Workshop on Visualisation in Environmental Sciences (EnvirVis) (2022)
S. Dutta and K. Feige and K. Rink and D. Zeckzer (Editors)

AtmoVis: Web Based Visualization of Air Quality Data with Interconnected Windows

Benjamin T. Powley[1], Craig Anslow [1] and David J. Pearce [1]

[1] Victoria University of Wellington, New Zealand

Abstract

Air quality has an adverse impact on the health of people living in areas with poor quality air. Hence monitoring is needed to understand the extent of poor air quality. Little work has been done on the effectiveness of visualization techniques for air quality data analysis. Few tools are developed specifically for air quality analysis and many practitioners use general purpose tools, such as spreadsheets or programming. This paper investigates which visualization techniques are most effective in analysing air pollution data. A user study was performed with 20 experienced or expert participants. The participants used a domain specific prototype visualization tool we developed, AtmoVis, to compare spatio-temporal trends among air quality variables using pre-existing visualization techniques. AtmoVis allows experts to explore data without the difficulties of programming, or working with spreadsheets. AtmoVis has a windowed layout that connects 6 different visualizations: heat calendar, line plot, monthly rose, site view, monthly averages, and data comparison. The results of the study demonstrated that the heat calendar, line plot, site view, monthly averages, and monthly rose visualizations were effective for analyzing the air quality through AtmoVis. The line plot and the heat calendar were particularly effective for temporal data analysis. AtmoVis was also effective for accessing air quality visualizations and inferring relationships among air quality variables at different monitoring sites. This research can help inform the design of future domain specific interactive tools for air quality analysis. AtmoVis could be extended to include other datasets in the future.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: Evaluation/methodology—

1. Introduction

Several air pollutants in the atmosphere, including particulate matter, ozone, nitrogen oxides and sulphur dioxide can cause health issues for people [noa18]. For example the health effects from particulate matter include cardiovascular disease [BWS*06] and respiratory mortality [HST*]. In order to understand the effects of air pollution on health and on the environment, a thorough exploration of the air quality data is needed to find relationships and patterns. Understanding the effectiveness of visualizations should assist in choosing the most effective visualizations for analysing the air quality dataset. Our industrial partner NIWA (National Institute of Water and Atmospheric research) [noac, noab] uses a mixture of spreadsheets, R, and ArcGIS [arc] to research, analyse and explore data. The use of general purpose tools for data analysis presents challenges as different people have different skill sets. An interactive web-based domain specific tool for air quality analysis could allow visualizations to be generated without programming skills. Interactive data exploration allows datasets to be explored more rapidly than non-interactive methods, which could help improve the response time of organizations monitoring air quality and enable more effective data analysis [BKSL16].

NIWA provided a dataset of air quality and meteorological data to develop an air quality visualization tool. The data has some issues that can make data analysis challenging. Problems include monitoring sites that use different types of measuring equipment over different time frames, a sparse and incomplete dataset.

As low-cost sensors become more prolific, an increase in the amount of air quality data available motivates the need for more effective visualization techniques. Evaluating visualization techniques for air quality analysis informs the development of visualization tools for analysing air quality.

This research paper makes the following contributions.

- An evaluation was conducted with 20 air quality and GIS expert participants. The task oriented evaluation collected quantitative and qualitative feedback. Statistical evaluation was performed on this data to look for significance.
- AtmoVis, a novel, web based, domain specific visualization tool

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

for environmental analysts to more effectively understand spatio-temporal air pollution datasets.

2. Related Work

We now review several visualization systems paying attention to their interactive design, and how the different visualizations were integrated. Relatively few existing papers explore the effectiveness of visualization systems on air quality data, so in addition to atmospheric visualization systems, a wider range of systems are reviewed. The play button (Fig. 1) allows users to click and see the time frame change for data in the visualizations.

Breadth First A breadth first interface promotes an exploration across many variables in a dataset so that an analyst can form questions and goals during the process of interacting with the data [WMA*16]. Voyager, Trelliscope and Visage allow multiple visualizations to be compared interactively [WMA*16, HGM*13, RLS*96]. Voyager and Trelliscope both display a large number of graphs inside facets [WMA*16, HGM*13]. A usability study on the Voyager system measured the number of variable sets that participants accessed while exploring a dataset. Participants accessed more variables using Voyager compared to a drill down interface [WMA*16] demonstrating that a system like Voyager would assist breadth first analysis.

An objective of the Visage system [RLS*96] was to find methods to combine different visualization tools and applications into a workspace for analyzing information. Visage draws frames around applications, and one of Visage's strengths is the ability to coordinate these applications, manipulate data and transfer information between them through scripts attached to the frames. A multi-window layout similar to Visage could assist with breadth first analysis.

