic sensor, on top of the water tank, monitors water level changes. Actuators, i.e., the pump, the lamp and the fan, are autonomously controlled by the microprocessor according to predefined time schedules that are locally stored in device properties. Relevant watering, lighting and ventilation tasks are also triggered periodically.

2.3. Collected Data

For this study, data measured by sensors in a controlled lab environment over three weeks was collected (Figure 2, right). These sensors usually generate large volumes of data and understanding such data is a complex task [16]. Several environmental parameters, (e.g., air temperature, humidity, carbon dioxide) are continuously being measured and stored in the database. For this study, twelve variables were measured: humidity, air temperature, carbon mono- and dioxide level, water level (tank), visible light, infrared light, ultraviolet light, pump period, fan speed, ammonia level, and nitrogen dioxide level.

![Figure 2. (Left) The system is split into two parts; cloud server part and monitoring (Cloud), and local hardware and control part (Local). An access point in between connects the two parts. (Right) The lab environment setup with a vertical plant wall. Water tank at the bottom, lights hanging from the ceiling. Sensors are placed around the plant. (Photo courtesy of Thor Balkhed).](image)

3. Visual Representations

Many different techniques have been developed for temporal multivariate data visualization [17]. However, three techniques, namely Line graph, Stacked area graph and Horizon graph (Figure 3), are investigated in this work. The following sections details why these three visual representations are of particular interest for the active plant wall data and the proposed study.
3.1. Line Graph

It is one of the most common time-series representations to display quantitative values over a certain time period, especially to laypeople [7]. Rendering several variables in a shared Line graph could introduce visual clutter and difficulties finding correlations between them. Javed et al. [7] concluded that Line graphs and other similar visual representations are more efficient for comparisons tasks over smaller visual spans with little or no overlaps or clutter. Furthermore, small multiples [18] which are split into individual graphs (Figure 4), one for each variable, could reduce the visual clutter. Hence twelve line graphs were implemented in this study, one for each variable, to keep the technique simple (Figure 3a). Although such implementation solves the clutter issue, it could introduce difficulties for comparison between the variables.

3.2. Stacked Area Graph

The Stacked area graph is an extension of the Area graph [3] and also uses shared-space to display multiple variables (see Figure 5). Each variable is represented by a ribbon, stacked on top of the previous variable, with the height of the ribbon representing the value at a given time step. The ribbons are drawn from bottom to top which means that the baseline for the ribbon on top relies on the ribbon below (see Figure 5c). Reordering the variables would change the shape of the representation substantially, which can lead to confusions, and also, with twelve variables the representation becomes a bit challenging (see Figure 3b). This type of graph is often used by practitioners and has been used in several domains [4,8] which are related to the work presented in this paper. Design options for area graphs are also often discussed in the literature (e.g., see [9]).

The Figure 3. Screenshots of the three visual representations used to display the collected indoor climate data. (a) a Line graph using twelve small multiples, (b) a Stacked area graph using twelve stacked ribbons, (c) a Horizon graph using twelve separate stripes.

The Figure 4. Four Line graph representations showing Humidity, Temperature, Carbon Dioxide Level, and Water Level in the tank used in this study.

The Figure 5. Stacked area graph representation showing two of the variables used in this study. (a,b) show Nitrogen Dioxide Level and Humidity, respectively, while (c) shows the two variables stacked on each other. The height at different time steps represent the value of the variable and here a negative correlation over time can be seen.
3.3. Horizon Graph

The Horizon graph is specifically used to display multiple time-series. Each row starts with a standard time-series and a baseline, normally the mean (see Figure 6a). Values above and below this baseline are colored for separation (see Figure 6b) and values below are horizontally mirrored around the baseline (see Figure 6c). Then the time-series is split into \( N \), normally four, bands and wrapped into a single space (see Figure 6d). The idea behind the Horizon graph (see Figure 3c) was introduced by Saito et al. [5] as two-tone pseudo-coloring, and further developed by Reijner [6]. Variations of Horizon graphs have been investigated in the literature for time-series visual representations (see for instance [11,19–22]).

