

AtmoVis: Visualization of Air Quality Data

by

Benjamin Thomas Powley

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the
requirements for the degree of
Master of Science
in Computer Science.

Victoria University of Wellington
2019

Abstract

Air quality has an adverse impact on the health of people living in areas with poor quality air. Monitoring is needed to understand the effects of poor air quality. It is difficult to compare measurements to find trends and patterns between different monitoring sites when data is contained in separate data stores. Data visualization can make analyzing air quality more effective by making the data more understandable. The purpose of this research is to design and build a prototype for visualizing spatio-temporal data from multiple sources related to air quality and to evaluate the effectiveness of the prototype against criteria by conducting a user study. The prototype web based visualization system, AtmoVis, has a windowed layout with 6 different visualizations: Heat calendar, line plot, monthly rose, site view, monthly averages and data comparison. A pilot study was performed with 11 participants and used to inform the study protocol before the main user study was performed on 20 participants who were air quality experts or experienced with Geographic Information Systems (GIS). 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 the most effective for temporal data analysis. The interactive web based interface for data exploration with a window layout, provided by AtmoVis, was an effective method for accessing air quality visualizations and inferring relationships among air quality variables at different monitoring sites. AtmoVis could potentially be extended to include other datasets in the future.

Acknowledgments

I would like to gratefully acknowledge the support of my supervisors Dr Craig Anslow and Dr David Pearce, School of Engineering and Computer Science, Victoria University of Wellington.

I would also like to thank Dr Guy Coulson from the National Institute of Water & Atmospheric Research Ltd (NIWA) for helpful discussions and support during the research study.

Sincere thanks to all of the participants who volunteered their time and effort to assist.

Thanks to my mother Dr Janet Webster for helpful encouragement and patience.

Contents

1	Introduction	1
1.1	Research Questions	4
1.2	Contributions	5
1.3	Thesis Outline	6
2	Background	7
2.1	Air Pollution	7
2.2	Information Visualization	14
2.2.1	Taxonomies	14
2.2.2	Exploration	16
2.2.3	Drill Down	19
2.2.4	Breadth First	19
2.2.5	Metrics	22
2.3	Evaluation Methods	23
2.3.1	Visualization User Study	24
2.3.2	Summary	27
3	AtmoVis	31
3.1	AtmoVis: Design	31
3.1.1	Personas	32
3.1.2	Kath: Air Quality Scientist Persona	34
3.1.3	Oliver: Data Analyst Persona	35
3.1.4	Mathew: Student Persona	36

3.1.5	AtmoVis: System Goals	37
3.2	Architecture	38
3.2.1	Front-End	39
3.2.2	Back-End	40
3.3	Implementation	41
3.3.1	Data	43
3.4	User Interface	44
3.4.1	Windowing	44
3.4.2	Control Options	46
3.4.3	Heat Calendar	46
3.4.4	Line Plot	48
3.4.5	Monthly Rose	50
3.4.6	Site View	51
3.4.7	Monthly Averages	53
3.4.8	Data Comparison	54
3.5	Design Analysis	56
3.6	Summary	57
4	User Study	59
4.1	Study Participants	59
4.2	Study Procedure	60
4.2.1	Study Methodology	60
4.2.2	Pilot Study	61
4.2.3	Main Study	62
4.3	Study Tasks	62
4.3.1	Task 1: Mapping The Data	63
4.3.2	Task 2: Aggregate Data	64
4.3.3	Task 3: Parallel Coordinate Data Comparison	65
4.3.4	Task 4: Temporal Pattern	65
4.4	Data Collection	66
4.5	Measuring Effectiveness	67

4.5.1	Analysis of the Data	68
5	Results	71
5.1	Pilot Study	71
5.1.1	Heat Calendar	73
5.1.2	Line Plot	73
5.1.3	Site View	73
5.1.4	Monthly Averages	74
5.1.5	Data Comparison	74
5.1.6	Windowing	75
5.1.7	Summary	77
5.2	Main User Study	78
5.2.1	Participants	79
5.3	Visualization Evaluation	80
5.3.1	Visualization Effectiveness	81
5.3.2	Heat Calendar	82
5.3.3	Line Plot	86
5.3.4	Monthly Rose	91
5.3.5	Site View	96
5.3.6	Monthly Averages	101
5.3.7	Data Comparison	105
5.3.8	Windowing	110
5.4	Discussion	111
5.4.1	Accuracy of AtmoVis for Persona Scenarios	111
5.4.2	Heat Calendar	115
5.4.3	Line Plot	116
5.4.4	Monthly Rose	117
5.4.5	Site View	118
5.4.6	Monthly Averages	119
5.4.7	Data Comparison	121
5.4.8	Windowing	122

5.5	Suggested Improvements	123
5.6	Summary	124
6	Conclusions	127
6.1	Contributions	128
6.2	Research Questions	130
6.3	Future Work	131
	Appendices	135
A	Information Sheet	137
B	Pre-Study Questionnaire	141
C	Study Instructions	149
D	Post-Study Questionnaire	155
E	Instructional Slides	167

List of Figures

1.1	A temperature inversion layer with hot air above and cold air underneath. The inversion layer prevents particulate matter from rising. Drawn By: B. Powley.	4
2.1	Two different methods for drawing calendar heat maps. . .	8
2.2	PM_{10} U-Air machine learning compared with linear interpolation between monitoring sites [68]. The linear interpolation is less accurate. The green at the top right monitoring site has been inaccurately spread out over a large area. . . .	9
2.3	AOT interpolated over Makkah Mina and Arafah in Saudi Arabia using ordinary Kriging interpolation [45].	11
2.4	Tree map containing various pollutants on separate nodes [38].	13
2.5	Two different views inside the TIARA system.	17
2.6	An example of the Trelliscope system being used to detect generator trips [30]. The two subsets of the data are presented for comparison, both have been recommended by Trelliscope as possible generator trips. The recommendations can then be categorized depending on whether they are generator trips.	18
2.7	A cognostic which provides an interactive way to select a section of a bar chart to filter on [30].	22

2.8	A cognostic which provides an interactive way to filter on a section of the line chart [30].	22
3.1	The AtmoVis user interface. Map: openstreetmap.org [8] . .	32
3.2	Architectural diagram describing AtmoVis.	39
3.3	AtmoVis interface controls: a) window border, b) window resize handle, c) delete button, d) options panel button, e) text title, f) options panel button, g) play button, h) time selector, i) load button, j) menu, k) menu button.	45
3.4	A heat calendar with the mouse hovered over a day, the average ozone pollution for that day is shown at the bottom of the screen.	47
3.5	A line plot showing ozone measurements at Musick Point with the mouse hovering over a measurement showing that the value of ozone on 02/01/2012 at 17:00 hours was $40\mu\text{gm}^{-3}$	49
3.6	The monthly rose diagram shows the concentration of ozone in colour and the frequency of counts by wind direction as the length of each sector from the centre.	50
3.7	The site view centred over Auckland with 19 monitoring sites visible and 3 of the monitoring sites have ozone data available coloured in red. The other sites are clear and have no data available.	52
3.8	The monthly averages for a given site with the mouse hovering over a label, the label can be dragged and dropped to relate the data to other visualizations.	53
3.9	A data comparison parallel coordinate plot with the mouse hovering over one measurement to show the details.	55

5.1	The window frame border for the visualizations and the options panel. The top image is the frame for the visualization with a highlighted gear icon which the mouse has hovered over. The bottom image is the frame for the options panel with a back button instead of a gear icon.	76
5.2	Effectiveness of visualization measured by the post-study questionnaire Likert scales aggregated by the method described in Section 5.3.1	81
5.3	Heat Calendar: Likert scales from post-study questionnaire.	84
5.4	Line Plot: Likert scales from post-study questionnaire. . . .	88
5.5	Monthly Rose: Likert scales from post-study questionnaire.	92
5.6	Site View: Likert scales from post-study questionnaire. . . .	98
5.7	Monthly Averages: Likert scales from post-study questionnaire.	102
5.8	Data Comparison: Likert scales from post-study questionnaire.	107
5.9	Four different usage styles in AtmoVis, Figure 5.9a shows window splattering, Figure 5.9b shows border to border careful tiling, Figure 5.9c shows a near maximized window and Figure 5.9d shows a stack of windows.	123

Chapter 1

Introduction

There are several different air pollutants in the atmosphere including particulate matter, ozone, nitrogen oxides and sulphur dioxide that can cause health issues for people. These pollutants have a variety of different adverse health effects [15]. For example the health effects from particulate matter include cardiovascular disease [18] and respiratory mortality [31]. Sources of air pollution include traffic, domestic fires, industrial sources, and shipping [15]. In order to understand the effects of air pollution on health and on the environment, data containing air pollution information needs to be collected and analysed. Analyzing a large multivariate dataset is difficult, a thorough exploration of the data is needed to find relationships in air pollution which are suitable for further analysis.

Data is visualized by looking at an image representation of the data, the image can be produced by any method and the data is interpreted and understood using human perception to produce an understanding of the image [56]. Exploration of a dataset using a visualization can help to form questions about data even when there is little knowledge of what the data is like at the start of the process [28]. In this thesis a prototype system for visualizing spatio-temporal data related to air quality is designed, built, and evaluated. Evaluating the visualization system through a user study determines the effectiveness of the system for answering questions related

to air quality and the results of the study could be used to inform the design of other visualization systems for environmental data. A novel web-based exploratory visualization system, *AtmoVis*, was developed as part of this research. *AtmoVis* was designed for breadth first exploration of air quality data.

Data visualization can improve on the current methods for data analysis by making data more understandable and by providing an interactive way to search, filter, and find relationships within the dataset. The data collected on air pollution has some problems and limitations that can make data analysis challenging. Research sites can collect different variables over different time frames. For example, upgrades at a measurement site can cause different data collection methods to be used for a given substance. Not all of the variables are collected at every site and some sites are only partially complete. Finding data collected using a visualization interface would be easier because the sites collecting a given pollutant could be visualized. The National Institute of Water and Atmospheric Research (NIWA) has provided a dataset containing air quality and meteorological data for the purpose of developing an air quality visualization interface. NIWA is a Crown Research Institute that performs research about New Zealand climate change and weather forecasting [6, 7]. The atmospheric dataset provided by NIWA for this research has a large number of sparsely collected variables over a range of time frames. In New Zealand there are some key locations where air quality is poor and there is seasonal variation in air quality, though in general air quality is good. For example, there are sites located at Auckland, Taupo, Wellington, Christchurch, and Dunedin. An effective visualization system would improve the analysis of the dataset provided by allowing the sites in the dataset to be explored interactively and analyzed through a graphical interface. An exploratory data visualization system could potentially help to form questions about air quality in New Zealand for future research. Air quality experts were interviewed to create personas and the personas were

used to identify the goals and functionality of the AtmoVis system. Pilot testing was conducted to identify necessary adjustments to system functionality and evaluation protocol before a main user study was conducted on air quality experts and participants with experience in Geographic Information Systems (GIS).

In the NIWA air quality dataset the meteorological measurements provide for wind speed, wind direction, rainfall, solar radiation, air temperature and humidity. Atmospheric measurements are for small particulate matter (PM_{10} , $PM_{2.5}$), nitrogen monoxide (NO), nitrogen dioxide (NO_2), both nitrogen monoxide and nitrogen dioxide (NO_x), carbon monoxide (CO), sulphur dioxide (SO_2), and ozone (O_3). PM_{10} is a measure of particulate matter concentration of particles smaller than 10 microns, and $PM_{2.5}$ is a measure of particulate matter concentration of particles smaller than 2.5 microns [34]. Weather conditions can affect the amount of a pollutant measured. For example, in a temperature inversion (Figure 1.1) there is hot air on top of cold air and the warm air prevents the pollution from rising which traps particulate matter underneath [15]. The data collected is in the form of numerical measurements at hourly or daily recorded times, and the time frame for data collection at a given site is specified in a metadata file. A way of visually comparing measurements from different sites and observing temporal trends in the data would make air quality data potentially more understandable. NIWA was consulted on their current methods for viewing data so that a visualization could be produced that would improve and augment the process of analysing data. Currently, NIWA has no specialized interactive visualization software in use, instead, software such as R is used to analyse data as well as spreadsheet software which is used to plot trends. R is a scripting language and the data can be analyzed by writing and modifying programs. The disadvantage of R is the difficulty of writing scripts and the amount of time required. An exploratory visualization system allows the dataset to be analyzed interactively without modifying program code. Spreadsheets are also used to

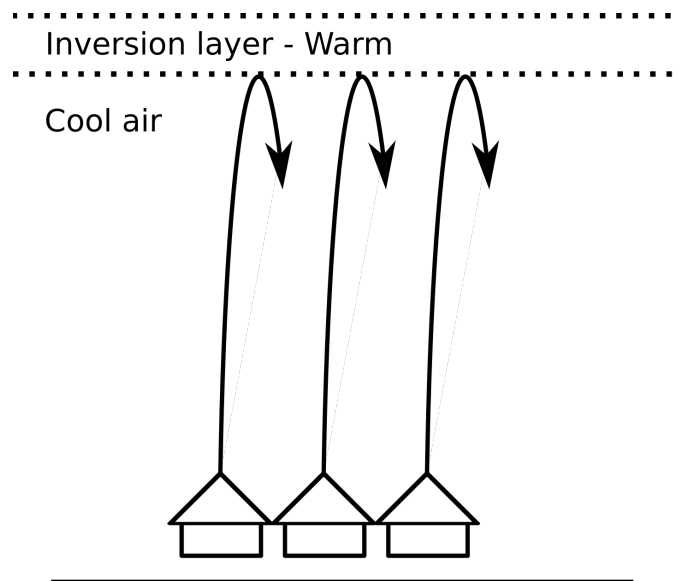


Figure 1.1: A temperature inversion layer with hot air above and cold air underneath. The inversion layer prevents particulate matter from rising. Drawn By: B. Powley.

analyze the data, however, spreadsheets are limited to the general visualizations that are provided in the software which are not developed specifically for air quality analysis. AtmoVis is a domain specific air quality application developed specifically for the dataset and allows the data to be explored interactively rather than by modifying scripts.

1.1 Research Questions

The following research questions were addressed in order to design and evaluate AtmoVis:

RQ1 How effective are the visualization techniques in AtmoVis for exploring air quality data?