Drag and Drop Drag and drop is an easy way to move data between visualizations, or to configure variables displayed by an interface. Polaris visualizes multi-dimensional databases and uses drag and drop as an interaction technique to create the visual specification of a table [STH08]. Visualizations are created with Lyra by drawing marks on a canvas, adjusting with drag handles, then binding data to them [SH14]. Lyra uses drag and drop for adding fields to the data pipeline and filtering. AtmoVis uses drag and drop to transfer data between visualizations and set axes. Using AtmoVis is fast because visualizations are provided pre-build, and the user adds them to a canvas rather than editing them in a system like Lyra.

Filtering Filtering data allows subsets to be analysed in more detail [Shn96]. The Polaris, Trelliscope, Visage, and Lyra systems employ filters to adjust the data [STH08,HGM*13,RLS*96,SH14]. Trelliscope uses histograms to provide an interactive filter for panels on the visualization [HGM*13]. The panels are filtered both on variables and on cognostics. Data in the Lyra system can be adjusted with filters and other transformations [SH14]. Filtering is a key design aspect of AtmoVis.

Visualization Techniques Heatmaps are a method for visualizing temporal air pollution data [ZWMT16] [LFM16] [DWC*20]. Calendar heatmaps can show pollutant levels as colours for each day of the calendar [ZWMT16] [LFM16]. Heatmaps can also be coloured according to the frequency of air pollution events [DWC*20]. AtmoVis does not interpolate the propagation of air pollutants between different sensors due to the sparse distribution of air quality monitoring sites in New Zealand, and the localized effect of air pollution [noa18]. Heatmaps are also used to compare the variability of temporal climate trends at different locations globally [JBMS09] and have been used to visualize the spatial trends in air pollution [ZLH13,LFM16,BKSL16]. Colour is interpolated over an area containing data measured at discrete locations. Calendar heatmaps were placed on top of a geographical map by Zhou et al. to indicate the location of the temporal air quality measurements on the heatmaps [ZWMT16]. A heatmap of data from a Toxic Release Inventory (TRI) was used to interactively filter the data on other connected temporal visualizations by brushing over a section of the heatmap [BKSL16]. AtmoVis has an interactive heatmap calendar that can change the displayed date time on other connected visualizations by clicking on a day. We did not find other air quality visualization literature which used a heat calendar for time navigation.

Openair is a library for the R programming language [CR12] with visualization and statistical functionality, for example, a calendar heatmap and rose plots. A disadvantage of R with Openair is that programming knowledge is required. AtmoVis incorporates wind rose plots generated with Openair into a graphical user interface so that domain specific experts without programming experience can access them.

A visualization system containing a parallel coordinate plot (PCP) was developed for analysing air pollution in Hong Kong [HWA*07]. The PCP visualized data at a given station. The axis for the wind direction was stylized with an 'S' shape to make it identifiable. The air quality experts provided feedback on the tool identifying that the polar plot was the most useful visualization and that experts preferred to use the polar plot with the PCP, however, details of a formal methodology for analyzing the usability of the system were not provided.

Evaluation The Lyra system was evaluated through a usability study where 15 participants were required to design three graphics [SH14]. The usability study measured task completion, length of time, and Likert scale responses in a post study questionnaire. Qualitative feedback was collected from participants about their experience with the tool, but did not compare Lyra with other tools. Koua, Maceachren, and Kraak [KMK06] evaluated the usability of Self-Organizing Maps (SOM), maps and PCP's for visualizing geospatial data. A usability study was performed with 20 participants. The user test measured the accuracy of the users' response, time taken, responses to visualizations based on questionnaires, interviews and a survey. Study tasks were marked, the same tasks were used for different visualizations, and the results compared performance and efficiency.

Although other visualization tools [DWC*20] have applied web based visualizations to air quality, New Zealand's air quality data are sparsely monitored, so different analysis tools are required. In