Most of the literature focuses on improving the Horizon graph and conducting quantitative user studies to understand, most often with time and error measured, which variation is best (see, e.g., [8,11,19,23]). It seems that only a handful of previous research [15] has focused on the use of this visual representation for domain experts in the form of qualitative user study and subjective feedback.

![Figure 6. The construction steps of a Horizon graph. (a) Standard time-series centred around its mean value (baseline). (b) Coloring the graph: red for above and blue for under the baseline. (c) Horizontal mirroring around the baseline. (d) Split and wrap the bands into one single space.](image)

3.4. Interaction

A majority of previous work, considering time-series visualization with these three visual representations, do not implement interaction but rather focuses on the static benefits and limitations of each visual representation and its variations. Moreover, even if an interaction might facilitate exploratory analysis (see for instance [12,24,25]), its use could possibly be detrimental in the case of Stacked area graph because of the additional memory load [8]. Consequently, and to facilitate a fair comparison between each visual representation, an interaction was not implemented in any of the visual representations used in the present study. While these visual representations are usually interactive, the choice to not exactly mimic their use in fully implemented systems to remove possible confounds in a controlled experiment is not uncommon in the fields of Human-Computer Interaction (HCI) and visualization. Examples of such cases can be found in non interactive visual representations [7,23], non-wrist-worn smartwatches [26], or touch-based systems that do not copy their everyday use (e.g., non-hand-held smartphones) [27–29]. The choice not to provide interaction techniques also helped to force the participants to analyze data shown in the visual representation, keeping the focus of the study on the visualization techniques for reaching an understanding of the data by the visual representation alone. As such, it was possible to investigate the domain experts perception of the representations at their simplest, static, form.
4. Method

To explore the experts’ experience of the three chosen visual representations, a qualitative study was conducted. Domain experts and a qualitative approach were selected because exploratory data analysis is a difficult task to assess quantitatively [30]. The qualitative approach, using Think-aloud protocol [31], audio recordings and transcripts from semi-structured interviews and notes, would give a better understanding of the factors that influence the experts’ impression (e.g., [32,33]). In the following subsections, the research method is described in detail.

4.1. Apparatus and Viewing Conditions

A 4K 27 inch HP monitor set at a resolution of 3840 × 2160 pixels was used together with a MacBook Pro. Participants were seated at a distance of approximately 60 cm from the monitor and were invited to adjust the height of the chair in order to provide the optimal viewing position. The application was implemented using Flask1 together with the JavaScript libraries D3.js2 and Chart.js3.

4.2. Stimuli

Changes in the indoor climate take approximately 48 h to affect the plant wall system. Consequently, the collected 60,000 measurements were subdivided into 48-hour intervals. This was then down-sampled to 288 points, showing every 20th measurement in order to not over-clutter the representations.

In each graph, time was placed on the X-axis and the twelve variables were presented. Axis labels were also removed to minimize any confusions and only focus on the development of the variables. Finally, each stimulus display was scaled responsively according to the 4K display.

The colors, taken from ColorBrewer 2.04, were chosen to be equally perceptible and clearly distinguishable. The Line graph was presented in a grid system with each variable as a separate graph and variable title on top (Figure 3a). The Stacked area graph had the variable names on the top, each presented with one color and ordered from bottom to top (Figure 3b). The Horizon graph was presented with the variables in order from top to bottom (Figure 3c), and each variable was displayed as a band with the title far to the left.

4.3. Participants

A domain expert, in this study, is defined as a person who have worked in a domain different from data visualization for a long time and have many years of experience in the field. Because of this expertise, the expert is viewed as a trusted source of valuable information about the topic and the domain [34,35]. However, the question remains: how many domain experts are needed?

Nielsen and Molich [36] found that with six people one can find, on average about 80% of all usability problems. Moreover, Wongsuphasawat and Shneiderman [37] conducted a usability test on eight participants, Isenberg et al. [38] showed that on average between 1–5 participants are used in evaluation research projects, Lundstrom et al. [39] conducted a user study including five orthopaedic surgeons, and Sousa et al. [40] conducted a study on six radio-diagnostics on how virtual reality can assist them in their work. Therefore, based on this knowledge, five domain experts (two female) from VPS, ages 38–53 and median = 48, were recruited for this study. In addition, a low number of participants is sufficient when conducting domain specific qualitative studies, as qualitative research is not concerned with making statistically significant statements and if investigators report qualitative feedback from domain experts [15,39,41–43].