It is important to know which visualizations are effective because using ineffective visualizations can prevent the data from being understood properly and reduce the efficiency of using a visualization tool. In order to measure the effectiveness of AtmoVis a user study was conducted relating to the air quality in New Zealand urban settings. We evaluated the usability of the design by comparing it to interface design taxonomies [42, 52, 66].

RQ2 How effective is the user experience of AtmoVis for exploring air quality data?

The user experience was measured based on participant responses to the post-study questionnaire. Analyzing user experience is important for producing effective domain specific tools and visualizations to ensure that they are comfortable and satisfactory to work with.

RQ3 How accurate are experts when using the visualizations in AtmoVis?

The accuracy of the results was measured by marking the results of the tasks that each participant completed as part of the user study. Accuracy is important as a visualization that is difficult to understand is less effective.

1.2 Contributions

The following list outlines the contributions of this thesis.

- A new web-based prototype data visualization system was developed to visualize air quality data.
- A user evaluation was conducted to find out information about the effectiveness of the visualization system and to make recommendations for the improvement of the tool.

1.3 Thesis Outline

The contents of the thesis is outlined as follows.

Chapter 2: Background Literature about visualizing air pollution, visualization systems and evaluation methods, is summarized, compared and contrasted.

Chapter 3: AtmoVis The personas for the target audience are described. System goals and the design analysis of the AtmoVis system are also presented.

Chapter 4: User Study The user study design is described and related to the personas and design of the AtmoVis interface.

Chapter 5: Results The results of the pilot study and main user study are presented then discussed.

Chapter 6: Conclusions The findings of the thesis are presented as well as suggestions for future work.

Chapter 2

Background

This chapter reviews and discusses previous methods used for the visualization of air quality, interface design used for visualization systems and evaluation techniques for measuring the effectiveness of a visualization system. Visualization is the process of interpreting data to produce an understanding of the information [56].

2.1 Air Pollution

Air pollution has previously been visualized using several different methods. In this section visualizations such as heat maps, line plots, minimum spanning trees, network flow graphs, hierarchical views and wind rose plots are reviewed and discussed.

A heat map is a visualization which uses colour to show the value of a variable displayed in a grid. A heat map can be used to visualize a variable as a colour overlay on a geographical map where the coordinates are continuous [37, 68] and it can also be used to visualize a variable in a discrete matrix by colouring the cells [37, 69]. Air pollution was visualized using heat map techniques to view temporal patterns in air quality (Figure 2.1). The colour of the entry in the cell represents the pollution level of a substance. Both Zhou et al. [69] and Li et al. [37] used heat

maps to display a calendar which allowed temporal relationships to be shown by highlighting days. In addition Zhou et al. also placed the heat map onto a geographical map so that spatial relationships could be compared [69]. Minimum spanning trees were used by Zhou et al. to position




Figure removed for copyright reasons.

(a) The “calendar view” heat map from Zhou et. al. [69].




Figure removed for copyright reasons.

(b) The circular heat map from Li et. al. [37].

Figure 2.1: Two different methods for drawing calendar heat maps.

cities on a grid map so that the calendars did not overlap [69]. The minimum spanning tree was used in a similar way to a force directed layout system. When the minimum spanning tree was not used the close cities had obscured calendar values. The minimum spanning tree was an effective way to re-position the heat maps to avoid overlap. The calendar displayed as a heat map would be useful for displaying a time series for air quality in New Zealand in the present study. The circular heat map

was found to be an effective visualization for finding the temporal characteristics of air pollution and times of the year were identified where air pollution was particularly serious [37]. Combining the minimum spanning tree with the heat map was effective for visualizing spatio-temporal air quality data [69].

Heat maps were used to visualize the spatial trends in air pollution over a geographical map [37, 68]. Since air quality is measured at discrete locations, the heat map uses an interpolation algorithm to calculate filled in colour over an area containing the data. There are many different interpolation algorithms so suitable methods need to be chosen for the data. For example, Li. et al. used Ordinary Kriging interpolation for calculating the values of the heat map between monitoring stations on a map of China [37]. In contrast, data from traffic, air quality records, points of interest and road networks were used to find the pollutant levels between monitoring stations in Beijing using machine learning [68]. The pollutants monitored include PM_{10} as shown in the heat map (Figure 2.2).



Figure 2.2: PM_{10} U-Air machine learning compared with linear interpolation between monitoring sites [68]. The linear interpolation is less accurate. The green at the top right monitoring site has been inaccurately spread out over a large area.

Of particular interest, the results from linearly interpolating PM_{10} across Beijing were significantly different from the machine learning result and the accuracy of the machine learning result and other interpolation techniques were calculated by removing a monitoring site and estimating the air quality at the monitoring site using the algorithm. Linearly interpolating a heat map in a city did not provide good results because when one monitoring station was removed, the air quality was interpolated incorrectly. The poor results for linear interpolation were due to the non-linear nature of air pollution [68].

Aerosol Optical Thickness (AOT) is a measure of how much light is blocked by aerosol particles in the atmosphere [45]. Kriging interpolation was also used to interpolate a heat map in a study comparing AOT to PM_{10} [45]. The study found that AOT was linearly correlated with the level of PM_{10} [45] and two heat maps were drawn to visualize the level of PM_{10} and the AOT in the area surrounding Makkah, Mina and Arafah in Saudi Arabia. One of these heat maps is shown in Figure 2.3. The linear correlation between AOT and PM_{10} allows PM_{10} air quality categories to be inferred optically with different equipment [45]. The heat map visualization is an effective way to show that the AOT and the PM_{10} are showing a similar pattern. Kriging interpolation was tested for validity on predicting the levels of O_3 and PM_{10} inside counties in the United States [64]. When interpolating O_3 the study found that, of the counties with suitable monitoring stations, Kriging interpolation was only valid in California due to difficulties with spatial autocorrelation at other sites. When interpolating PM_{10} , different regional splits were required. The issues with applying Ordinary Kriging interpolation for PM_{10} in different counties suggest that validity would need to be checked before applying Ordinary Kriging interpolation to PM_{10} in areas of New Zealand. Any interpolation algorithm would need to be checked for validity before being applied to drawing a heat map. The research performed by Liu et al. showed that there are time-lagged trends in $PM_{2.5}$ in different cities

in China which are influenced by weather patterns [39]. The time-lagged trends show that in the regions measured the air pollution is transported along particular paths [39]. Heat mapping strategies that attenuate air pollution with distance from the sensor would not be valid between different cities as air pollution being transported could be predicted differently than the actual level.

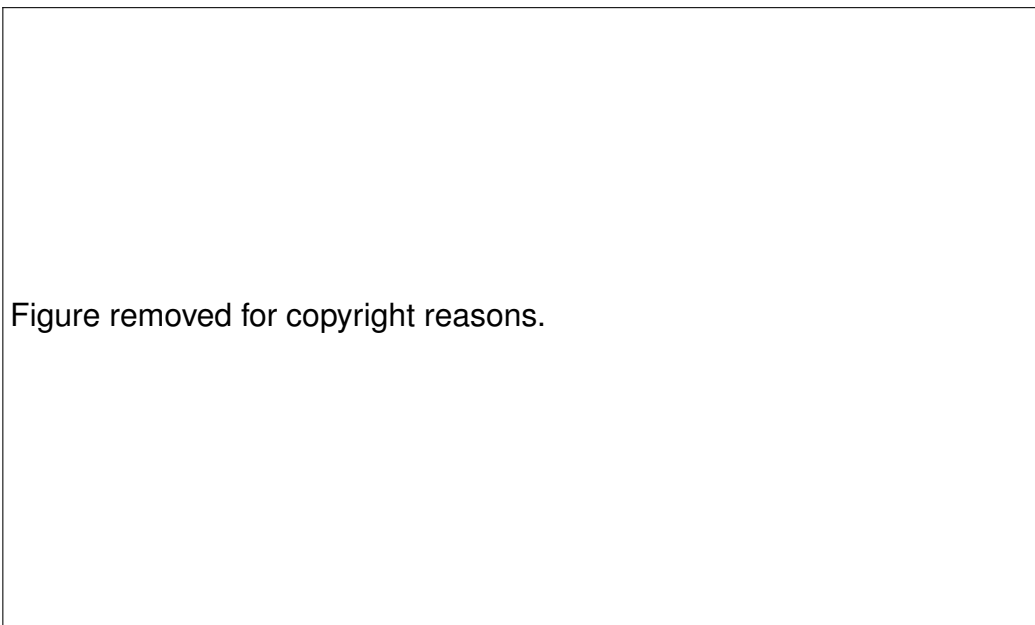


Figure 2.3: AOT interpolated over Makkah Mina and Arafah in Saudi Arabia using ordinary Kriging interpolation [45].

As an alternative to heat maps, the correlations of $PM_{2.5}$ time series was visualized graphically by using a weighted network flow graph [39]. In the network flow graph, each city was represented as a circle and the colour of the lines between the cities represented the correlation strength between $PM_{2.5}$ time series. The line had a direction representing the time lag between two cities [39]. One advantage of the network flow graph was that temporal pollution information was visualized, the heat map only showed spatial information.

Some air pollution visualization systems have been targeted at the wider

community. A study of the air pollution at a coke plant in Pittsburgh aimed to provide the community in the surrounding area with a tool for detecting, monitoring, and presenting the impact of the coke plant on the air quality so that the community could engage with policymakers to reduce air pollution in the region [32]. The system was built iteratively with input from the community and the community input made the system more effective and usable. Initially freeze frames were taken at timed intervals from a camera pointed at the coke plant to collect data, however, the system was later extended to use machine vision to identify short animated sequences. The system was presented as an online tool which allowed smell complaints about the coke plant to be recorded with statements from members of the community. The data could be visualized using the online tool. In the user study data was recorded about the web site use. The data provided quantitative statistics about the use of the tool. There was also a survey which contained questions about the way that the system impacted on the local community as well as questions about its operation. The coke plant was closed down as the result of community pressure after the visualization system was used to show air pollution demonstrating that the data visualization was a powerful tool for social change [32].

In yet another approach, visualization using hierarchical views to show air pollution time-series made use of coloured bar charts to display air pollution over the course of a year (Figure 2.4). The bar charts were on the leaves of the tree with each leaf node representing a collection station and parent nodes representing cities and provinces. The hierarchical time series was effective for comparing the levels of different pollutants however the hierarchical time series does not show a spatial relationship in the position of the cities. There was a time-series bar chart for each pollutant with the darkness of the colour indicating the severity of the air pollution [38]. The colour coding on the time series bar chart made the data easier to read and pollutants easier to compare in the same city. Colour coding is an

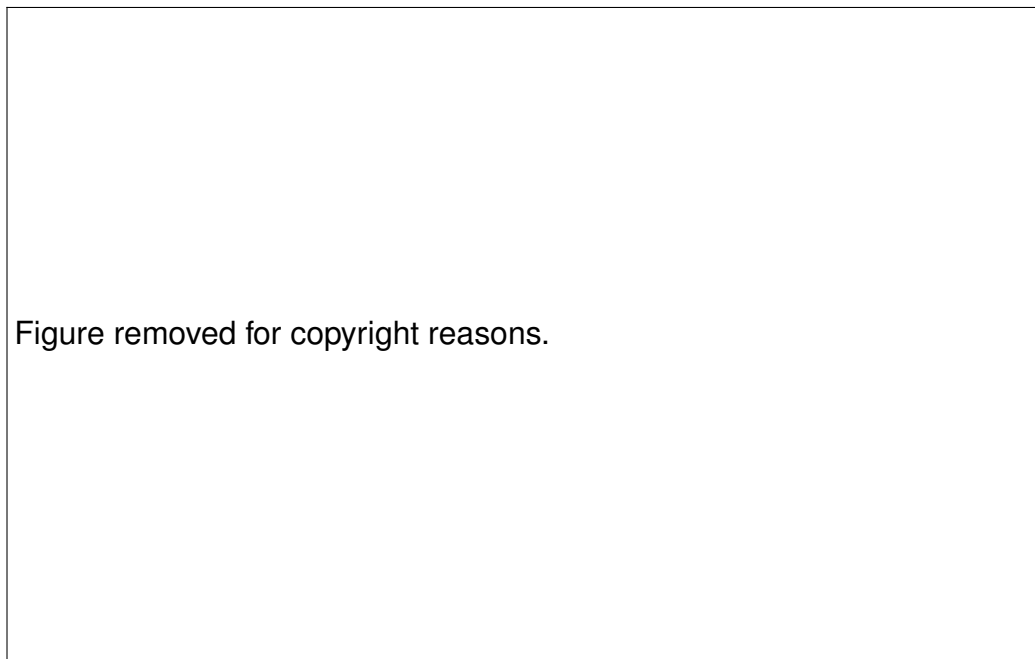


Figure 2.4: Tree map containing various pollutants on separate leaf nodes [38].

effective way to communicate the intensity of the air pollutants.

Not all visualization systems use graphical interfaces, the Openair tool for data visualization is a library for the R programming language and allows a variety of different statistical techniques to be applied to the data [23]. Openair has plotting functionality built in, for example, a calendar heat map can be drawn from a data set. The calendar heat map draws a monthly or a yearly calendar view with the day coloured according to the average level of a pollutant. If the wind direction is available then the calendar can draw an arrow inside the day for the average wind direction. Openair also contains a wind rose plot which plots the average pollution level for a given wind direction over a time frame. The wind rose plot was effective for visualizing how the wind direction and the level of a pollutant were related. Wind direction and air pollution were compared in an air quality report prepared by the Waikato Regional Council using a line plot [22], so

visualizing air pollution differences with wind direction is useful functionality to include in the prototype system being developed for this research and a wind rose plot could be incorporated into the system. The plots are built on top of other R libraries like Trellis which allows the plots to be customized with overlays [23].

2.2 Information Visualization

Three aspects of interface design were investigated for use during the development of this software project. Exploration, Drill Down, and Breadth First. A “Breadth-First” interface design can encourage a less focused view of the data and the exploration of a larger selection of the points in the dataset [65]. “Drill-down” allows for a more focused view at a smaller selection of the points in the dataset [16]. The drill down design principle can be used to filter the data from an overview to a smaller collection. Data exploration is the process of looking at data to see what it contains and the requirements for an exploration interface are different from the requirements for communicating data to others. Taxonomies for user interaction are reviewed for use when developing the system design. Taxonomies can be used to evaluate whether an interface can perform tasks that are frequently required when exploring data.