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

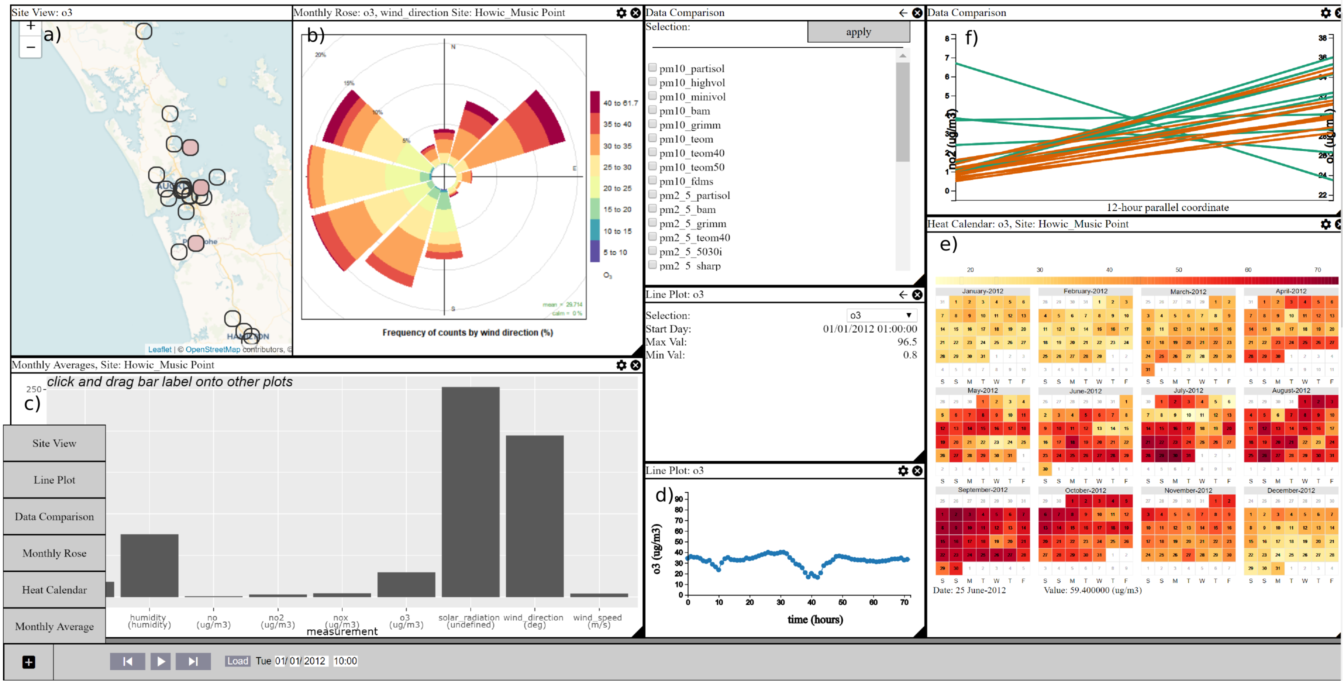

Figure 1: *AtmoVis visualizations: a) Site view, b) Monthly rose, c) Monthly averages, d) Line plot, e) Heat calendar, and f) Data comparison.*

the literature, we were not able to find a visualization tool that combined the necessary selection of visualizations with multiple linked views and a web based interface. We hope that the combination of visualizations with interactivity will improve the experience of analyzing air quality data.

3. AtmoVis

AtmoVis is a domain specific tool for environmental analysts to explore air pollution data in urban environments across different time frames and sites.

3.1. Design

AtmoVis has a breadth first interface to allow a large number of variables among different sites to be explored rather than drilling down into the variables at individual sites. This approach was taken as breadth first interfaces are more effective for exploring a large number of variables in a dataset [WMA*16]. AtmoVis integrates New Zealand air quality data into a web-based front end with the functionality and goals identified based on Personas [PA05]. Three experts were interviewed to create three personas using the programmer's perception of the target audience, an air quality scientist, a data analyst, and a GIS student. Two system goals were developed by analysing persona air quality focused scenarios and goals. Firstly, the system should both allow the discovery of pollutants measured by a site and identify sites measuring a selected pollutant. Secondly, the system should allow the comparison of temporal trends for a pollutant, and the comparison of spatial trends in air quality. The design of AtmoVis is based on other visualization systems for data exploration, for example, Voyager, Trelliscope and Gapminder [WMA*16, HGM*13, noaa] and adopting user interaction principles from Shneiderman [Shn96].

3.2. User Interface

We now describe the visualization techniques and features of the AtmoVis user interface.

The **Site View** (Fig. 1a) is the starting point for data analysis showing the monitoring sites on a map of New Zealand. Sites are represented as coloured circles with the colour depicting the intensity of the air quality value and additional information is shown when the mouse hovers over (e.g. name and value). Sites are added to other visualizations by drag and drop.

The **Monthly Rose** (Fig. 1b) depicts directional trends in the concentration of an air pollutant at a site. Dragging and dropping from the site view changes the site visualized. The month can be adjusted from a time selector.

The **Monthly Averages** (Fig. 1c) provides a monthly summary of all data recorded at a monitoring site. The labels for the bars can be dragged and dropped onto other open visualizations to display the variable. An options panel filters the data to fewer variables.

The **line plot** (Fig. 1d) displays hourly data to allow comparison between sites distinguished by line colour. Hovering over a measurement displays the time and value. One y-axis variable is shown at once for several different sites. The options panel can configure

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

the y-axis variable displayed. The line plot can be zoomed with the mouse scroll wheel.

The **Heat Calendar** (Fig. 1e) displays the yearly variation in air pollution. The day is coloured according to the mean value of the pollution measurements. Hovering the mouse over a day shows the mean value of the pollution on that day in a labelled text field under the calendar. Clicking on the day changes the time selector on the user interface to the start of that day.

The **Data Comparison** (Fig. 1f) (PCP) visualization shows correlations between different pollutants and other recorded variables (e.g. wind speed). An axis is provided for each variable in the data being compared. Different sites have a different line colour. A line for every hour in a 12 hour time frame is displayed with different site colours allowing relationships among variables to be discovered based on whether the lines point in the same direction intersect.