1 http://flask.pocoo.org/.
2 https://d3js.org/.
3 https://www.chartjs.org/.
4 http://colorbrewer2.org/.
Finally, all participants in the present study had in one way or another seen Line graphs before, however, they had not seen or used Horizon graphs or Stacked area graphs. They had normal or corrected-to-normal vision and no prior knowledge of the study structure. No compensation for participating in the study was provided.

4.4. Task

The tasks, inspired by Dimara et al. [44], focused on finding the correlation between variables to better understand how the experts’ reason when making decisions. The participant’s task was to read and answer questions regarding relationships in the data presented with the visual representations. These questions concerned two or three variables in the data, for example, “What is the relationship between Water Level and Humidity?” and “What is the relationship between Humidity, Pump period, and fan speed?”. To answer these questions, participants had to find the correct variables, and understand the patterns and relationships between these variables. By analyzing the data this way, experts are able to draw conclusions and predict the status of the plant system. Two levels of difficulty were used in the study to better mimic a real-world scenario as well as to demand both a low and a high level of analysis of the visual representations. Each participant analyzed the data present in the representations by answering a total of eighteen questions, three less complex and three more complex questions for each visualization, in a counter-balanced order.

4.5. Procedure

Each session started with a basic introduction to the three visual representations. Simplified versions of the representations were printed on paper, and during the introduction, the test administrator described each representation in detail. The introduction was followed by a training session with the dataset and the representations displayed on the computer screen. During the training session, experts had to analyze the data in the representation and answer one question in both difficulty levels (the relationship between two variables, and the relationship between three variables), in order to get familiarized with the test process. During these two initial parts of the test session, experts could ask questions and discuss the visual representations as well as the test procedure with the administrator.

After the training part, participants answered three questions in each difficulty level in each visual representation. The order of the visual representations was counter-balanced between the experts to avoid order effects and the data presented was from different time steps of the complete data set to render new patterns for each question. The experts were instructed to read the question aloud, in order to fully understand and remember the task during the trial [45] and think-aloud protocol [31] was used to collect their feedback. When they felt they had answered the question, the next trial began. After each visual representation, participants had to answer questions about their experience with the representation in a semi-structured interview. Finally, after all three representations, an interview was conducted regarding the general experience of all representations. The test session took approximately fifty minutes to complete.

4.6. Interviews

After completing the trials for each representation, a semi-structured interview followed. In this interview, four different questions were asked to understand how the participants’ experienced the used visual representation. These questions were divided into four areas and covered the understanding of the representation, finding the necessary information to solve the task, whether or not the representation displayed relevant information, and finding relationships in the data with the representation. After finishing all trials for all representations, a more general questionnaire was filled in by the participant and the answers where discussed. The questions examined which visual representation (1) was the easiest to read and interpret, (2) best presented the data and relationships, (3) was the fastest to find information in, and (4) the best to provide help for the participant in their
work tasks. The answers to all interview questions were audio recorded but also written down by the test administrator. A second test administrator also took notes throughout the whole test session. The data collected from all five participants was transcribed and analyzed via categories found, as described in the following section.

5. Results

The results, categories found in this process, were either derived directly from the interview question responses or comments given during the interviews. This way of categorizing the information in the interviews supported the formation of an understanding of the domain experts’ experience and how they would analyze their data for gaining insight about the relationships in the data. As Tweedie [46] states; sometimes there is a need for precision to be able to provide an exact answer, and sometimes there are also situations where less precision is needed. Four categories were found for further analysis:

General overview of the data: This category considers how fast and easily domain experts think that they perceived an overview of the data present in the visual representation.

Details in the data: This category concerns the experienced simplicity with which the experts find details in the presented data in the different representations.

Perception of visual elements: This category considers how the experts perceived the visual elements in the visual representations, such as color schemes used, order and position of variables, as well as the position of the variables labels.