2.2.1 Taxonomies

Creating a taxonomy is a way of categorizing how visualization systems allow the user to interact with a visualization in order to explore data and find information in the system. Comparing a visualization to a taxonomy is a way of ensuring that necessary functionality has been included in the system. A taxonomy was created by Yi et al. [66] to classify user interaction techniques for visualization systems. The taxonomy categorizes techniques according to whether they *select, explore, reconfigure, encode, abstract/*

elaborate, filter or connect. Selecting the data allows points of interest to be identified. *Exploring* the data allows some of the data on the screen to be replaced with other data to build a better picture of what the data means. Re-configuring the data visualization system will allow the user to change how visualizations are positioned. *Encoding* changes the way that data is presented, for example providing different types of charts representing the same data or adjusting colour. *Abstracting* or elaborating on the data will change the level of detail of the data presented. *Filtering* the data will allow end users to focus on the part that is of most interest based on filter criteria. *Connect* can associate different parts of the data with each other based on relationships [66].

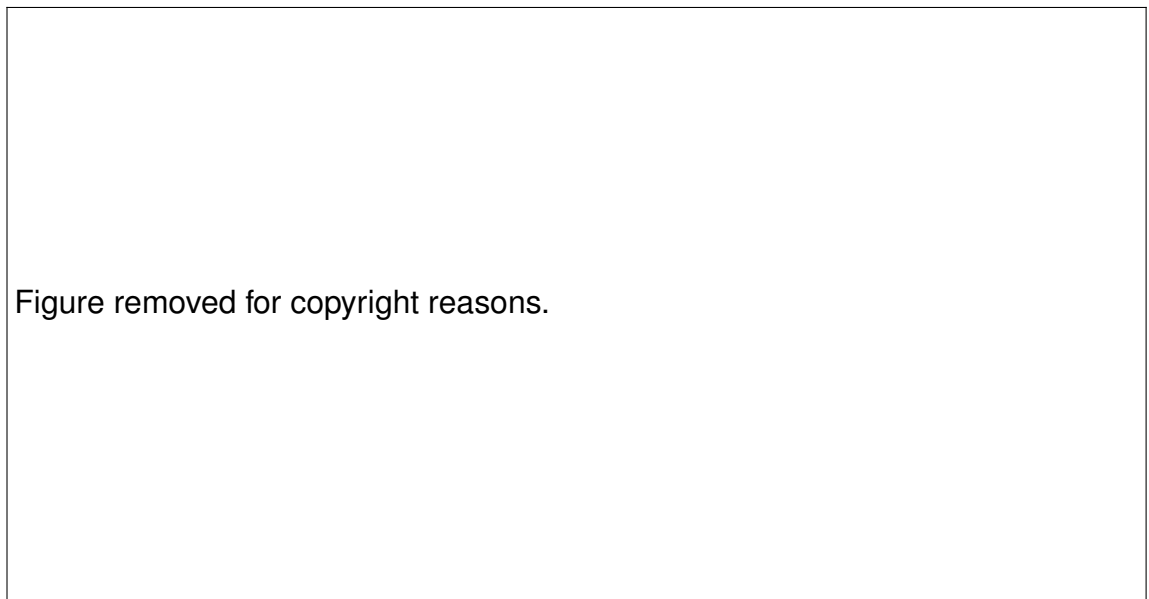
The information seeking mantra describes the tasks performed when exploring data. Shneiderman's taxonomy [52] described the following tasks: *Overview, zoom, filter, details-on-demand, relate, history* and *extract*. Each task was related to an activity that an end user performs when searching for data in a visualization system. The *overview* task is where the user looks at all the data. The *zoom* action allows the user to view the most useful part of the data. The *filter* task removes unneeded information so that the user can concentrate on what is important. The *details-on-demand* task allows the user to select parts of the data to view in more detail. The *relate* task happens when a user wishes to compare data and view relationships. The *history* option allows a user to undo unwanted actions. *Extraction* refers to the action of saving a subset of the data. The two taxonomies [52, 66] are similar because both have a way of inspecting the information at different levels. The taxonomy by Yi et al. has an explore task and the taxonomy described by Shneiderman has a zoom task. Filtering is featured in both taxonomies and connecting data is similar to relating data [52, 66]. There are some parts that do not directly correspond to each other, for example, the reconfiguration and encoding tasks do not have a directly corresponding categorization in the taxonomy described by Shneiderman. We can use techniques discussed in the taxonomies to

build a more effective visualization. In the visualization system prototype produced for this thesis, Breadth First techniques would be useful for presenting data to the user so that the entire dataset can be analyzed to find locations for further analysis based on air quality variables.

2.2.2 Exploration

The Voyager, TIARA, Polaris and Trelliscope systems were designed for data exploration. [30, 57, 62, 65] Different data exploration tools support a variety of different data sources. The exploratory interface can be produced to take advantage of the source, such as the interfaces produced for TIARA and Polaris, described below.

The TIARA system [62] was designed specifically to group a large collection of text documents based on “time sensitive” keywords and to automatically summarize them unlike the Voyager system [65] which was designed as a data exploration tool for multivariate tabular data rather than as a text exploration tool. The TIARA system was applied to both analysing patient records and summarizing emails. The patient records contain a small number of text fields and other structured fields which can be navigated using a facet view (Figure 2.5). The Polaris system was designed for exploring multidimensional databases specifically using a “Pivot Table” interface [57]. The pivot table allows different combinations of fields to be used for the row and column labels. Trelliscope is a system designed for visualizing trellis displays [30]. In a trellis display, a different subset of the data is displayed on each visualization and a variable is used to condition the visualization. Both trellis plots and the pivot table are ways of comparing a large number of visualizations. Trelliscope is implemented on top of tools supporting data extraction from a variety of sources making the tool more general than interfaces such as Polaris [30].



(a) The word cloud used by the TIARA system for describing emails inserted into the system [62].



(b) A zoomed view of the emails [62]

Figure 2.5: Two different views inside the TIARA system.

The Trelliscope system was demonstrated on a data set containing measurements from Phasor Measurement Unit (PMU's) on a power grid. The Trelliscope system was used to analyse the data for events (Figure 2.6) and to detect erroneous records for removal [30]. Trelliscope could potentially be used to detect temporal events in air quality monitoring data since the data extraction method allows the tool to be used generally. The TIARA system is designed as a text exploration tool, so would not be applicable to the dataset for air quality monitoring, though, the use of a “drill down” interface with faceted views demonstrates the Shneiderman’s information seeking mantra [52] with an “Overview first” of the email data “zoom and filter” to inspect the email data more closely then “details on demand” accessed through facets.



Figure 2.6: An example of the Trelliscope system being used to detect generator trips [30]. The two subsets of the data are presented for comparison, both have been recommended by Trelliscope as possible generator trips. The recommendations can then be categorized depending on whether they are generator trips.

2.2.3 Drill Down

The TIARA , Voyager, Visage and Tioga-2 systems were designed with a “drill-down” design principle [16, 49, 62, 65]. In a “drill-down” design a user can select variables to filter and extract information. The user interface of the TIARA system is based on a “stacked graph” visualization of keywords from the text [62]. The drill down can be performed by selecting words on the visualization. The Tioga-2 user interface supports “drill-down” behaviour using “wormholes” and by performing the “Set Range”, “Overlay”, and “Shuffle” operations on relations and composites. A wormhole allows a user to jump from one canvas to another canvas. Set range filters the displayed relation, overlay combines two composites, and shuffle changes the drawing order for relations [16]. The use of data exploration systems can be compared to the taxonomy performed by Shneiderman [52]. Both the TIARA system and the Tioga-2 system made use of the drill down principle. The drill-down principle would be useful in the design of a visualization system prototype for air quality monitoring data so that monitoring sites can be presented in an overview and then filtered and to find sub selections.

The Visage system allows users to drill down through variables in a table. The user can select variables from a menu to filter the display. Relations between database objects can also be used to drill down [49]. Data in the Lyra system can be adjusted with filters and other transformations [51]. The other tools such as the Polaris [57] and Trellescope [30] systems have made use of filters and transformations to adjust data.

2.2.4 Breadth First

Voyager, Trellescope, and Polaris have a breadth-first aspect to interface design. Voyager displays a large number of graphs inside facets [65].

In the Trelliscope system there are a large number of plots displayed at once on different panels [30]. The Polaris “Pivot Table” can contain rows columns and layers. An algebra was formally defined for generating tables. The algebra supports Concatenation, Cross product, and Nesting [57]. The algebra allows different scales to be produced for the plot visualization and also specifies layers [57]. In the Tioga-2 system, there are several windows containing visualizations. Windows represent “Tioga Programs” and contain boxes-and-arrows, a canvas, and a menu bar. Displayable visualizations can be “extended-relations”, “Composites” and “groups”. Extended-relations are relations from a RDBMS. Composites join relations together into a plot by superimposition, and groups combine composites together into a plot by displaying them in different parts of the visualization [16]. Both Tioga-2 and Polaris operate on data from a relational database. A visualization system with several different windows or facets like Voyager or Trelliscope would assist in the exploration of air quality data by allowing different visualizations to be related, compared and interacted with. An algebra to combine different visualizations is out of the scope of this research project but the prototype visualization system could be extended in the future to combine different air quality visualizations.

Faceted navigation is a technique for displaying information in different parts of a layout. The facets can be “superimposed” or “juxtaposed” [42]. The FacetMap visualization makes use of a bubble map layout where nodes are nested in bubble-shaped categories [55]. Clicking on areas of the facet bubble map can expand by producing SQL queries to get the next level of detail. The ResultMap visualization makes use of faceted navigation [27], The data is classified into nested categories using a tree map. ResultMap is specially designed to work with digital repositories of information and “search engine result pages”. The ResultMap was similar to the facet map in it’s presentation however the ResultMap did not use bubble shaped categories. Both the tree map and the bubble map were

effective at presenting hierarchical information, however, the leaf nodes can get very small to read. Also, the tree map layout nodes are not always a suitable shape for text so tree maps can be subdivided alternately between vertical and horizontal to make the shape more suitable [56]. Tree maps can be used to visualize a file system [42]. The RB++ system is an interface for faceted browsing. Text based overviews are provided to assist with navigation. Histogram style comparisons of the numerical data are also provided [67]. The “High-Dimensional Visual Analytics” system produced by L Wilkinson et. al. [63] consists of a scatter plot matrix of the data and a Features Plot. There are n^2 plots for n variables. Scatter plot matrices were used for visualizing air quality data. [37, 68]. The disadvantage of scatter plot matrices is the amount of space required to construct the matrix. It can be difficult to read scatter plot matrices when there are too many variables. There are interaction techniques for selecting and highlighting points to inspect an entry [63]. Using a histogram or a bar plot to compare and filter numerical data would assist with the visualization of air quality data and each air quality visualization could have a facet which allows the visualization to be filtered and configured. The air quality dataset provided for this research project has a large number of variables, and some pollutants such as PM_{10} and $PM_{2.5}$ have been recorded as several variables due to changes in monitoring equipment, so scatter plot matrices may not be suitable for visualizing the entire dataset. A relational database was built into the back end of the visualization system to provide access to the data so that the dataset can be accessed through a web page remotely. Visualizations were able to interact with the relational database by sending queries to a server to retrieve subsets of the data.

2.2.5 Metrics

A recommender system can be used to augment the analysis of the data by making recommendations, directing the exploration through the dataset [65], ranking visualizations and reducing the amount of time searching through the dataset [30]. Recommender systems can recommend variables and charts to look at. The Voyager system is an example of a visualization tool with a recommender engine. The recommender engine is called Compass and it can recommend visualizations to look at based on variables that the user has selected. The recommendation is performed by suggesting additional variables, performing a data transformation, encoding the data into visualizations, then ranking the visualizations based on metrics [65]. These visualizations are then rendered by Voyager’s interface. Trelliscope uses “cognostics” to decide which panels are displayed. A cognostic is a sort of metric. It represents a calculation performed on the data to describe the behaviour of the data. The cognostic metrics are used to decide on which panels are displayed [30] (Figures 2.7, 2.8).



Figure 2.7: A cognostic which provides an interactive way to select a section of a bar chart to filter on [30].



Figure 2.8: A cognostic which provides an interactive way to filter on a section of the line chart [30].

Voyager's compass recommender engine uses metrics to rank responses to be displayed by Voyager [65]. The Features Plot for "High-Dimensional Visual Analytics" uses metrics to rank scatter plots according to their visual appearance [63]. The metrics used to rank the plots are referred to as scagnostics by Wilkinson et. al. Each point on a cell of the scatter plot matrix represents a scatter plot. Selecting a point on the scatter plot matrix will allow the corresponding plot to be visualized.

2.3 Evaluation Methods

There are several different ways to evaluate visualizations, *A cognitive walk through* is where a person looks at the tasks that a system can perform and steps through them mentally, A heuristic evaluation can be used to measure the effectiveness of the system based on some predefined rules, and user studies can be used to evaluate a system based on testing. Evaluations can be performed on mock-ups rather than complete systems, for example, paper and cardboard prototypes can be used to evaluate the functionality of a user interface [61]. Rapid prototyping techniques can be used for designing systems so that different ideas can be tested. The prototyping for AtmoVis was performed using program code rather than lo-fi modelling, though SVG images were created in an image editor to visualize the look of the interface early in the project. There were three techniques considered for the evaluation of the system: A cognitive walk-through, a heuristic evaluation and a user study. A cognitive walkthrough could be used to evaluate the system by stepping through the process of using the system. A weakness of the the cognitive walkthrough is that it is performed by a programmer so the programmer could still misjudge the requirements of the target audience even after the cognitive walkthrough is complete. During user testing, the test participant could do something that was unexpected or misunderstand the interface.

2.3.1 Visualization User Study

This section discusses methods for conducting a user study. The discussion focuses on the method for evaluation, the participants chosen and the definition of usability.

A review of visualization papers was conducted by Isenberg et. al. in order to determine what evaluation techniques are widely used in the field of visualization. That study grouped research publications into categories based on the evaluation technique and what was being evaluated then the number of publications per category was compared, Some of the evaluation techniques reviewed were based on user studies and others were based on measuring aspects of the system. One of the findings of the research was that a small number of participants are used for user studies in general with a mean of 23.8 participants and a median of 9. Conclusions were developed about how rigorously evaluations are reported and conducted [33].