4. User Study

We conducted a user study to measure the effectiveness of AtmoVis in presenting air quality information to New Zealand environmental analysts and scientists so that inferences could be made, and to gather the information that can be used to make the tool more suitable for the target audience. Ethical approval was obtained from our university's human ethics committee. The following research questions were addressed in order to evaluate AtmoVis with environmental scientists:

RQ1 How effective are the visualization techniques in AtmoVis for exploring air quality data?
RQ2 How effective is the user experience of AtmoVis for exploring air quality data?
RQ3 How accurate are experts when using the visualizations in AtmoVis?

4.1. Participants

After a pilot study with 11 participants, 20 participants were recruited for the main user study. User study participants were selected experts in air quality analysis, and university graduate students with GIS experience. Most studies were done in person and a few online over video conference. For the participants 17 had experience with visualization software requiring programming experience such as R but stated they needed a more domain specific tool for better analysis. 13 participants identified as researchers, 5 participants identified as analysts and 3 participants identified as students. Categories are not mutually exclusive.

4.2. Procedure

The procedure involved a pre-study questionnaire, study tasks, and a post-study questionnaire. The design of the user study reflected the system goals (Section 3.1) to ensure that the study tasks were testing use cases that could occur outside the laboratory settings.

Each participant was given a brief description of the study goals. Instructional material was provided through videos and a web-based slide show describing AtmoVis. A think aloud protocol was followed [Nie93]. Participants were referred to parts of the question sheet, instructional slides, or video in response to their questions to ensure consistent responses by the researcher. The researcher took observation notes in a log book. A post-study questionnaire gathered information about the participant's experiences. We allowed up to two hours to complete the study.

4.3. Study Tasks

The study tasks were devised from the development of the tool, requirements analysis with our domain experts at NIWA, and through our pilot study. Initial tasks were short and direct asking the participants to perform actions, and record the results. These tasks introduced the participants to the functionality of the interface. Later tasks were open-ended and asked the participant to use a selection of visualizations to answer questions or to describe the data. All participant numbers are out of 20 unless otherwise stated.

Mapping the data Participants familiarised themselves with AtmoVis, identified the number of sites visible on a site view, identified the maximum level of a pollutant at a given site and compared pollutants at different sites on both the line plot and the PCP.

Aggregate data Participants used the heat calendar to 1) identify the month with the highest average level of the pollutant, 2) read the monthly rose plot to identify the relationship between a pollutant and wind direction, 3) compare two monthly rose plots to comment on the distribution of a pollutant with wind direction at two locations, 4) comment on similarities and differences between two pollutants at two locations, and 5) identify the worst day in the year for a pollutant and describe the trend on a line plot.

Parallel coordinate data comparison Participants used the data comparison in AtmoVis to observe and compare data from two sites over a 12-hour period. 12-hours was chosen so that users can see daily changes in the PCP when the play button is pressed Three pollutants were added to the data comparison and relationships among the pollutants at the sites were described.

Temporal pattern Participants used the line plot, site view, heat calendar, and monthly averages chart to observe trends and temporal patterns from a start date and commented on the line plot. Given a site and a pollutant participants used the site view, monthly averages, and heat calendar then commented on seasonal trends or days with the highest pollutant levels. Participants chose a site and used the monthly averages, site view, and monthly rose then commented on different pollutants and identified spatial trends.

5. Results

5.1. Visualization Perceived Effectiveness

Fig. 2 shows the distribution of Likert scale means for each visualization task covered in the post-study questionnaire as perceived effectiveness. Every visualization, other than data comparison, had a median Likert scale value over the centre of the scale (>3). The order of visualizations from best median response to the worst is heat calendar (4.38), line plot (3.60), monthly rose (3.33), site view (3.17), monthly average (3.10), and data comparison (2.75). The

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

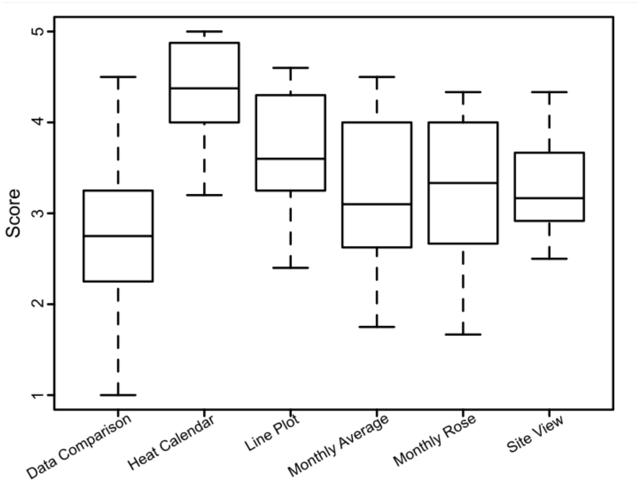

Figure 2: *AtmoVis: perceived effectiveness of each visualization technique measured by the post-study questionnaire Likert scales aggregated by the method described in Section 5.1.*

range for the heat calendar (1.80) and the site view (1.83) were smallest respectively showing more consistency in the responses for those visualizations. The IQR of the heat calendar was 0.88 and the IQR of the site view was 0.75. The results clearly show that the heat calendar was perceived to be more effective than the other visualizations and participants struggled with the usability of the data comparison visualization regardless of experience.