Complexity of the representations: This category covers how the complexity of the visual representations was perceived by the experts, with regards to visual representation, number of variables, and familiarity with the type of visual representation.

The order of the visual representations in the following text is based on the ranking of the appropriateness based on the experts’ answers from the interviews, with the highest ranked representation first.

5.1. General Overview of the Data

An overview could be seen as a qualitative awareness of the data, and that this awareness is acquired swiftly and without cognitive effort [47]. In the present study, an overview is seen as when a user easily obtains an understanding of the scope of the data, of the variables present in the visual representation, and of major relationships within the data. For getting a general overview of the data, the Line graph was perceived as the best, followed by the Horizon graph and then the Stacked area graph.

The Line graph was generally experienced as easy to understand and quick to find information in. Although it is a commonly used representation for visualizing temporal data, three out of five of the domain experts, expressed that they were not specifically used to work with this representation for plant wall data visualization in such a structure. Yet, they found this representation to be an easy way of getting an overview of the data without experiencing a particularly high cognitive load. One of the experts commented that “although I have not worked with this [type of] representation before, it still was quite easy to find relationships [in the data], it felt logical to me”. As the Line graph only displayed one variable in each graph, this might have supported the perception of overview and been experienced as less cluttered (similar to Javed et al. [7]). This was also expressed by one expert: “the representation is easy to understand because it gives a simple overview of the data, and you can easily find relationships between variables as the graphs were separated and not cluttered in one single graph”.

The Horizon graph was new to all the experts but was in general reported to be easy to find relationships with. According to Javed et al. [7], using the unfamiliar Horizon graph was slightly slower than the Line graph when evaluated with laypeople. The representation was experienced to be slightly difficult to comprehend in the beginning but when they learned how to interpret it they could...
perceive an overview of the data and easier find answers to the questions. One expert noted that “there were too many variables on the screen and it required some learning but after a few questions it became easier” which also mirrors the results found by Janetzko et al. [15]. Apart from providing a general overview (see Figure 7), the representation was also found to be useful for revealing global patterns as well as showing temporal changes. A reason for this could be due to the color coding, as one expert said “it was very easy to understand, mostly because of the simple color code. The representation gave a good overview of low and high values”.

![Figure 7](image7.png)

**Figure 7.** An example of an overview provided by Horizon graph. The number of variables is visible, individual developments of the variables over time are discernible, as well as major relationships within the data.

The Stacked area graph was experienced as the most difficult representation for providing an overview, mostly due to its complexity. The experienced learning curve was a bit higher for this representation, however, after working with it for a while the experts understood how to read and compare the variables. One expert stated that “the representation is very messy and difficult, but if you work with it for a while, it might become easier”. The stacking of the variables might have put higher demands on the analysis, which in turn crippled the perception of an overview, as one expert stated “it is very difficult to find relationships between variables because it is very challenging to see which variable that has a higher value”. This result mirrors past claims from Thudt et al. [8], who showed that Stacked area graph have aesthetic qualities but probably should not be considered when efficiency is a criterion. In addition, the number of variables and the coloring of these might also have affected the perception of overview more negatively.

### 5.2. Finding Details in the Data

In this step of the analysis, relationships and trends are analyzed in detail which can be seen as the aggregation of several features of an object, which is a conscious cognitive process that does not occur quickly but require cognitive effort [47]. In the present study finding details is seen as when a user finds, recognizes and can describe detailed relationships in the visual representation (see Figure 8). For finding details in the data, the Horizon graph was perceived as the best, followed by the Line graph and then the Stacked area graph.

![Figure 8](image8.png)

**Figure 8.** A section of the representations showing a detailed development of one variable over a short period of time. From left to right: Line graph, Stacked area graph (dark green ribbon), Horizon graph.

Overall the Horizon graph was experienced as the most useful representation for revealing detailed patterns and trends in the data, as well as supporting the analysis of detailed relationships
between multiple variables (see also [7]). The color scheme indicating negative and positive trends relative to a baseline as well as the space efficiency with closely stacked variables might have supported the detailed analysis. One expert stated that with the Horizon graph “it was easy to see high and low values and compare the variables”. However, even if the learning curve for the Horizon graph might be slightly steeper as compared to the Line graph, it was considered by the experts as an easier representation compared to the Stacked area graph.