Usability was defined by Nielsen according to criteria which require that the user should produce only a small number of mistakes when using the system (“errors”), the user should be pleased with their experience of using the system (“satisfaction”), the user should retain knowledge of the system after having used it (“memorability”), the user should be able to use the system rapidly once they have become proficient (“efficiency”) and the system should be “learnable” so that the user can quickly gain the level of proficiency required to operate the system. Nielsen suggests that usability testing can be used to ensure that a system is usable as defined by the criteria and testing can be performed in a laboratory setting. User studies can be carried out with a small sample size or large sample size, There are different ways of conducting user studies for data visualization. User studies can be qualitative or quantitative, qualitative user testing can be performed even with a small number of participants [43].

A previous study evaluated the usability of self-organizing maps, maps and parallel coordinate plots for visualizing geospatial data [36]. The evaluation was based on a usability study on 20 participants and a visualization taxonomy was used to derive the questions from tasks that the participants can perform using different visualizations. The test measured the accuracy of the users' response, the amount of time taken, and the users' response to the visualizations based on questionnaires, interviews and survey responses. The responses to the study tasks were marked based on a binary marking system. The same tasks were used for different visualizations and the results of the task were t-tested to compare the performance of the visualization [36]. The definition of usability was based on the accuracy of the users results and the number of mistakes made during the test ("effectiveness/user performance"), the feedback from the users on their experiences with the system ("user reactions") and whether the system's functionality was what the user expected from the system ("usefulness"), this definition is similar to the definition of usability by Nielsen [43] as both the usability study by Koua et. al and Nielsen used the accuracy of the users' results as a measure of usability and the users' responses to the system in order to measure the usability of the system. Nielsen's definition contains definitions of "memorability" and "learnability" which are not covered by the user study by Koua et. al. and Nielsen did not have separate usability criteria for the users' expectations about the functionality of the system.

Multi-dimensional In-depth Long term Case studies (MILC) can be used to evaluate visualization tools. The method is ethnographic and involves surveying participants, interviewing them and using logs from the visualization system. An advantage of using a MILC study is that the participant is using the tool to perform day to day work so the tasks are more applicable than a laboratory usability test [53]. Participants in a MILC study are observed for a long period of time so MILC is not suitable for the visualization system being built for this thesis.

A previous research experiment has statistically analysed the way that participants use window systems over a three week period using logging software. The users were classified into different categories based on their style of window management [59]. The categories were “piler”, “splatterer”, “maximizer”, and near maximizer. A piler stacks windows on top of each other with a large amount of occlusion. A splatterer stacks windows around the screen with a large proportion of the windows visible. A maximizer has one window maximized on top and the other windows underneath and a near maximizer has one window which takes up most of the screen and leaves some of the desktop visible.

Data visualization tools can be used in an educational setting, a study was performed on the usability of an interactive map for teaching students about climate change [46]. The students were required to answer a variety of questions about the data presented by the map. The study aimed to collect qualitative information so that the interface could be further improved and found that middle school students had difficulty understanding map overlays when completing the study tasks. The task completion and the number of errors were used to evaluate the usability of the interface and the study had a pre-study questionnaire, study tasks and a post-study questionnaire which asked Likert scale questions about the students use of the software. The definition of usability used for evaluating the software was based on Nielsen’s.

The sv3D tool was a study about the usability of a source code analysis system. The participants in the study were all university students and the participants were split into groups, some participants used the visualization tool and others used an Integrated Development Environment. Data was recorded about the speed and accuracy of completing the set tasks then the results were statistically compared, the results found that the participants using sv3D were not as fast to complete the tasks compared to the participants using the IDE, though there was no significant difference between the task correctness for participants in the two groups [41]. Other

user studies have been performed using students as participants. A study was performed on meshes in computer graphics where the effect on quality from reducing the number of faces and vertices was evaluated in a user study on student participants [50]. The study also evaluated whether students believed that carrying out a study was effective for teaching empirical methods [50]. The sv3D tool was evaluated based on the accuracy of the participants' responses to the user study tasks and the length of time taken for the response to be given by the participant. This definition of usability is consistent with part of Nielsen's definition of usability [43] because Nielsen defined "efficiency" and "accuracy" as part of a usability definition.

Studies can be carried out using a mixture of qualitative and quantitative information. A study evaluating the Newdle clustering EVS system advocated using both qualitative and quantitative information when performing user evaluation. The study of the Newdle clustering system also aimed to apply cognitive load theory to the design of the user study so that different types of cognitive load could be measured when using different interfaces [40]. Users were separated into groups and different groups were tested on different systems. The qualitative test questions were categorized according to the type of response given.

2.3.2 Summary

A literature review was conducted identifying and discussing different interface taxonomies, visualization systems and evaluation techniques. The comparison of visualization systems was used to identify functionality that would be useful in a visualization system for air quality data. The most useful techniques and features for visualizations in those systems were reviewed. These techniques and features are relevant to RQ1: "How effective are the visualization techniques in AtmoVis for exploring air quality data?" (Section 1.1).

The following features are used in the visualization prototype system:

“Breadth-first” The Voyager, Trelliscope, and Polaris [30, 57, 65] systems reviewed had a “breadth-first” design principle. Breadth first design is suitable for the prototype system because it allows a large number of sites in the dataset to be explored inspecting a pollution variable first. This approach allows the dataset to be searched for areas with the highest pollution instead of looking at areas first and drilling into the areas to find out how much pollution is at the one location. “The breadth-first” method assists with exploration so that new areas of interest can be discovered based on the pollution level.

“Drill down” The TIARA, Voyager, Visage and Tioga-2 [16, 49, 62, 65] systems had a “drill-down” design principle which allows data points to be filtered and selected based on criteria. Visualizations inside an air quality data visualization system could contain filter criteria for drilling down and sub selecting the data.

“Windowing” Tioga-2 displayed the visualizations in several different related windows. Voyager and Trelliscope displayed different visualizations in facets. An air quality data visualization system which displays the visualizations in different windows and facets would assist with the exploration of the data by allowing visualizations to be related and compared.

“Database Interaction” The RB++, Tioga-2, FacetMap and Polaris [16, 55, 57, 67] systems operate on data from a relational database. Generating queries from a web based front end would assist in managing large amounts of data by retrieving subsets of the data available. The air quality prototype visualization system will use visualizations which query a database to reduce the amount of data loaded when the visualization system starts.

“Visualizations” The calendar heat map visualization was included in Openair [23] and it is effective for visualizing air pollution data and would be useful in a prototype air quality visualization system. Line plots and bar charts were used as part of the Voyager and Trelliscope system for data visualization. Line plots are effective for finding trends and events in temporal data. Voyager and Trelliscope also integrated a recommender engine to find effective visualizations which is outside the scope of this research.

“Taxonomies” Taxonomies allow a system to be evaluated against criteria describing the functionality and tasks that should be achievable by the system. Shneiderman’s [52] taxonomy could be used to evaluate the functionality of a prototype visualization system for air quality before a user study is performed in order to identify potential issues with the effectiveness of the system.

“User Study” The definition of usability by Nielsen [43] provides a method for measuring effectiveness which could be used with either small sample size or large sample size usability testing. A similar method of measuring effectiveness would be useful in conducting a user study on a prototype visualization system. User studies can be qualitative or quantitative and the user study performed by Koua [36] on the usability of self-organizing-maps, maps and parallel coordinate plots contained quantitative and qualitative measures on a sample size of 20 participants. A review of visualization papers conducted by Isenberg et. al [33] indicates that a sample size of 20 participants would be above the median for reported usability studies in the field of visualization. The same sample size and similar methods could be used for measuring effectiveness for the prototype visualization system for air quality data. A window based interface could be used in the prototype system and previous research experiments have been performed on window systems which classify usage into different

categories [59], results could be compared to find out whether more or fewer windows were used than the average. Using both qualitative and quantitative methods produces more feedback information that can be used to improve the system and evaluate the users perspective on their own experiences. The qualitative information and quantitative Likert scales will help to answer RQ1, RQ2: “ How effective is the user experience of AtmoVis for exploring air quality data?”, and marking the responses of study tasks will contribute to RQ3: “How accurate are experts when using the visualizations in AtmoVis?” (Section 1.1).

In Chapter 3: AtmoVis, personas are developed and the design, system goals, architecture, and implementation are described.

Chapter 3

AtmoVis

AtmoVis is a web based system for visualizing atmospheric data. Web-based visualization systems can be used for collaborative work where the tool can be used to display information in an easily digestible way, assist in collaborative work as an explanatory aid between researchers who are on site and researchers who are off site, and assist with the timing of operations involving many people [60]. Though collaboration is outside the scope of the study, web pages can be accessed asynchronously and remotely. This chapter contains the requirements of the AtmoVis visualization system, personas, system goals, system design, architecture, and implementation.

3.1 AtmoVis: Design

AtmoVis is based on other visualization systems for data exploration, for example, Voyager, Trelliscope and Gapminder [3, 30, 65]. The taxonomy developed by Shneiderman, called “the information seeking mantra” [52], inspired the user interaction design process. AtmoVis has a breadth-first interface because the goal of AtmoVis is to allow a large number of variables among different sites to be explored rather than to drill down into the variables of individual sites. Breadth-first interfaces are better for ex-

ploring a large number of variables in a dataset [65]. Figure 3.1 shows the AtmoVis interface with different visualizations open demonstrating an example session.

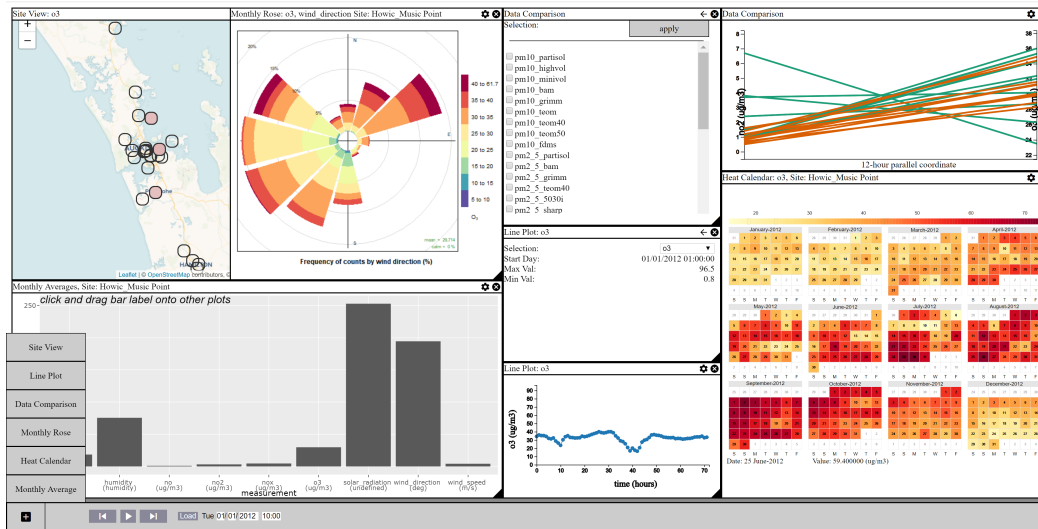


Figure 3.1: The AtmoVis user interface. Map: openstreetmap.org [8]

3.1.1 Personas

A persona is a fictitious person that can help to build an understanding of a group of users and the necessary functionality of the system [43]. Personas can help to build an understanding of a user's goals and requirements and to build empathy with the target group [44]. Several personas can be used to design a system when there are several groups of users with different goals. Using a persona reduces the risk of a software developer building a system targeting themselves rather than the target group [47]. For this project, some preliminary interviews were carried out with two air quality experts from the Ministry For Environment (MFE) and from NIWA to help determine what sorts of systems were already in place and were used to identify the functional requirements of AtmoVis for use in developing personas. A persona consists of a personal description of the

target audience, goals that the persona wants to achieve, and scenarios describing possible uses for the system.

Performing pilot studies on prototypes of AtmoVis and talking to participants allowed the persona scenarios to be produced in a way that more accurately reflected how AtmoVis could be used. Personas can be developed at different stages during a project's development [26]. A persona can be explicitly specified or used implicitly in the mind of a software developer. Research on the target audience should be conducted before producing personas, however, there can be constraints which limit the research which can be conducted before the persona is created [26]. For AtmoVis, interviews were conducted prior to the development of the personas, then the software was developed and pilot studies were conducted while the personas were being refined. There is some contention over whether a persona should represent an individual or a group, however, the persona should not be produced using a stereotype and should be a realistic representation of the individual or the group [44]. To make the persona feel realistic there are three aspects to the construction of a persona suggested by Nielsen [44], the "psyche", the "social background" and the "appearance". The psyche description can contain information about the personality and mindset of the persona. Personality and mindset allow the description to seem more realistic, making the individual easier to empathize with and communicating their motivation. The personality is important so that the persona can be thought of as a real person and can also help to find their goals by studying the mindset and the way that they may want to tackle a problem. People with different social backgrounds can behave differently, there could be educational differences between two personas which affect the way that they would approach a given task. Communicating the level of education can keep functionality in perspective. The visual appearance can convey important background information about the persona [44]. It is not necessary to supply a picture, though pictures can be used, otherwise, the description of appearance can be pro-

vided in the text.

The personas were developed using the programmer's perception of the target audience, based on information gathered in interviews. The information provided by the personas does not demonstrate how a user actually interacts with AtmoVis, so personas are a good technique for system design but they are insufficient to judge whether the design suits the target audience. The following personas were developed to represent the target groups that could use AtmoVis, a data analyst, a climate scientist and a university student.

3.1.2 Kath: Air Quality Scientist Persona

Kath is an air quality scientist. She is 40 years old and has a Bachelor's degree. Kath dresses smartly, she wears tidy black pants, a light coloured shirt and dark work shoes. Kath is tidy at work, likes coffee and sees herself as a casual person. Her motivation for using AtmoVis is to reduce the amount of time taken to produce relevant plots for air pollution. She is interested in technology and has some basic experience using statistics software. Kath has previously used parallel coordinate plots to identify correlation, and wind rose plots to identify spatial patterns, however, it is time-consuming to produce the correct data plots and a more interactive solution would be preferable as an interactive tool could also be used to introduce the dataset to a junior scientist. Kath has access to several datasets on air quality as well as information about the emission sources of air pollution to supplement the use of the system.

Kaths main goals:

Goal 1 To identify spatio-temporal trends in air quality over a given region.

Goal 2 To use air quality visualizations to investigate temporal trends in air quality at a single location.

Goal 3 To communicate air pollution findings with data analysts and policy experts.