5.2. Statistical Analysis

Welch's t-test was chosen for determining significant difference between rankings as the samples are small, and variances are not necessarily equal between groups being tested. Table. 1 groups the mean Likert scale test scores based on participants experience analysing air quality data. The table describes the mean, standard deviation, and the number of participants in each group. Welch's two-sample t-test was used to determine whether the two groups "no air quality experience" and "air quality experience" have the same mean.

The heat calendar ($t(17.54) = 1.04$, $p = .311$, $d = 0.45$), line plot ($t(14.51) = 1.80$, $p = .093$, $d = 0.83$), monthly rose ($t(14.29) = 0.62$, $p = .546$, $d = 0.29$), and site view ($t(15.53) = 0.51$, $p = .620$, $d = 0.23$) did not display statistically different means. This result indicates the participant's statement of their experience with air quality data did not significantly alter the participant's Likert scale rankings of their experience with the visualizations.

The monthly averages ($t(15.65) = 3.70$, $p = .002$, $d = 1.67$) and data comparison ($t(11.40) = 2.20$, $p = .049$, $d = 1.09$) are statistically different means. This result indicates participants stated experience with air quality data significantly altered the participant's Likert scale rankings of their experience with the monthly average plot. Participants with no experience in air quality data analysis gave more positive responses.

Table 1: *Summary statistics for the mean Likert scores for each visualization in the post-study questionnaire grouped by experience.*

	Air Quality Experience					
Vis	No Experience			Experience		
	Mean	SD	n	Mean	SD	n
Heat Calendar	4.5313	0.4519	8	4.2875	0.5905	12
Line Plot	3.9458	0.6794	8	3.4000	0.6407	12
Monthly Rose	3.3750	0.8626	8	3.1389	0.7972	12
Site view	3.3333	0.5419	8	3.2063	0.5610	12
Monthly Average	3.8750	0.6682	8	2.7250	0.6995	12
Data Comparison	3.2375	0.9661	8	2.3727	0.6424	11

5.3. Temporal Overview (Heat Calendar)

All participants responded that the calendar was useful for identifying times with high pollution and 19 participants indicated the colour coding was useful in the heat calendar view. 17 participants responded that the heat calendar was useful for time navigation. 18 participants felt they did not need much assistance. All participants found the month with the highest daily mean level of PM_{10} in Woolston using heat calendar, demonstrating that the heat calendar was effective at showing the daily mean level of a pollutant. Participants made some positive comments about how the heat calendar shows yearly trends.

> "The heat calendar is very clear in terms of understanding some important trends there." - PID 7

Participants responded positively to how the heat calendar allows different days to be selected and how it interacts with other visualizations.

> "It automatically responds when you click on the day, I got ya, that's pretty cool." - PID 20

5.4. Temporal Comparison (Line Plot)

The line plot was considered effective for temporal data analysis. 13 participants found the line plot effective for finding temporal patterns. 11 participants found the play button effective for finding temporal patterns and 11 participants responded that the line plots use of colour was effective for interpreting the data. Mouse interaction received positive responses with 13 participants indicating the mouse navigation was easy or very easy. 12 participants stated that they did not require assistance with the line plot. 19 participants

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

were able to insert a location from the site view onto the line plot. 11 participants correctly identified the date and time for the highest pollution visible on the plot. 3 participants found the correct date and incorrect time, while 5 participants looked for a higher pollution level using the play button. 18 participants correctly identified the relationship between a specified pollutant measured at two different sites demonstrating that the line plot is effective for trend comparison. 16 participants found the relationship between a specified pollutant and solar radiation by comparing two line plots. The starting time was confusing to 4 participants because the axis was labelled with one hour offsets from the current time. 6 participants suggested adding functionality to support more than one dependent variable on a line plot.

5.5. Monthly Rose

Responses were positive with 14 participants stating that the monthly rose was effective for identifying pollutants, and 12 participants stated that the monthly rose was effective for finding relationships between data variables. 11 participants stated that they needed little or very little assistance with the monthly rose. 15/19 participants found some trend in the wind direction and concentration of a pollutant at a single site location and 9/19 found the correct relationship between both variables. 5/19 participants needed assistance to comment on the monthly roses but were able to read the visualizations with assistance. 5/19 participants correctly compared both wind direction and the level of a pollutant at two locations. 14/19 participants gave partially correct solutions. 4 participants commented that the monthly rose was effective for analysing the data. 4 participants were unfamiliar with rose plots and were unsure of the information displayed. The monthly rose was effective for describing the trend for a single variable in wind direction and concentration. The colour scale for a single monthly rose was effective, however, 6 participants commented that distance between colour categories was not identical across multiple monthly roses making the comparison more difficult. The interaction between the monthly rose and the heat calendar received some positive responses. Two participants stated that they would like to use AtmoVis to use the monthly rose.