Although the Line graph was found to be useful for getting a quick overview of the data, it was not perceived as equally good for revealing detailed relationships. A reason could be the structure of the Line graph, which displayed twelve smaller graphs simultaneously creating distances between variables needed to discern the relationships in the task. One expert expressed “I don’t understand the relationships”, and further suggested that the link between the variables displayed was possibly illusive. Another expert stated, in a similar way, that the distance between the separate graphs made a detailed analysis more difficult, but that it was “easier when the graphs were close to each other”. While this structure reduces clutter and overlapping issues, the distance between each graph forced the experts to keep the development of one or more variables in memory while looking at another variable to determine the relationship between these. This process might have caused a higher cognitive load and increased the risk of misinterpretations.

The Stacked area graph was experienced as the hardest representation to find and understand details with. Although the representation is more compressed and space-efficient than the Line graph, the layout seemed to make the analysis hard since one variable is stacked on top of a lower one. While this could be challenging already with two variables, twelve variables were used in the study which might have caused confusions and incorrect analyses of the variables and the relationships between them. As one expert stated that “sometimes it is confusing to think about how the variables are stacked on top of each other, and still try to find a pattern in the data”. Another expert stated that the “relationship between variables are not clear because the development of the lines in the graph are spiky, which confuses me”. Generally, the representation was perceived as difficult and overwhelming to work with, even if it was better understood by the experts after a considerable longer time compared to the other two representations. One expert said that when “first introduced to the method the concept of area width was difficult to understand, but when worked with it for a while it became easier”.

5.3. Perception of Visual Elements

Three important aspects could be considered when choosing the type of visual representation: the type of data being represented, the type of users, and the type of task being performed. However, it is also important to consider the way in which a user interprets the encoded data [47]. The interpretation of the visual representation is affected by aspects such as the color scheme and the labels used, the structure in regard to the position and the number of the visual elements, and the use of tool-tips and interaction. A visual representation needs to be as intuitive as possible in order to be simple to work with, and to minimize the cognitive effort for the user [47].

The color scheme might be used in a visual representation to, for example, help to distinguish between variables or identify trends in the data, (see Figure 9). The Line graph had a monochrome color scheme which was perceived a simple and easy to understand, while in the Horizon graph positive and negative values are colored in red respectively blue. One expert stated that “Horizon graph has a nice color scheme which made it easy to distinguish between high values and low values”. However, with the Stacked area graph, each variable has an individual color which was perceived as more complex. The number of colors seemed to increase the perceived complexity. A specific color scheme is highly dependent on the underlying data, as well as the context in which it is presented [9]. Therefore, coloring the Stacked area graphs with many layers might be challenging, as colors might become visually distracting and make the representation challenging to read. The selected colors were also experienced as possibly problematic. One expert expressed that “It would make the representation more
intuitive if there was some logic behind the colors”, therefore suggesting a more ecological connection between the actual variables and the colors used, for example using the color blue for the water.

![Color schemes example](image)

**Figure 9.** The color schemes used. From left to right: Line Graph, Stack area graph, and Horizon Graph.

In both the Line graphs and in the Horizon graphs the label for each variable in the visual representation was presented in each individual graph or on each representation. This created a more closely link between label and representation compared to the Stacked area graph where the legend with variable names and color code was presented at the top of the display. One expert commented that “Line graph was intuitive and simple to follow”.

Even if the structure in the Horizon graph was experienced as new and unfamiliar, it was still experienced as easy to understand and the domain experts expressed that the Horizon graph provided more information for detailed analysis compared to the other visual representations. The Line graph was perceived as simple, but the experts expressed that the separation of the variables in different smaller individual graphs was a hinder for careful analysis of the data. The structure of the Horizon graph was, on the other hand, preferred by the experts for detailed analysis of the data, but also for this representation the ordering of variables could obstruct the analysis. In the Stacked area graph, the structure was experienced as most demanding and problematic for analyzing and interpret the data, and relationships in the visual representation.