Scenario 1 A policy change was implemented which limits the output of SO_2 pollution from a given industrial source. Kath wants to find out whether the change has an effect on directional pollution in a given area.

Scenario 2 Two pollutants are believed to be produced by the same source. Kath wants to find out how related the pollutant levels are in a given region.

3.1.3 Oliver: Data Analyst Persona

Oliver is a data analyst working for a company processing air quality monitoring data. He is 34 years old and has a Bachelor's degree in Statistics. Oliver dresses smartly and sees himself as a tidy person. He is wearing a polo shirt, light coloured pants and brown walking shoes. He is clean shaven and takes some effort over his appearance though he does not follow fashion trends. Oliver sees himself as productive and looks for fast ways to produce understandable results. He uses a programming environment to query data rather than a basic spreadsheet due to the nature of his work. Oliver has experience at programming macros, however, he finds it difficult to identify the right information when using a programming environment or a spreadsheet due to the presentation of the data. Oliver's rationale for using a visual tool is to do some preliminary investigation and to save some time which would otherwise be spent programming. AtmoVis should allow Oliver to rapidly identify areas in New Zealand with the worst pollution and also to find pollutants which may require further research with a programming environment. The interactive nature of AtmoVis could help to build an overall understanding of the data set and to allow unusual aspects of the data to be found more easily.

Oliver's main goals:

Goal 1 Preview up to date information about air quality as soon as it is entered.

Goal 2 Communicate efficiently with scientists who work in the air quality field.

Scenario 1 Oliver wants to find out when a particular region is breaching air quality standards for SO_2 . The software tool should make peak pollution levels visible.

Scenario 2 An environmental policy has changed an emission limit for an industrial source of air pollution. Oliver wants to find out whether the policy has had a meaningful effect on the air pollution of the region. The tool should make aggregate data available for visualization so that differences can be compared.

3.1.4 Mathew: Student Persona

Mathew is a student living in Auckland. He is 20 years old and wears tidy jeans, a polo shirt and walking shoes. He has a neatly trimmed beard and looks after his appearance at university. Mathew is interested in the air quality around where he lives because there is an odour caused by the emissions from an industrial plant. These emissions can aggravate the throat and cause health problems. Mathew is interested in technology but he is not a programmer. He can use basic features in a spreadsheet program but does not make use of macros. Mathew sees himself as community focused and motivated and he wants to use his knowledge of data and statistics to inspire change. He has studied some statistics courses and he is looking for a tool that will allow him to explore the regional air quality data collected and find the most concerning pollutants in the atmosphere which would be suitable for further investigation. A visual tool would allow Mathew to find the information most relevant to air pollution in his region without using a programming environment or macros.

The air quality data could be read into a spreadsheet, however, finding interesting features to comment on is difficult. The tool has an exploratory feel which could reduce the amount of time needed to find interesting features.

Mathew's main goals:

Goal 1 To improve understanding of air quality monitoring.

Goal 2 To access air quality information on the web without contacting an expert.

Goal 3 To use familiar understandable visualizations to view and communicate air quality information.

Scenario 1 Mathew opens up the web page for the visualization tool looking at the Auckland region. He wants to use the visualization to explore the monitoring sites at the surrounding area and find out what pollutants are available to view and analyse, then find a trend over time for a particularly interesting pollutant.

Scenario 2 Mathew's environmental activist group says that a particular building site was creating offensive smells in a region building a subdivision. A new air quality monitoring site was installed by the group. Mathew works remotely and can access the database using the software tool. He wants to find the difference between air quality at a particular site location before and after work started at the site.

3.1.5 AtmoVis: System Goals

The following system goals were developed by analysing persona air quality focused scenarios and goals.

Goal 1 To allow pollutants measured in a region to be discovered.

Goal 2 To allow regions measuring a pollutant to be discovered.

Goal 3 To allow temporal trends for a pollutant to be compared.

Goal 4 To allow spatial trends in air pollution to be compared.

Goal 5 To encourage a breadth first exploration of air quality data and to reduce the barrier for investigating air pollution.

Table 3.1 shows that the functionality required by the persona scenarios is covered by the system goals.

Table 3.1: The relationship between scenarios and system goals.

Persona	System Goal				
	1	2	3	4	5
Kath: Scenario 1				yes	
Kath: Scenario 2			yes	yes	
Mathew: Scenario 1	yes				yes
Mathew: Scenario 2	yes		yes		
Oliver: Scenario 1		yes			
Oliver: Scenario 2			yes	yes	

3.2 Architecture

The system architecture describes how the components of AtmoVis interact with each other at an abstract level. AtmoVis has a front end and a back end so that the system can be used as a web-based visualization resource with the front end in the browser and the back end on the server. The architecture diagram (Figure 3.2) also describes how the air quality data was pre-processed for insertion into the database which is contained within the back-end. The system was designed incrementally using several prototypes and an evaluation review before the final version.

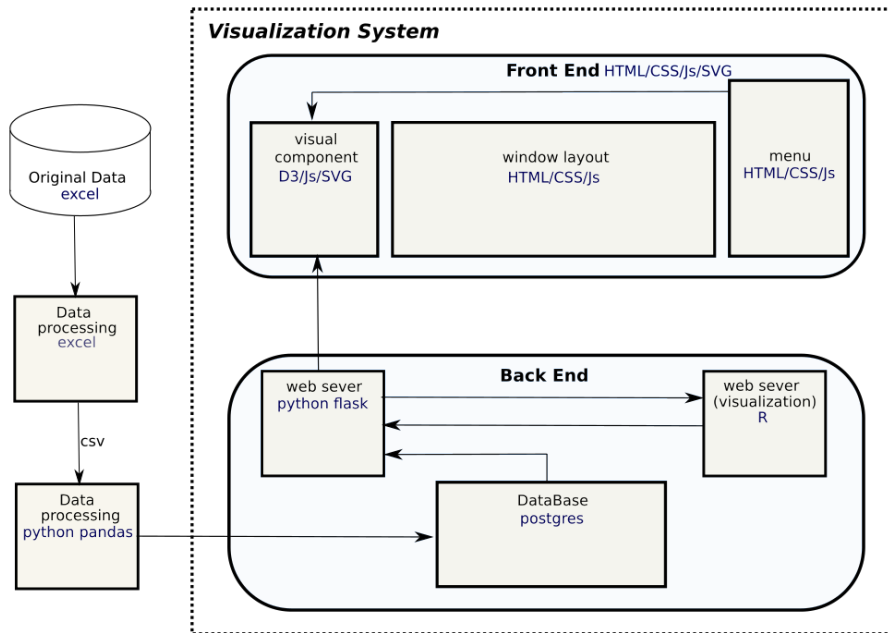


Figure 3.2: Architectural diagram describing AtmoVis.

3.2.1 Front-End

The AtmoVis user interface was implemented in HTTP, CSS, JavaScript, and SVG using D3 and Plotly libraries for drawing visualizations. HTML, CSS, and JavaScript were chosen so that the interface can be built into a web page. Separating the front end from the back end allows the front end to run on a different machine with the back end connected remotely over HTTP. The front end consists of a presentation layer, a business layer and a data layer. The data layer contains front end caches which store information which was retrieved from the server, the business layer communicates data between different visualizations and communicates with the data layer to get the data. The presentation layer builds visualizations and renders them to the screen.

The visualizations communicate with the back-end to retrieve data as necessary so that the system does not need to load a large dataset immediately. Front end caching was needed to reduce the number of messages that were being sent to the server and to manage network latency. The data caching was separated from the visualizations so that different caching strategies could be applied depending on the type of visualization. Some visualizations require the same caching strategy but others do not. The observer design pattern is used to propagate events through the interface to different visualizations.

D3 is an industry standard JavaScript library for producing visualizations that can be rendered in HTML, CSS, and SVG inside a web page [20]. The use of systems like D3 allows for more flexibility in the type of visualization than charting software such as GNU plot and Google plot [19]. Plotly is based on D3 and can produce interactive data visualizations using a declarative JSON based specification language [54]. The front end is responsible for rendering visualizations, displaying the data and implementing user interaction using the D3 library for drawing interactive visualizations [20]. The interface design was based on a window layout so that the mouse could be used for positioning and selecting different visualizations.

3.2.2 Back-End

The system was implemented as a web-based service with data stored on a pre-configured Python Flask server. Flask is a web server microframework for the python programming language [13]. Python was chosen as the programming language for the web server because it is frequently used for data analysis and there are many libraries, such as Pandas, that can be used for data processing [11]. Flask was chosen for its convenience. The back end processes data from a database and sends it to the front end. The data was provided in separate Excel files which needed to be preprocessed

with Excel and Python, before insertion into the database using the Python Pandas library. The database interfaces with flask so that the database is not directly exposed to the front end. This ensures that the Flask server is producing the SQL queries.

R was also used as part of the back end. R is a statistical package which can produce visualizations through the use of different R libraries. The Openair library is an R library which can produce static visualizations of air quality data. Running an R web server using the plumber API [9] allows the visualizations produced by Openair to be sent to the front end. Static visualizations from openair were made interactive in the front end using JavaScript. The Python flask back end communicates with R to request visualizations for the front end to display. The Plotly R library also runs on the R back end and produces interactive plots for display using the plotly JavaScript API [54]. The R back end retrieves data through flask so that there is only one interface to the database.

3.3 Implementation

The following section compares different libraries and programming tools for implementing visualizations, and discusses the choices that were made for the libraries used to implement AtmoVis.

D3 [20], Protovis [19] and VEGA [12] are three systems for building visualizations inside a web browser. D3 and Protovis are embedded DSL's in JavaScript [19, 20]. D3 is implemented as a JavaScript library [20]. VEGA is based on JSON [12]. D3, Protovis, and VEGA have all taken a declarative approach to the specification of a visualization [12, 19, 20]. VEGA's JSON grammar is interpreted by a JavaScript library that can target HTML5 Canvas or SVG for rendering [12, 19, 20]. The JavaScript library can be used to interact with the JSON specified visualization through the use of signals, listeners, and other methods [12]. In D3 the visualization is specified in JavaScript [20]. Selectors are an important part of D3. DOM nodes

can be selected, then data can be bound to nodes. There are methods that select different parts of the binding: an *enter* method, which selects data with no node attached, and an *exit* method, which selects nodes with no data attached. The enter and exit methods are used to assist with the process of updating a visualization [20]. The Protovis visualization can target HTML5, SVG, and Flash for rendering targets [19]. Primitives in the Protovis DSL are called marks which can be Area, Bar, Dot, Wedge, Rule, Link and Label [19]. Panels are used to nest content into more complex visualizations [19]. JavaScript can be used to register event handlers with marks so that user interaction can be added [19]. The VEGA visualization grammar uses signals, data, scales, projections, axes, legends and marks. The signals are used to add interaction from the mouse or from values inserted by the JavaScript runtime. The data specifies how information is loaded into the system for visualization. Data can be filtered, imputed, or adjusted by a variety of operations on being loaded. Scales allow transformations to be applied to the data that specify pixels colours and sizes when the data is drawn. The axes are labelled co-ordinate axes, and legends can be used to indicate scaling, colour and other attributes. The marks indicate how data should be drawn by the system [12]. D3 was the visualization system chosen for implementing AtmoVis because D3 works well with other web technologies and is more customizable than the VEGA grammar.

The Shiny library [25] for R can be used to build interactive web-based interfaces. Also, the Plotly R library can be used to produce interactive R charts [54]. Shiny runs a web server from R and only supports the use of R as a programming environment. However, Plotly provides API's for other programming environments and an account can be produced to host charts and data online [54]. Plotly has a JSON based grammar [54] for specifying plots in a similar way to how plots can be specified in VEGA [12] however the grammar is not required as plotly can convert some charts built in ggplot2 adding interaction. Plotly supports a wide variety of dif-

ferent charts, including histograms, heat maps, contour plots and line charts since plotly makes use of the D3 library underneath, many plot types can be extended using D3. For this project, D3 was used to extend a Plotly histogram plot to add drag and drop functionality for the histogram bar labels. The drag and drop functionality allowed the histogram plot to interact with other plot types by changing the pollutant drawn by the other plots.

3.3.1 Data

Input data was provided as separate Excel spreadsheets which were organized folders with a folder for each region. There was a spreadsheet for each monitoring site that contained hourly measurements and there were also some spreadsheets containing daily averages. The spreadsheets contained two worksheets, a worksheet with metadata and a worksheet with the actual recorded information. The metadata contains information about the monitoring site, for example, the start and finish times for the measurements contained in the data from the site. The data measured was not complete, there were missing values and the time frames that the measurements are recorded over were different for each site. The data was converted to CSV format before processing, Additionally, there was some re-tabulation for several sites in the dataset.

The Pandas API [11] was used to process the input data from the CSV spreadsheets into an SQL database. Additional tables were produced for the units of measurement and the site names. When reading in the data using the Pandas API there were data entry errors. The use of the NZTM_X and NZTM_Y parameters was inconsistent and, there were measurements which were clearly invalid. Extreme outliers were removed and a basic check to ensure that the coordinates were within the expected parameters was performed. There were several different ways of representing coordinates so all coordinates were converted to latitude and longitude before

being read into the database. The Pyproj tool was used to convert the coordinates [10]. The data visualization system needed a map of New Zealand so Carto [24] was used to provide a street map layer for the Leaflet [5] map.

3.4 User Interface

3.4.1 Windowing

In order to implement an exploratory interface for AtmoVis which supports many different visualizations, a way of viewing and adding visualizations to a canvas was required. A windowing system was implemented using D3 which provides a draggable window border. The window border (Figure 3.3 a) can be dragged anywhere on the main canvas however the canvas is subdivided into a grid so that space is used efficiently. Dragging and dropping a window frame snaps the window onto the nearest location on the grid. The windows are resizeable. There is a draggable triangle (Figure 3.3 b) in the corner which changes the window size. The window frame also provides two buttons, a delete button (Figure 3.3 c) and an options panel button (Figure 3.3 d). Clicking on the window border raises the window to the top. The window border contains a text title briefly summarizing the visualization (Figure 3.3 e). For example, the line plot visualization contains a text title summarizing the pollutant which is being drawn. The Google material icons library [4] was used to provide interface icons. Visualizations can provide both a main visualization panel and an options panel accessible through the options button. The delete button removes the visualization from the canvas and is styled with an 'x'. The options panel button is styled with a gear when the visualization is displayed, or a back arrow when the options panel is displayed (Figure 3.3 f). During the pilot testing, the gear icon was replaced with a back arrow when the options panel is visible, leaving the gear icon on the options panel made the layout confusing.