> *"That's quite useful, I quite like that feature, it's certainly a lot quicker as a data visualizer, I mean you could do all this in R of course but its obviously a lot quicker than mucking around with the R script itself."* - PID 20

5.6. Site View

Responses to the site view Likert scale questions were positive with 14 participants indicating the mouse was easy or very easy to use for navigation, suggesting that participants were able to navigate and use the map to access sites effectively. 9 participants reported that the information displayed in the options panel was effective. 8/18 indicated colour change was effective or very effective for identifying temporal patterns. 9 participants indicated that the use of colour on the map was effective for representing the data collected at each site, though the site view was effective for temporal analysis. Participants generally felt that they did not need much assistance, 7 participants felt that they did not need assistance with the site view while 8 felt that they neither needed little assistance nor much assistance. The use of colour was generally effective. 6 participants recognised the number of sites with a specified pollutant demonstrating that interesting sites can be identified based on colour value, however, 6 participants found the colour scale difficult for comparing sites with low levels of pollution. The sites could be found easily when location was known, however, 5 participants discussed the difficulty of finding particular sites.

5.7. Monthly Averages

The participant responses to the monthly averages visualization were generally positive. 12/19 participants responded that the monthly averages visualization was effective for changing variables on the site view and 9 participants found the monthly average plot effective for identifying pollutants of interest. 10/19 participants responded that monthly averages were ineffective for finding relationships between pollutants. The monthly average plot was easy to use; 11 participants responded that they required little or very little assistance with the monthly averages. 17 participants were able to use the monthly averages to drag and drop a pollutant onto a monthly rose. 19 participants used the monthly averages to report some of the measurements available at a site, however, a large number of measurements were available and participants did not list all measurements. Dragging and dropping variables onto the site view was a strength and one participant responded with the following comment after dropping a label.

> ''Right so you can change what the site is viewing based on the labels, ok cool, that's cool." - PID 18

5.8. Data Comparison (PCP)

The data comparison tool was not favoured by participants. 10/19 participants stated that the data comparison was ineffective or very ineffective for finding relationships among pollutants. 11/19 participants correctly interpreted the relationship between two specified pollutants. The data comparison was difficult for the participants to interpret. 4 participants commented on the scales chosen for the data comparison. The scales were automatically chosen based on the highest measured value for the variable displayed on the axis, the scales could be changed by dragging and zooming the axis. One participant chose not to answer this question, no reason was given.

6. Discussion

RQ1 Participants used the heat calendar, line plot, monthly rose, monthly averages, site view, and data comparison visualizations to explore the data and identify relationships among variables in the dataset. The heat calendar was ranked the most effective with a high median score and the other visualizations, except data comparison, were considered effective. The heat calendar and line plot were most effective for assessing and comparing temporal trends, and required the least assistance. Navigation with the heat calendar received positive feedback. The data comparison and the monthly average were the most contentious visualizations. A larger IQR or range on the box plot diagram (Fig. 2) shows a wider range of responses regarding the effectiveness of these two visualizations.

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

AtmoVis was effective for determining directional trends for a single variable at a monitoring site. Comparing monthly roses was less effective and only one month could be visualized at once. The system could be improved by allowing visualizations at different time frames for easier comparison.

The heat calendar was more effective than the monthly averages for temporal comparison. However, the monthly averages visualization provided a method for finding out which pollutants were available at a given site.

Comparing different pollutants with the line plot at the same site is an important feature. Consequently, building more comparative functionality into the other visualizations would improve the usability of the interface.

The data comparison received the lowest overall effectiveness score and there was a statistically significant difference between users with experience analysing air quality data and users without. The participants without air quality experience gave more positive responses compared to users with experience. The data comparison required the most assistance.

RQ2 The user study identified that the user experience of AtmoVis was very effective for the exploration of air quality data. Participants were observed dragging, dropping, and clicking window borders to organize the layout. The use of visualizations within the windows indicates that windowing was effective but three participants suggested functionality for automatically tiling the visualization windows would assist with organising them.

> "It would be great if the tiles that I put them in were not overlapping each other" - PID 15

Dragging labels from the monthly averages onto another visualization to configure the variables received some positive feedback, though participants did not find this as intuitive as we expected based on the results of the Lyra [SH14] user study where participants were asked whether they felt drop zones were natural to use.

RQ3 The heat calendar and the line plot were found to be the most accurate for temporal analysis. Comparing monthly rose plots for different time frames was less accurate because participants were not able to see visualizations for both time frames at once and the data comparison was more difficult to read than the line plot when comparing two variables over a time frame. The line plot and the monthly rose were accurate for spatial analysis.

Limitations. The number of air quality experts in New Zealand is limited so recruiting is challenging hence participants with domain specific knowledge in GIS were also included who may not have had the same previous experience with the dataset. AtmoVis only supported the use of one monitor and there were differences in window placement for application windows in single monitor and multi-monitor systems. The laboratory style test was chosen because incorporating AtmoVis into the daily work of a group of participants was not practical at this stage. Additional testing is required to determine whether the windowed layout is more suitable than a tiled layout. Due to the scope of the project, the computational performance of generating the visualizations and the network performance were not rigorously tested.