A common issue for all visual representations was that the domain experts would have preferred some kind of tool-tip and interaction when working with the data analysis. One domain expert expressed after working with the Horizon graph that “I would have preferred selection and ordering of the variables which would have made the analysis deeper”.

There are, in both the Line graph and Horizon graph, some shortcomings in the structure of the representation, in regard to the positioning of variables in the Line graph and the use of bands layered on top of each other in the Horizon graph. However, both were experienced as easier to use than the Stacked area graph. Thudt et al. [8] concluded that the Stack area graph is aesthetically pleasing to the eyes if the colors are chosen correctly but choosing “correct” colors is a difficult task [9]. Furthermore, the color scheme used in the Stacked area graph was designed to make the representation of the variables distinguishable but the number of variables and colors increased the experienced complexity rather than facilitated interpretation of the representation.

The main task for the domain experts was to analyze and investigate the data to find relationships, patterns and trends. In general, two aspects of the visual representations affected this task; the color scheme and the positioning of variables. The aim of the present study was to investigate which of the visual representations, in their simplest and most basic form, would be best to use for active plant wall data, and as a consequence no tool-tips or interactions were implemented. Nonetheless, interaction techniques were asked by domain experts. Tool-tips and interactions such as zooming, filtering, and positioning of variables could considerably facilitate data analysis [12,24,25]. The number of variables could then be reduced, and specific variable selected and positioned to only display the variable in the right order needed for an analysis.
5.4. Complexity of the Representations

A visual representation consisting of many different and connected parts that might be defined as complex. The complexity is further increased as the number of variables increases but also by an increased number of time steps. The perceived complexity might also be affected by the positioning of variables, the selected color scheme, or the familiarity of the visual representation. The experienced complexity most probably influences the perceived difficulty of visual representation, and negatively affect understanding the overview and finding details in the representation. The Line graph was generally perceived as the least complex visual representation, followed by the Horizon graph and then the Stacked area graph.

The number of variables affected the complexity of all representations (see Figure 10). One expert expressed that “there were too many variables on the display”. In the Line graph, the number of variables affected the distances between the graphs on the display, the positioning of variables, and the size of the graphs. In the Horizon graph, it influenced similar aspects as the for the Line graph, however to a less extent and this was found to be especially challenging in the Stacked area graph. As more time steps are presented on the time axis, developments and changes in the variables will be displayed more densely which will create a more noisy and complex representation. This was experienced to be most disadvantageous in the Stacked area graph, as the noisiness of one variable would affect the representation of other variables.

The positioning of the visual elements on the display affected the perceived complexity of the visual representations. In the Line graph, each variable was presented in individual graphs, which facilitated the identification of each variable. One domain expert stated that “I experience this method to be simple to understand, I clearly see the graphs”. In the Horizon graph, the positioning was not particularly commented on in the interviews. As the Horizon graph is designed to be space-efficient [7], the positioning of the variables did seem to affect the complexity. However, the mirroring and overlaid bands in this representation is experienced as somewhat complex, as one expert expressed it “was a simple visual representation, but I got a little bit confused about positive and negative values” suggesting that guidance is the beginning is necessary. The positioning of variables seriously affected the experienced complexity in the Stacked area graph. One expert experienced the Stack area graph as “Extremely difficult, it feels like the variables on top have higher value due to their position”.

The colors used also affected the experienced complexity (see Figure 9). The monochrome color scheme in the Line graph made it less complex as one domain expert stated it “the single color makes it easy”. In the Horizon graph the bands of values overlaid onto each other, and the bands were then color-coded in magnitude as well as in positive and negative values relative the baseline. This color scheme reduced complexity as trends and patterns in the data are more easily recognized. One domain expert expressed that it was simple “to see differences between high and low values”. In the Stacked area graph, each variable was colored with a different color to distinguish between variables but the increased number of colors was experienced as more complex.
How familiar a user is to a visual representation, most probably affects how complex (and difficult) a representation is experienced. None of the domain experts had experience in Information Visualization or worked with the Horizon graph nor the Stacked area graph. However, one expert stated that the Line graph was easy to use, “because I am used to these types of graphs”. However, the experts expressed that the more unfamiliar visual representations could be useful for certain types of tasks, such as deep analysis of the relationships between variables and finding patterns.