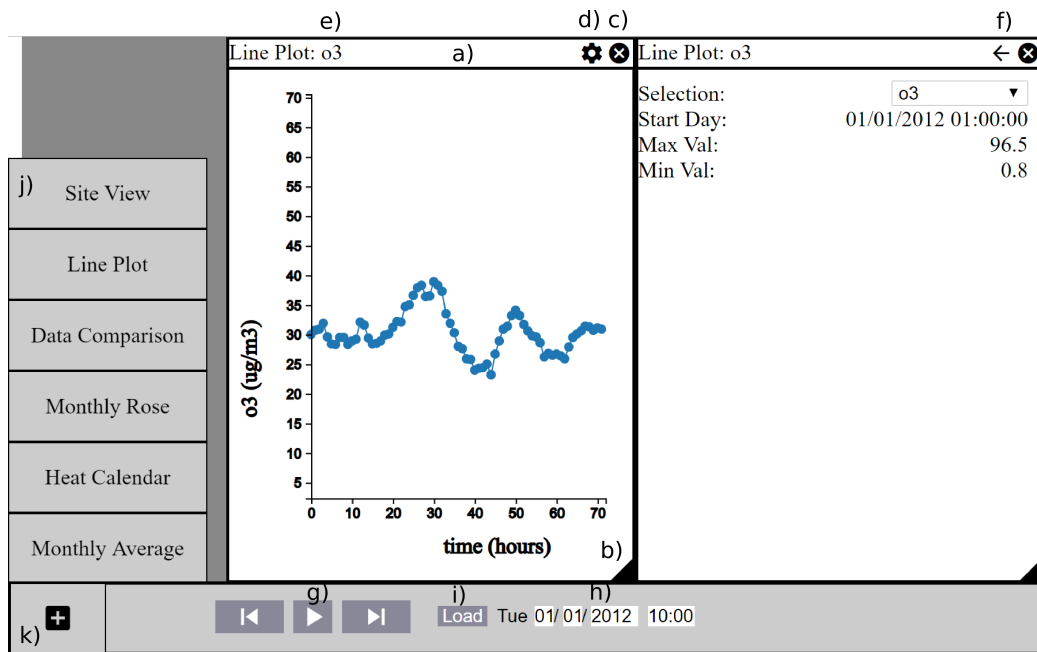


Figure 3.3: AtmoVis interface controls: a) window border, b) window re-size handle, c) delete button, d) options panel button, e) text title, f) options panel button, g) play button, h) time selector, i) load button, j) menu, k) menu button.

There were two options considered for the windowing. One option was to use floating windows and the other was to use tiles. In a tiling layout, the grid can be selected or subdivided into panels and the visualizations occupy the different tiles. The tiles do not necessarily have a top window border with buttons and tiles can be split or merged to change the layout. The windowed option was chosen because windows are more intuitive than tiling and the tiling option would leave some functionality hidden from view. A tiling window layout was developed but not used. The tiling layout allowed regions of the interface to be selected and merged before visualizations were dropped onto the canvas. An additional mode was needed to move around the visualizations.

3.4.2 Control Options

AtmoVis has control options for the time shown on all of the visualizations. The play button (Figure 3.3, g) was provided so that temporal trends in the data can be inspected. The advantage of using a play button is that several plots fixed to the same date can be added to the canvas area and then animated with the play button to show the temporal data. A text area containing the exact date and time is provided (Figure 3.3, h) so that dates and times can be inserted to change the date and time selected. In my own experience, the play button alone was too restrictive because it was difficult to find interesting sections of data without playing right through, so the heat calendar can be used with the play button to make data more explorable. The load button (Figure 3.3, i) is provided so that the data can be reloaded and a notification becomes visible when the data has been reloaded which can confirm the absence of data to a user. A menu is provided (Figure 3.3, j) which contains a list of the visualizations which can be inserted by clicking once, or by clicking and dragging into the desired position on the interface. A menu button (Figure 3.3, k) is provided to hide and show the menu.

3.4.3 Heat Calendar

The heat calendar (Figure 3.4) draws a yearly calendar for a single variable at a monitoring site. The day is coloured according to the mean value of the pollution measurements for that day on the calendar. Hovering the mouse over a day shows the mean value of the pollution on that day in a labeled text field under the calendar. Clicking on the day will change the time slider on the user interface to the start of that day. The month headings can also be clicked. Clicking on a month will change the calendar from a yearly calendar to a single calendar month.

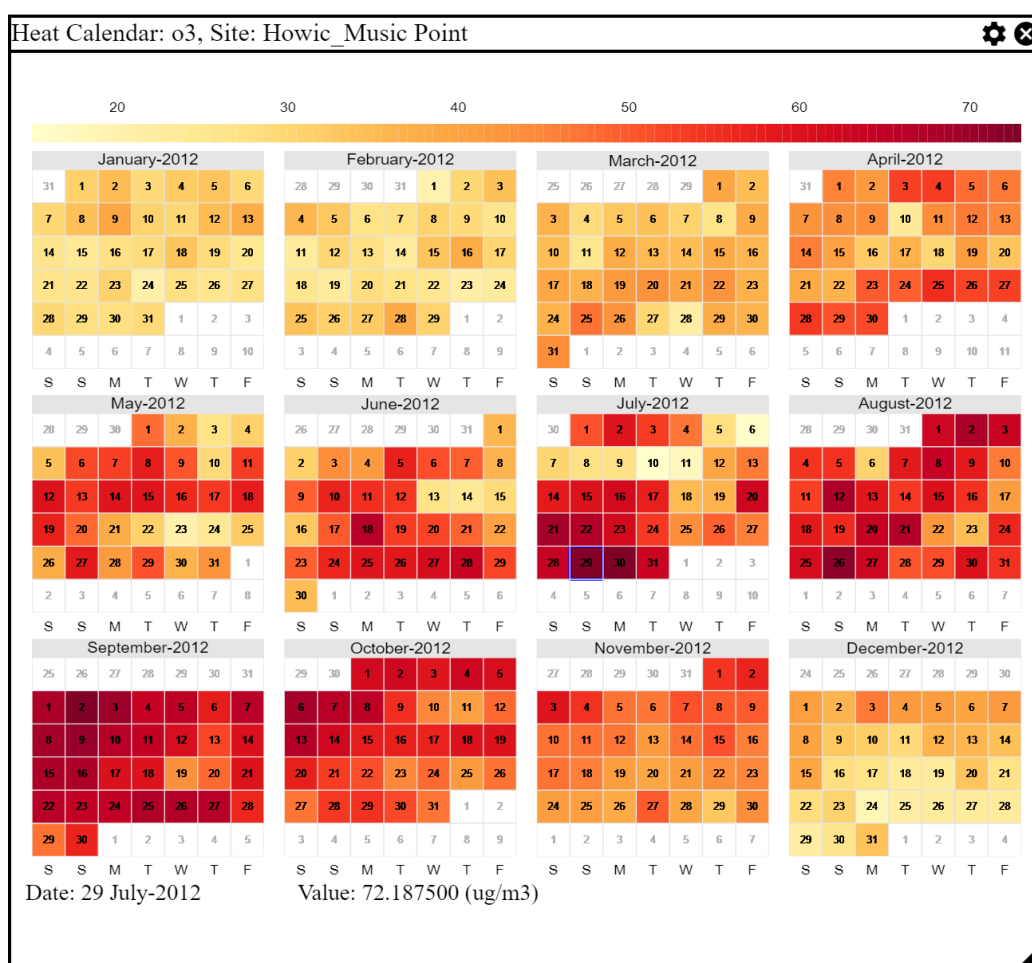


Figure 3.4: A heat calendar with the mouse hovered over a day, the average ozone pollution for that day is shown at the bottom of the screen.

The heat calendar improves the utility of the data visualization tool by allowing days of interest to be found quickly and efficiently. Several calendars can be displayed at once in different windows and clicking on the title of the month will display a monthly view removing information that is not required for the analysis. The heat calendar (Figure 3.4) shows the mean daily O_3 pollution at Musick Point in 2012. The mouse is positioned over 29th of July 2012 and there is a small blue box showing that the day has been selected and the mean value for O_3 on that day is $72.19\mu\text{gm}^{-3}$.

The Openair calendar discussed in Chapter 2 was included in AtmoVis. Openair was chosen as it contains a selection of pre-existing air quality visualizations. However the visualizations produced by Openair are not interactive, so D3 was used to post-process SVG images produced by Openair. D3 provides selectors which can be used on the SVG image output from the Openair library. The heat calendar (Figure 3.4) was produced by extending a calendar generated by Openair with mouse interaction.

There were a few solutions considered for making the heat calendar visualization interactive. The Shiny library allows interactivity to be added to plots, so does the Plotly library, however, D3 was chosen as it is the most flexible and allows any SVG image to be post-processed. The inclusion of the heat calendar required changes to the way that the data was represented on the back end because the averaging operation was found to be too time-consuming for interactive use. The Pandas library was used to perform the averaging of the data and additional tables were added into the postgres relational database containing the daily averages.

3.4.4 Line Plot

The line plot (Figure 3.5) allows air quality data comparison between different sites. Time in hours is plotted along the x-axis and the measured data is plotted along the y-axis. Colour is used to distinguish between monitoring sites. There is a point plotted where each measurement was taken. Hovering over the point brings up a dialogue box containing the time and the value at that point. The options panel is used to change the data which is being displayed. Different pollutants and atmospheric measurements can be selected using the options panel. The line plot can be interacted with through the use of drag and drop. Dragging and dropping a site from a map will add the site to the line plot. When the point is dragged from a map which has a pollutant or meteorological variable displayed, the y-axis variable on the line plot is changed so that the line plot can be configured through the use of drag and drop as well as through the use of the options panel.

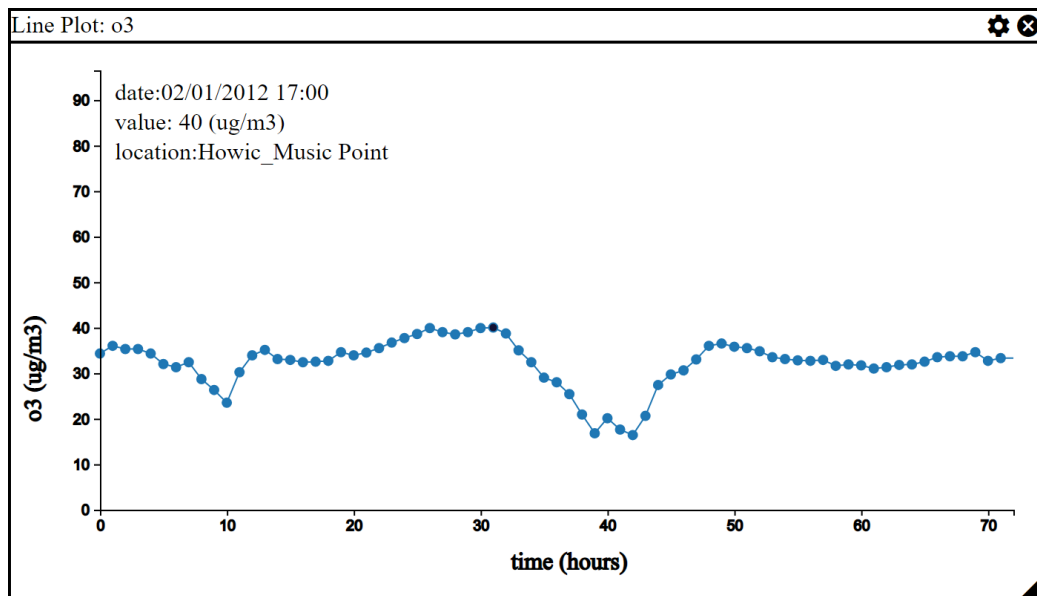


Figure 3.5: A line plot showing ozone measurements at Musick Point with the mouse hovering over a measurement showing that the value of ozone on 02/01/2012 at 17:00 hours was $40\mu\text{gm}^{-3}$.

The line plot displays one y-axis variable at once from several different sites so that a site can be dragged from a map onto the line plot without re-configuring which variables are displayed on the site or on the map which helps speed up the interaction. Several different line plots can be viewed within AtmoVis at the same time so that different variables can be displayed at once. Using one y-axis on each line plot keeps the axis scaling simple and the plot can be zoomed by using the scroll wheel anywhere on the visualization. Using more than one y-axis would have made the zooming more complex. The zoom only scales the y-axis so that the time scales for different line plots are consistent and easy to compare. The line plot was implemented in D3 because it allows HTML and SVG elements to be selected, filtered and updated with a convenient syntax for binding data. D3 also allows interactive mouse functionality to be used.

3.4.5 Monthly Rose

The monthly rose visualization (Figure 3.6) integrates the Openair pollution rose into AtmoVis. The monthly rose plot is not interactive however other charts can be used to control what is being displayed and the monthly rose plot can receive a site dragged and dropped from the site view.