Future Work. Additional datasets such as water quality and traffic data could be added to AtmoVis. AtmoVis could be applied to data from other countries. Saved layouts could be applied to position windows on the screen to reduce the amount of time spent rearranging windows, also sessions could be saved and loaded. AtmoVis does not provide many statistical metrics for comparing variables in the data, though this functionality could be added. The site view could be extended with a search feature to make it easier for participants to locate, as some participants had difficulty finding sites. The site view could also be extended by adding overlays for the regions or allowing more than one pollutant to be visualized at once. A longitudinal study of the usability of the tool could be performed to evaluate AtmoVis over a longer time frame in a workplace environment to garner further richer insights [IIC*13, SP06, LBI*12].

7. Conclusion

Practitioners' experience using visualization tools for air quality data differ, programming tools are slow for exploration of datasets and require computing knowledge. AtmoVis supports exploring air quality data to compare spatio-temporal trends among air quality variables using pre-existing visualization techniques. We conducted a user study with 20 experts to evaluate the effectiveness of the different visualizations and interface for the analysis of an air quality monitoring dataset. AtmoVis provides a faster way to access the rose plot from Openair without needing programming experience. The heat calendar and the line plot were the most overall effective visualization techniques for data analysis and comparing temporal trends. The site view, monthly averages, and monthly rose were also effective. The window layout was an effective method for accessing visualizations and inferring relationships among air quality variables at different monitoring sites. Participants were able to effectively use several windows at once to answer questions. The drag and drop interactivity between different visualizations received positive feedback with a high rate of completion. AtmoVis received generally positive feedback in the post-study questionnaires and the responses also identified some parts of AtmoVis which could be improved. The results of the user study demonstrate that air quality data analysis could benefit from interactive visualization through a web-based interface and there is a need for domain specific tools like AtmoVis to make better informed decisions about air pollution.

8. Acknowledgement

The authors would like to acknowlege Dr. Guy Coulson[†], A/Prof. Mairéad de Róiste[‡], and Dr. Stuart Marshall[‡] for reading this research paper and providing helpful comments and feedback.

References

[arc] ArcGIS, https://www.arcgis.com/index.html. Accessed: 03/10/2019. URL: https://www.arcgis.com/index.html. 1

† The National Institute of Water and Atmospheric Research, New Zealand

‡ Victoria University of Wellington, New Zealand

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.

[BKSL16] Burlinson D., Koehrn K., Subramanian K., Lu A.: Are environmental regulations working? a visual analytic approach to answering their impact on toxic emissions. In *Proceedings of the Workshop on Visualisation in Environmental Sciences* (Goslar, DEU, 2016), EnvirVis '16, Eurographics Association, p. 17–21. 1, 2

[BWS*06] Barnett A. G., Williams G. M., Schwartz J., Best T. L., Neller A. H., Petroeschevsky A. L., Simpson R. W.: The Effects of Air Pollution on Hospitalizations for Cardiovascular Disease in Elderly People in Australian and New Zealand Cities. *Environmental Health Perspectives 114*, 7 (July 2006), 1018–1023. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1513338/, doi:10.1289/ehp.8674. 1

[CR12] Carslaw D. C., Ropkins K.: openair — An R package for air quality data analysis. *Environmental Modelling & Software 27–28*, 0 (2012), 52–61. doi:10.1016/j.envsoft.2011.09.008. 2

[DWC*20] Deng Z., Weng D., Chen J., Liu R., Wang Z., Bao J., Zheng Y., Wu Y.: Airvis: Visual analytics of air pollution propagation. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (2020), 800–810. 2

[HGM*13] Hafen R., Gosink L., McDermott J., Rodland K., Dam K. K. V., Cleveland W. S.: Trelliscope: A system for detailed visualization in the deep analysis of large complex data. In *2013 IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)* (Oct. 2013), pp. 105–112. doi:10.1109/LDAV.2013.6675164. 2, 3

[HST*] Hales S., Salmond C., Town G. I., Kjellstrom T., Woodward A.: Daily mortality in relation to weather and air pollution in Christchurch, New Zealand. *Australian and New Zealand Journal of Public Health 24*, 1, 89–91. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-842X.2000.tb00731.x, doi:10.1111/j.1467-842X.2000.tb00731.x. 1

[HWA*07] Huamin Qu, Wing-Yi Chan, Anbang Xu, Kai-Lun Chung, Kai-Hon Lau, Ping Guo: Visual analysis of the air pollution problem in hong kong. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1408–1415. 2

[IIC*13] Isenberg T., Isenberg P., Chen J., Sedlmair M., Möller T.: A Systematic Review on the Practice of Evaluating Visualization. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (Dec. 2013), 2818–2827. doi:10.1109/TVCG.2013.126. 7