6. Discussion

Analysis of time series data from the active plant wall field is important but also difficult. This paper provides a better understanding of the domain experts’ experience of the used visual representations when analyzing data. Remote monitoring and management of plant walls is important for them since they have a wide distribution of these plants and physically visiting each is time consuming and costly. This study resulted in new knowledge about the domain and how data analysis would be performed. Even though the experts know their data, guidance through these visual representations is necessary, which was also shown by Janetzko et al. [15].

While this study has focused on data from the active plant wall field, the results are also transferable to any domain working with time series data (e.g., sports data, biomedical data, weather forecasting, economic forecasting, stocks market data) for analysis, prediction, finding trends and patterns as well as livestream monitoring for alerts.

A summary of which visual representation the expert preferred is presented in Table 1. The fact that the experts found the Line graph easy is not surprising as it is one of the most used visual representation for displaying temporal multivariate data. The experts also found the representation to be easy for finding a general overview of the relationships between the variables, and that the representation was the fastest for finding answers to the questions.

Table 1. The experienced preference for the visual representations in the four categories found in the interviews with the domain experts, the stated preference of the representations. Recommendation is based on all categories and the discussion of these.

<table>
<thead>
<tr>
<th>Visual Representation</th>
<th>Overview</th>
<th>Details</th>
<th>Complexity</th>
<th>Presentation</th>
<th>Preference</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line graph</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Horizon graph</td>
<td>□</td>
<td></td>
<td>□</td>
<td></td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Stacked area graph</td>
<td>□</td>
<td></td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

= Preferred □ = Less preferred □ = Not preferred.

The Horizon graph was perceived as a good representation for perceiving a general overview of the data as well. In addition to this, it was also experienced as providing a deeper understanding of the details, how the variables change over time which is useful when analyzing historical data. Being a space-efficient representation [7], the Horizon graph provided a better comparison alternative between the variables, which could lead to better finding relationships. Although if interaction is not implemented, an intuitive ordering structure is necessary.

While the two above-mentioned representations where experienced as good representations to analyze temporal multivariate data with, the Stacked area graph was not. The experts perceived this representation to be the most complex representation to work with. This is most likely due to the structure of the representation as well as the color scheme, thus mirroring previous findings from Thudt et al. [8] from their study with laypeople. However, with a high number of variables in combination with being able to distinguish between them, this representation ends up with many colors which can be overwhelming for the participants.

While the representation can be cumbersome, experts did slowly learn how to interpret it. The conclusions derived from the transcribed data suggest that a Stacked area graph could be useful
for an analysis of relationships between variables if carefully designed. Grouping variables and using
different shades of one color could be such a design suggestion. This would decrease the number of
variables and also make the representation aesthetically pleasing. However, multiple graphs might
be needed.

In summary, the preferred visual representation was the Line graph. Although if domain experts
had worked more with the three visual representations, their answers might have changed. Therefore,
based on the findings and discussions in the present study, it seems that the Horizon graph can
be recommended for domain experts working with such temporal multivariate data. Indeed, they
(1) provide an overview over trends and relationships in the data, while (2) being able to present details
for further analysis, as well as (3) having a rather simple structure and intuitive way of presenting
data. Finally, Horizon graphs have already been argued in the literature as being space-efficient [7],
which represent another advantage of this visual representation over the other two.

7. Limitations and Future Work

Based on the work in this paper, this section presents several directions of immediate future
research targeting the development and evaluation of the visual representations. The suggested
directions could, hopefully, help users across several different domains, with similar tasks and data.

**Further evaluations** It would be of interest to make evaluations within other, similar, domains in order
to investigate the three representations' efficiency for similar/other tasks. The domain experts
in the present study can all be categorized as laypeople when it comes to using visualization
as an analysis tool for temporal multivariate data, so increasing the information visualization
community's understanding of how novice users act and respond to this type of visual analysis
would certainly be of use.

**Combining representations** The Horizon graph is challenging to understand at first. However, after
a few minutes, users are able to grasp the notion of overall positive or negative trends. Analyzing
details remains, according to presented results, difficult for a substantial time. A possible solution
for this would be to overlay a Line graph on top of the colored bands in the Horizon graph.
This would likely help users in simultaneously interpreting overview, trends and details more
efficiently (see Figure 11).