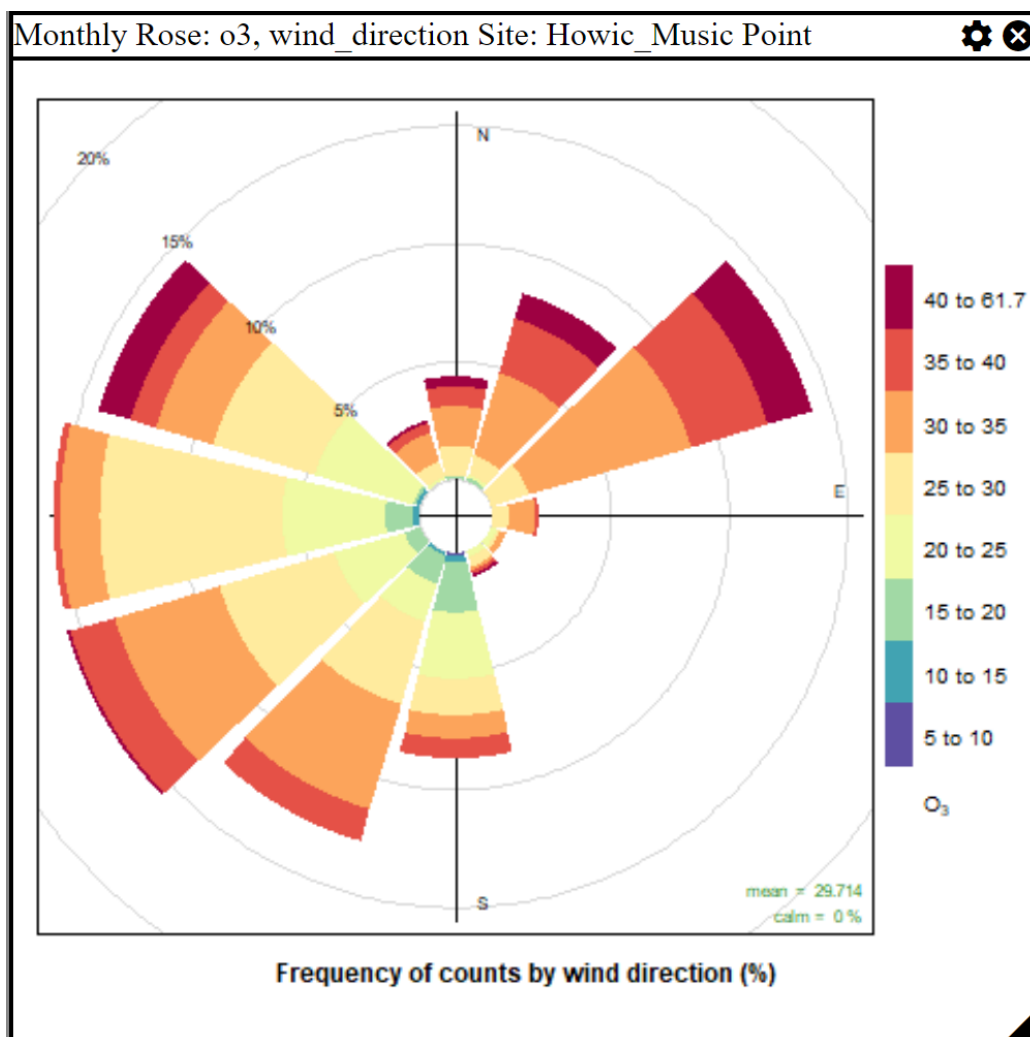


Figure 3.6: The monthly rose diagram shows the concentration of ozone in colour and the frequency of counts by wind direction as the length of each sector from the centre.

When a site is dropped onto the monthly rose the pollutant will be set to match the pollutant of the site view that it was dragged from, if a pollutant is set for the site view. The options panel for the monthly rose is the same as the options panel for the line plot so the pollutant can be reset the same way using the drop-down menu. Using the same options panel ensures that the two plots behave consistently. The monthly rose plot can be controlled from the time slider. When the month changes a new rose will be drawn. In Figure 3.6, a larger proportion of measurements with O_3 between 40 and $61.7\mu\text{gm}^{-3}$ is coming from a northerly direction than from a southerly direction indicating higher O_3 concentrations from the north. Openair runs on an R backend and in order to display the monthly rose plot (Figure 3.6) a PNG image of a pollution rose plot is produced and sent to the flask server to be forwarded to the front end. When the pollution rose plot is received by the front end, it is displayed using the windowing layout implemented as part of AtmoVis.

3.4.6 Site View

The site view (Figure 3.7) has two panels, the main panel containing the map and an options panel. The main panel contains an interactive map which can be navigated using the mouse and shows the position of sites overlaid on the map. The sites displayed are represented as coloured circles with the colour representing the intensity of the air quality data value. Hovering over a site location shows information about the site and the monitoring sites are added to other visualizations by dragging and dropping. The site view is designed to be the only way of adding a monitoring site to another visualization.

The sites displayed are coloured according to the intensity of the air quality data value allowing information about air quality to be read from the map. The options panel for the site allows the pollutant to be selected and shows information about the variable selected.

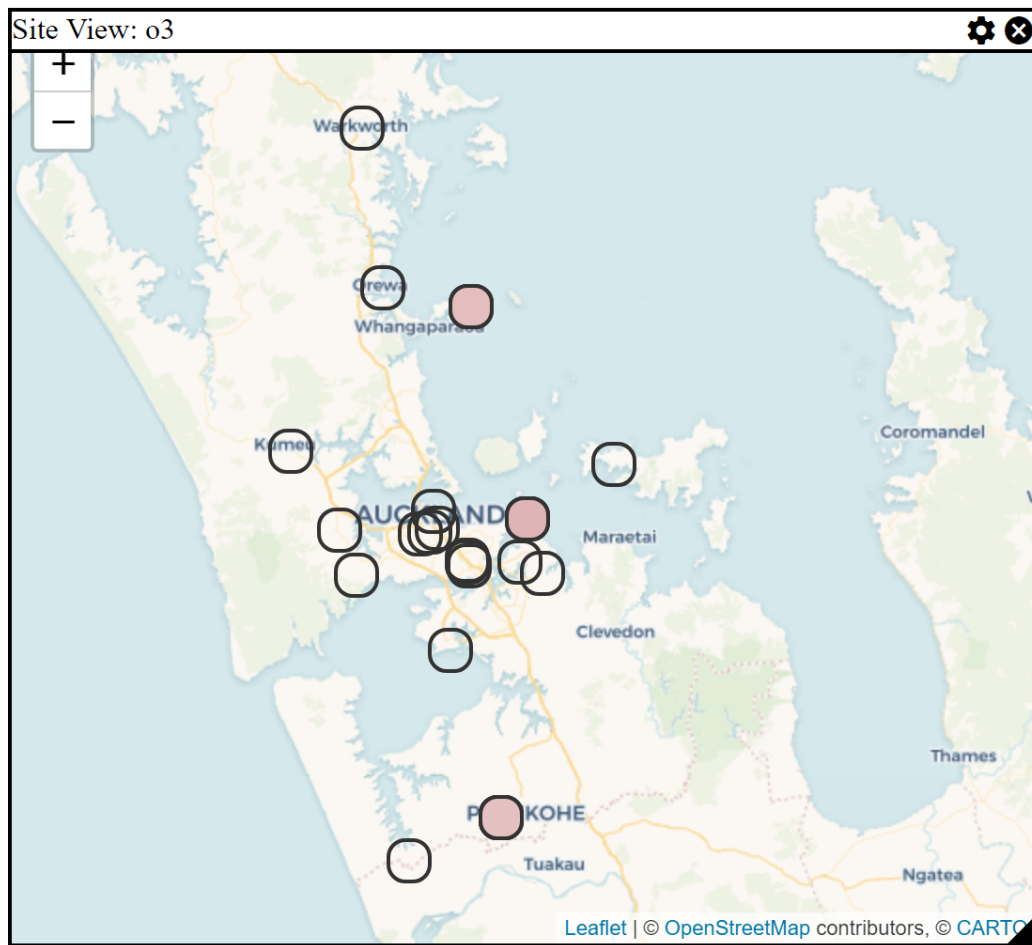


Figure 3.7: The site view centred over Auckland with 19 monitoring sites visible and 3 of the monitoring sites have ozone data available coloured in red. The other sites are clear and have no data available.

The site view (Figure 3.7) was implemented using Leaflet [5] and D3. Leaflet is a JavaScript library for drawing maps, it can draw maps described in the geojson format, draw map tiles provided by third parties such as open street maps, and add positioning icons. Carto [24] was used as a map provider and the map was based on data and images from open street maps. The options panel for the site view was implemented using D3 and allowed the pollutant to be selected. Drag and drop functionality was implemented using the D3 library to add mouse interaction to the site so that the site could be dragged between different visualizations.

3.4.7 Monthly Averages

The monthly averages plot (Figure 3.8) provides a monthly summary of all the data that is recorded at a monitoring site.

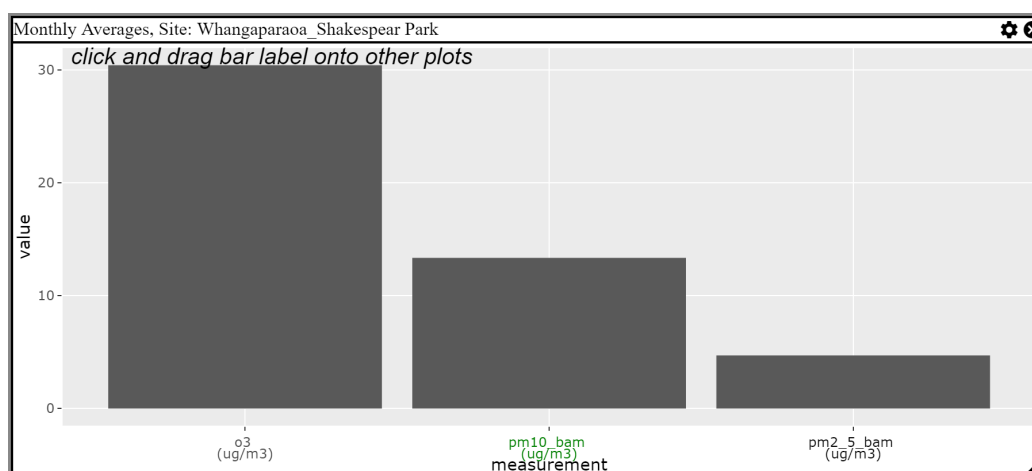


Figure 3.8: The monthly averages for a given site with the mouse hovering over a label, the label can be dragged and dropped to relate the data to other visualizations.

There is also an options panel which allows the data to be filtered. The options panel works in a similar way to the options panel for the data comparison, however, the default settings are different. The monthly averages plot displays a bar graph with the x-axis showing the variable measured by the bar, and the y-axis showing the averaged hourly measurements of the variable indicated on the bar over the month shown on the time selector. The units are displayed on the x-axis because the units for the bars are dependent on the variable being charted. The labels for the bars can be dragged and dropped onto other visualizations open in AtmoVis to reconfigure the visualization to display the variable being dropped. By default, the monthly averages plot will set all applicable data variables to be shown by the plot. This behaviour allows the plot to be read without going into the options to reconfigure. Reconfiguration is only needed to filter the data to a smaller number of variables. In Figure 3.8, the site monitored is in Whangaparaoa and the monthly average levels of ozone, PM_{10} mea-

sured with a BAM and $PM_{2.5}$ measured with a BAM are displayed. The month is January 2012 which is displayed on the time selector (not shown).

The Plotly histogram was used because Plotly has built in mouse interaction which allows for histogram bars to be zoomed for a closer comparison. There were other options considered for the monthly averages plot. One option was to implement the histogram in D3 however more implementation work would have been required as Plotly provides zooming functionality automatically. Plotly is based on the D3 library and plots can be extended using D3 [54]. The functionality of the Plotly histogram was extended to allow the plot to be used as a navigational tool. Mouse interaction was added to the labels of the histogram bars using D3 selectors. The interactive labels allow variables to be dragged and dropped onto other visualizations to reconfigure the data displayed.

3.4.8 Data Comparison

A parallel coordinate plot called the “Data Comparison” visualization (Figure 3.9) is provided to show relationships between numerical variables. There is an axis for each variable in the data which is being compared. A measurement time at a monitoring site is represented as a line which joins together the different axes for every variable measured at that time. A parallel coordinate plot is provided as part of AtmoVis and each site is represented using a different colour. A line for every hour in a 12 hour time frame is displayed on the plot in the appropriate colour allowing relationships among variables to be discovered based on whether the lines are pointing in the same direction or whether they are crossing each other in a pattern.

Figure 3.9 shows 12 hours for sites at Musick Point and Pukekohe in Auckland. The Pukekohe site is in purple and Musick Point is in green. When a line is hovered over details are displayed. The line highlighted in brown on the image has been hovered over and the details identify the

site as located in Pukekohe. The values for the variables measured at the site are displayed, NO_2 is measured at $1.7\mu\text{gm}^{-3}$ and O_3 is measured at $28.3\mu\text{gm}^{-3}$. The measurements were taken on 3/1/2012 at 19:00.

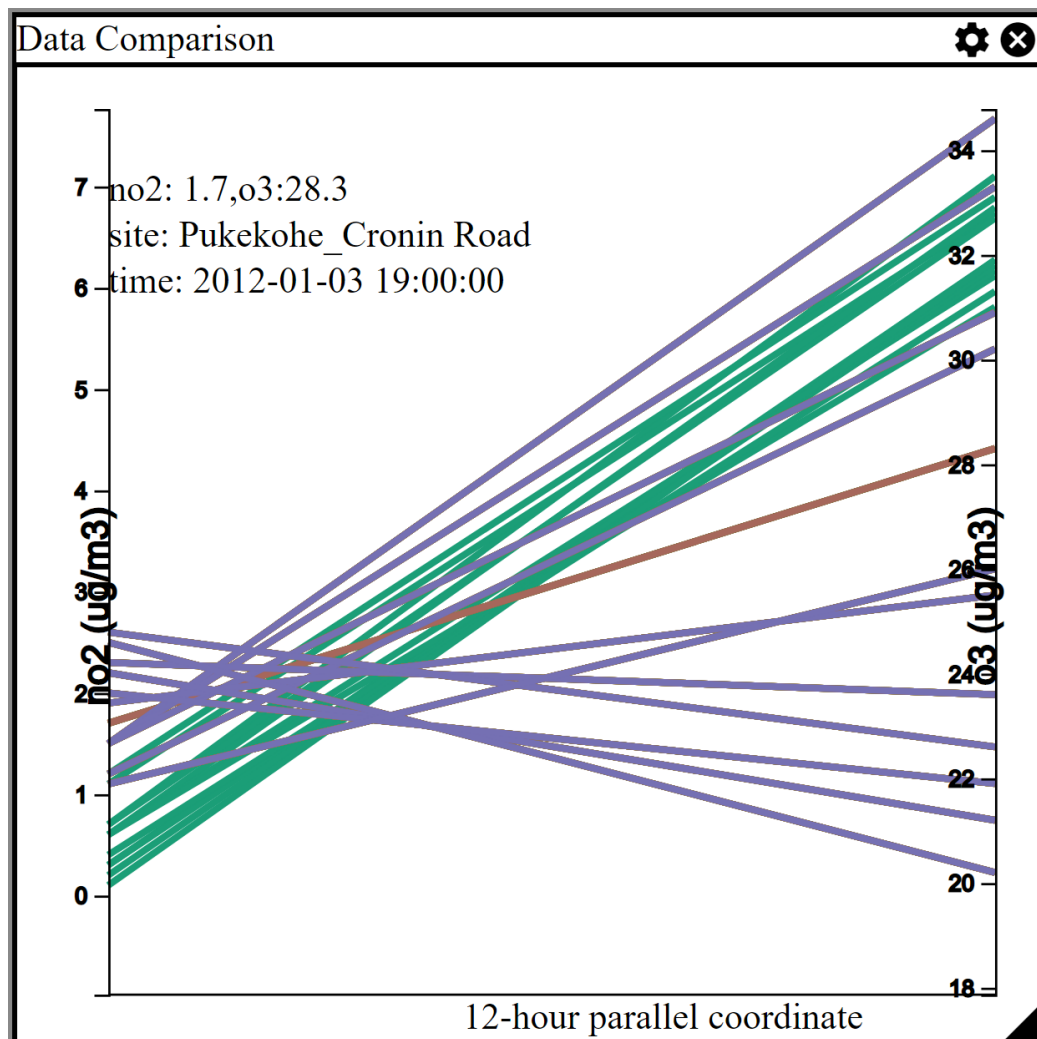


Figure 3.9: A data comparison parallel coordinate plot with the mouse hovering over one measurement to show the details.