[JBMS09] Janicke H., Bottinger M., Mikolajewicz U., Scheuermann G.: Visual exploration of climate variability changes using wavelet analysis. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (Nov. 2009), 1375–1382. URL: https://doi.org/10.1109/TVCG.2009.197, doi:10.1109/TVCG.2009.197. 2

[KMK06] Koua E. L., Maceachren A., Kraak M. : Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographical Information Science 20*, 4 (Apr. 2006), 425–448. URL: https://doi.org/10.1080/13658810600607550, doi:10.1080/13658810600607550. 2

[LBI*12] Lam H., Bertini E., Isenberg P., Plaisant C., Carpendale S.: Empirical Studies in Information Visualization: Seven Scenarios. *IEEE Transactions on Visualization and Computer Graphics 18*, 9 (Sept. 2012), 1520–1536. doi:10.1109/TVCG.2011.279. 7

[LFM16] Li H., Fan H., Mao F.: A Visualization Approach to Air Pollution Data Exploration—A Case Study of Air Quality Index (PM2.5) in Beijing, China. *Atmosphere 7*, 3 (2016). URL: http://www.mdpi.com/2073-4433/7/3/35, doi:10.3390/atmos7030035. 2

[Nie93] Nielsen J.: *Usability Engineering*. Academic Press, Inc., 1300 Boylston Street, Chestnut Hill, MA 02167, 1993. 4

[noaa] Gapminder Tools, https://www.gapminder.org/. Accessed: 04/07/2019. URL: https://www.gapminder.org/. 3

[noab] NIWA, https://www.niwa.co.nz/. Accessed: 04/07/2019. URL: https://www.niwa.co.nz/. 1

[noac] Niwa Weather, https://weather.niwa.co.nz/. Accessed: 04/07/2019. URL: https://weather.niwa.co.nz/. 1

[noa18] New Zealand's Environmental Reporting Series: Our air 2018. In *New Zealand's Environmental Reporting Series: Our air 2018*. Ministry for the Environment & Stats NZ, 2018, pp. 12–40. URL: www.mfe.govt.nzandwww.stats.govt.nz. 1, 2

[PA05] Pruitt J., Adlin T.: *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005. 3

[RLS*96] Roth S. F., Lucas P., Senn J. A., Gomberg C. C., Burks M. B., Stroffolino P. J., Kolojechick A. J., Dunmire C.: Visage: A User Interface Environment for Exploring Information. In *Proceedings of the 1996 IEEE Symposium on Information Visualization (INFOVIS '96)* (Washington, DC, USA, 1996), INFOVIS '96, IEEE Computer Society, pp. 3–12. URL: http://dl.acm.org/citation.cfm?id=857187.857618. 2

[SH14] Satyanarayan A., Heer J.: Lyra: An Interactive Visualization Design Environment. In *Proceedings of the 16th Eurographics Conference on Visualization* (Aire-la-Ville, Switzerland, Switzerland, 2014), EuroVis '14, Eurographics Association, pp. 351–360. URL: http://dx.doi.org/10.1111/cgf.12391, doi:10.1111/cgf.12391. 2, 7

[Shn96] Shneiderman B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages* (Washington, DC, USA, 1996), VL '96, IEEE Computer Society, pp. 336–343. URL: http://dl.acm.org/citation.cfm?id=832277.834354. 2, 3

[SP06] Shneiderman B., Plaisant C.: Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth Long-term Case Studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization* (New York, NY, USA, 2006), BELIV '06, ACM, pp. 1–7. URL: http://doi.acm.org.helicon.vuw.ac.nz/10.1145/1168149.1168158, doi:10.1145/1168149.1168158. 7

[STH08] Stolte C., Tang D., Hanrahan P.: Polaris: A System for Query, Analysis, and Visualization of Multidimensional Databases. *Commun. ACM 51*, 11 (Nov. 2008), 75–84. URL: http://doi.acm.org/10.1145/1400214.1400234, doi:10.1145/1400214.1400234. 2

[WMA*16] Wongsuphasawat K., Moritz D., Anand A., Mackinlay J., Howe B., Heer J.: Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (Jan. 2016), 649–658. doi:10.1109/TVCG.2015.2467191. 2, 3

[ZLH13] Zheng Y., Liu F., Hsieh H.-P.: U-Air: When Urban Air Quality Inference Meets Big Data. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY, USA, 2013), KDD '13, ACM, pp. 1436–1444. URL: http://doi.acm.org/10.1145/2487575.2488188, doi:10.1145/2487575.2488188. 2

[ZWMT16] Zhou M., Wang R., Mai S., Tian J.: Spatial and temporal patterns of air quality in the three economic zones of China. *Journal of Maps 12*, sup1 (2016), 156–162. URL: https://doi.org/10.1080/17445647.2016.1187095, doi:10.1080/17445647.2016.1187095. 2

© 2022 The Author(s)
Eurographics Proceedings © 2022 The Eurographics Association.