**Extending representations** The Stacked area graph could be extended with smoothing of details in
variable not focused on. This could reduce the noise and detail in those variables, resulting in
more focus on the interesting variables to better see trends and patterns. However, the selected
variable(s) would have a full resolution, displaying all variations over time for detailed analysis,
(see Figure 12).

**Storytelling** An effective way of explaining complex concepts and engaging the users could be with
animation (showing changes over time in the data). It would be of interest to investigate whether
storytelling, an ordered sequence of steps each containing words, images, visual representations,
video, or any combination thereof, would increase users understanding of the data. This added
dimension could produce certain cognitive effects, however, appropriate use of motions might
empower users in exploratory data analysis [48].

**Application** Knowledge gained from these, or similar, studies could be implemented as an application,
a fully functional dashboard for monitoring of temporal multivariate data. Such dashboard
should be implemented with regards to which visual representation that is most suitable
to provide (1) overview and (2) details, as well as considering (3) the complexity of the
representation and (4) the use of visual elements. In this application, interaction, for example, to
select and position variables, would be both necessary to implement and to evaluate.
Figure 11. A possible example of a combination of a Line graph and a Horizon graph. The Horizon graph reveals patterns, discernible trends, and show how variables develop over time, while the Line graph provides an easier perceptible representation.

Figure 12. An example of how a Stacked area graph could be extended with an interactive slider for setting the detail level of not selected variables.

8. Conclusions

In the present study, three visual representations (Line graph, Stacked area graph, and Horizon graph) were qualitatively evaluated with five domain experts in the field of active plant walls. In their work these experts need to monitor and understand the status of the active plant wall in order to keep the plant wall healthy and maximize the effect on the indoor climate. The analysis of time-series data is important not just for this domain but also for other domains working with similar tasks and data (e.g., economic forecasting, climate forecasting, sales forecasting electroencephalography (EEG) analysis, and stock market analysis). Therefore, the results of the present study are of interest for the visualization community as well as experts working with similar tasks and data. Four categories, from audio recordings and transcripts from semi-structured interviews as well as notes with domain experts regarding their experience of the visual representations, were derived: (1) General overview of the data, (2) Finding details in the data, (3) Perception of visual elements, and (4) Complexity of the representations. Based on the analysis of these categories, the Horizon graph is recommended for presenting such data. This is because the Horizon graph is space-efficient, makes trends and relationships in the data easily discerned, and have the ability to display a multitude of variables on the screen without over-cluttering the representation. Directions for future research include further evaluations in other domains, investigating the combinations of the visual representations used in the present study, evaluating possible extensions of the representations used, and exploring how storytelling could improve the understanding of the representations.

Author Contributions: K.A.H. is the leading author of this work. He designed and conducted the study as well as analyzed the collected data. Furthermore, he also conducted an extensive literature search and wrote major parts of the article. Y.L. implemented all the software code as well as the hardware setup for collecting the data and wrote one section in the article. L.B., J.J., and N.R. contributed in the discussion and provided feedback on the study design. They also contributed to the writing and proof-reading of the article.

Funding: This research received no external funding

Acknowledgments: The authors wish to thank Vertical Plant System AB for information and inspiration, the domain experts for taking part in the present study, and Shaofang Gong for the opportunity of visualizing the active plant wall data. Last, but definitely not least, the authors wish to thank the reviewers as well as the editors for improving our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.
References

1. Liu, Y.; Hassan, K.A.; Karlsson, M.; Weister, O.; Gong, S. Active Plant Wall for Green Indoor Climate Based on Cloud and Internet of Things. IEEE Access 2018, 6, 33631–33644. [CrossRef]

2. Funkhouser, H.G. A Note on a Tenth Century Graph. Osiris 1936, 1, 260–262. [CrossRef]


45. Forrin, N.D.; MacLeod, C.M. This time it’s personal: The memory benefit of hearing oneself. *Memory* 2018, 26, 574–579. [CrossRef] [PubMed]