D3 was used to implement the parallel coordinate plot as the enter, update, exit model used by D3 is useful for adding the lines to an SVG image. A parallel coordinate plot can be read by first looking at the axis to identify any grouping in the data [35]. Then the colour can be used to

determine any categories that can be grouped. When a colour group is not present it can tell us something about the data as well. Relationships in the data can be identified by looking for lines which are parallel or intersecting to show an inverse relationship.

3.5 Design Analysis

The system was designed around the information seeking mantra [52] described previously. The following description shows how the system has been designed to perform functions described by the taxonomy.

“Overview first” The site view provides an overview of the data and displays the monitoring sites that are available.

“Zoom” Some of the plots allow zooming. The site view can be zoomed, so can the Line Plot, Data Comparison plot and the monthly averages. The study tasks should require users to perform a zoom.

“Filter” Some plots can be filtered using checkboxes, the parallel coordinate plot can use checkboxes to select which coordinate axes to view, the monthly averages can also use checkboxes to restrict the data shown as an alternative to the zoom.

“Details-on-demand” Hovering over different points on the site view, line plot, calendar, parallel coordinate plot and histogram will provide information about the data point in detail.

“Relate” Points from the site view can be dragged and dropped onto other plots as a selection mechanism which can relate the data between the two plots, Also dragging and dropping labels from the monthly averages plot allows the site view to show different pollutant information. The heat calendar relates to different plot types by changing the time selected for every visualization in use. There

is, however, no option to go back through a change history or save layouts.

3.6 Summary

In this chapter the architecture of a web-based system for visualizing air quality data, AtmoVis, was presented and the windowing system was described. Three personas were created to describe scenarios and goals of the target audience. These personas and goals were used to create system goals and the goals were related to the personas to ensure those persona scenarios were covered by the system goals. The data visualized by AtmoVis contains measurements of air quality and meteorological variables.

The visualizations included in AtmoVis are: Site view, line plot, heat calendar, monthly rose, monthly averages, and data comparison.

Site View The site view is designed to show an overview of all the sites with data available and one variable is displayed by shading the colour of each site according to the intensity of the measurement. Sites with no measurement are left clear and sites containing a measurement are coloured.

Line Plot The line plot shows a time series trend for an air quality variable at several different sites.

Heat Calendar The heat calendar displays daily averages for a selected variable inside each date on the calendar with the colour of the date indicating the intensity. The heat calendar is interactive and clicking on a date changes the time selector to that date allowing the heat calendar to be used to discover time frames with high air pollution on other visualizations.

Monthly Rose The monthly rose shows directional trends for a variable over a monthly time frame. The distance of a sector from the centre indicates the number of measurements counted for that wind direction and the colour was shaded according to the intensity of the variable being visualized.

Monthly Averages The monthly averages is a bar graph showing all of the variables available at a given site. The bars are labelled with the variable and the units. The height of the bar indicates the mean of all the measurements for that variable over the month on the time selector.

Data Comparison The data comparison is a parallel coordinate plot that shows the relationship between different variables each displayed on a separate y-axis. The data measured for a site, at a given time, is represented as a line that passes through each axis for every variable measured at that time. The sites are represented by different colours.

In Chapter 4, a user study is presented to evaluate the effectiveness of AtmoVis and the persona scenarios are used to show that the user study tasks are suitable. Use cases are produced using the personas.

Chapter 4

User Study

In this chapter, the user study design, user study participants, data collection methods, and measures of effectiveness are described. It is important to perform a user study to develop and inform the system design process. The objective of this user study was to measure the effectiveness of AtmoVis at presenting information to environmental scientists about New Zealand air quality so that inferences could be made, and to gather the information that can be used to make the system more suitable for the target audience. The results of the user study contribute to answering the research questions (Section 1.1). The personas and system goals were used to design the study and interface taxonomies were used in the analysis of the study tasks to ensure that all aspects of the system design were addressed by the study.

The user study required ethical approval which was obtained from the Victoria University Human Ethics Committee. The ethics approval number was #0000026810.

4.1 Study Participants

11 participants were recruited for the pilot study. After the 11 trials were conducted AtmoVis was ready for testing with air quality experts. The

participants for the pilot study were sourced from both statistical/data analysis backgrounds and computer science/engineering backgrounds. Some of the participants for the pilot study were sourced from within the School of Engineering and Computer Science (ECS), Victoria University of Wellington, and other participants were sourced from outside the university. The pilot study was used to identify some usability problems in the design and to identify any issues in the user study questions. 20 participants were recruited for the main user study. Participants were sourced via email through a GIS expert and also by sending an introductory email to air quality scientists at NIWA, GNS Science, and Regional Councils in New Zealand. An information sheet was provided to all participants (Appendix A). The participants for the main user study were selected experts in air quality analysis, and university students sourced from the School of Geography, Environment and Earth Sciences (SGEES), Victoria University of Wellington, with experience in GIS. AtmoVis design was targeted at users experienced in air quality data analysis.

4.2 Study Procedure

In the following section, the method for study, study tasks, data collection, and analysis methods are described.

4.2.1 Study Methodology

The user study was designed with a pre-study questionnaire, study tasks and a post-study questionnaire [43]. The pre-study questionnaire was used to get information about the participants' background (Appendix B). After the pre-study questionnaire was completed the study tasks were provided to the participants (Appendix C). The post-study questions were provided after the study tasks were complete (Appendix D).

The design of the user study reflected the system goals to ensure that

the study tasks were testing use cases that could occur outside the laboratory settings. The study tasks were based on the goals of the personas which were developed during system design (Section 3.1.1). Personas allowed the goals and requirements of users to be explored in detail through narratives. The persona goals define the sort of tasks that an analyst would want to perform with the system.

4.2.2 Pilot Study

A pilot study was used as part of the requirements gathering process [43]. The results of the pilot study were used to improve the user study questionnaires, tasks and protocol before conducting the main user study. AtmoVis was also refined and adjusted concurrently with pilot study using the results, and the pilot study was performed incrementally. Not all of the participants were using the same prototype of AtmoVis as the system was being improved between pilot tests. The pilot study consisted of a mixture of questionnaires and usability tasks, and was conducted to determine whether each AtmoVis prototype fitted the requirements of the target audience.

Each pilot study participant was given a brief description of the goals of the study, then instructional material for the use of AtmoVis was provided. The instructional material for the pilot study consisted of some instructional videos and a slide show provided on a web page (Appendix E) describing different aspects of the system. The participants were allowed to look at this material before and during the study. The pre-study questionnaire was presented at the same time as the study tasks. Since the pilot study was designed for the improvement of AtmoVis, participants were encouraged to voice their thoughts about the program and the instructional materials using a think aloud protocol [43]. Questions about the material and system were answered and noted. The researcher took notes about the usage of AtmoVis in a log book while the study was be-

ing performed. The screen recording software was trialled in some of the pilot studies. After the study tasks were completed the post-study questionnaire was given. The post-study questionnaire was used to gather information about the participant's experiences. Data responses from the question sheets were tabulated so that statistical analysis techniques for the main user study could be trialled. The study tasks in the pilot determined whether the correct questions were being asked by the study task sheet and whether AtmoVis was suitable for answering these questions.

4.2.3 Main Study

As with the pilot study, participants received a pre-study briefing with some videos and instructional materials and the structure and procedures used in the pilot study were the same. Participants were allowed to ask questions in the study and any questions and answers given were recorded in the log book for further analysis. Participants in the user study could be referred back to statements in the question sheet, instructional slides, or the videos in response to their questions to ensure that answers were given consistently by the researcher. After the tasks were completed a questionnaire was provided that contained ranked quantitative scales and some qualitative questions on the design of the interface and the completion of the tasks (Appendix D).

4.3 Study Tasks

The tasks were structured so that the first questions were short and direct asking the participant to perform actions on the interface and read the results. The reason for this was to introduce the functionality of the interface to the participant. Later questions were more open-ended and asked the participant to use a selection of visualizations to answer a question or to describe the data. The questions were grouped according to different

aspects of data visualizations.

AtmoVis was used by the researcher during the development of the user study to find areas of interest for use in the task construction which helped to ensure that the questions and tasks were manageable for the participants. Reading air quality reports helped to build an understanding of the air quality in the region. Tasks were constructed which required the participant to inspect air quality in Auckland, Masterton and Woolston. AtmoVis was used to find sites in Auckland to inspect for ozone air quality. The effect of pollution on air quality in Masterton and Woolston has been reported [14, 15]. Masterton is a site where particulate matter (PM_{10}) was recorded above the national short term guideline occasionally [14]. Woolston is a site where SO_2 levels have been reported above the national short term guideline occasionally and the SO_2 pollution has been attributed due to industrial sources[14]. Auckland is a location where ozone is measured. Ozone is a pollutant which is formed when other pollutants react in the atmosphere unlike PM_{10} and SO_2 . Ozone was not recorded above the national guidelines over the time frame used in the tasks [14], though the amount of ozone in the atmosphere has been increasing [15]. AtmoVis was used to inspect the data to find interesting and suitable tasks to complete based on persona scenarios.

4.3.1 Task 1: Mapping The Data

In the first section entitled “mapping the data”, the exercise on ozone in Auckland required the participant to view a particular pollutant over time at a site and to find information about a trend over a time frame. The participant was required to insert a site view and find the start date of the O_3 air pollutant, then they were required to insert line plots, read data from line plots and compare two different map views for points which are highlighted on both. The parallel coordinate plot was used to describe a relationship between different pollutants. The first section encourages a

“breadth first” approach to data exploration by introducing the site view as a central plot where the data can be found. The site view allows the data to be compared by location as well as by the selected pollutant level. The exercise of inserting the site view required the participant to pan or zoom the site view in order to find the data point which was situated outside the default bounds of the viewing area. This functionality can demonstrate the “Zoom” aspect of the information seeking mantra [52]. When comparing two line plots the participant must be able to use the window system. There could be four plots on the screen at once starting with the site view which provides an “overview” of all the possible data points. The plots were related by the drag and drop functionality which could be used to set the pollutant on the line plot. In the parallel coordinate plot, the data was initially very close together so the participant needed to zoom in order to fix the scaling. Kath’s scenario 2 (Section 3.1.1) is about finding out how related two pollutants are in a given region, so Kath would be interested in question 3b where the line plot is used to compare two different variables and question 4 where the parallel coordinate plot is used to compare different variables. Mathew’s scenario 1 (Section 3.1.1) required the analysis of time series trends so the line plot would be of interest.

4.3.2 Task 2: Aggregate Data

The second section entitled “aggregate data” required different visualizations showing aggregated data to be compared and analysed to find differences in the pollution between two monitoring sites in Christchurch, and also to compare the visualizations for two different time frames at the same site to see whether any trends are continued. The participant was required to insert a site view, monthly average, calendar and wind rose plots. The calendar was used as a navigational device which allows data to be related by clicking on a day and changing the time for other visualizations. The monthly average plot was also used to relate data by

allowing the label to be dragged and dropped onto the site view. The wind rose plots were compared. The use of the calendar for navigation was also a task that related well to Oliver's persona as the calendar could be used to find averages while comparing a variable over different time frames. Oliver's scenario 1 (Section 3.1.1) required peak pollution levels to be visible for inspecting air quality standard breaches so the use of the calendar would be of interest. Kath's scenario 1 (Section 3.1.1) requires Kath to observe the directional pollution change in an area over time, so looking at the wind roses by clicking on days on the calendar would interest Kath.

4.3.3 Task 3: Parallel Coordinate Data Comparison

In the third section entitled "parallel coordinate data comparison." The use of the site view was the same as for the other two sections, "mapping the data" and "aggregate data". However, the parallel coordinate plot was used to compare data. Also, the parallel coordinate plot was only used to look for patterns in a filtered selection of the data. Due to the range of the data displayed on the parallel coordinate plot, zooming the axes was necessary to interpret whether the data was related or not. Kath's scenario 2 (Section 3.1.1) requires kath to compare different pollutants in a region so the comparison of the data with a parallel coordinate plot would be of interest to Kath.

4.3.4 Task 4: Temporal Pattern

The fourth section entitled "temporal pattern" was based on Mathew's persona. A selection of plots and locations were given for each subtask but the subtasks were very open-ended and there was no information supplied about the pollutants and trends that were being searched for. In question 1a the participant is required to find an interesting site and report on temporal trends for variables at that site using a selection of visualizations, Mathews persona scenario 1 (Section 3.1.1) requires Mathew to find

a temporal trend in an interesting site. In question 1b the site and the pollutant is provided so the question is similar to Mathew's persona scenario 2 (Section 3.1.1) and Oliver's scenario 2 (Section 3.1.1) is also similar as the data is aggregate. Mathew's persona was more exploratory than Oliver's or Kath's. The breadth first nature of the interface could help with finding suitable areas with pollutants. The use of the open-ended tasks was designed to collect information about how different visualizations could be used together to explore the data.

4.4 Data Collection

The pre-study questionnaire collected qualitative and quantitative information on the background of the participants. Collecting background information about the participants allowed differences between groups of participants to be analysed.

The study tasks were split into sections based on aspects of the visualization system which were being studied by the task. The results from the study tasks were marked to determine whether the participant was answering the questions accurately when they were performing the study.

The post-study questionnaire was intended to record information from the participant about different aspects of each visualization as well as general information about their experience with the user interface and the dataset during the study tasks. The post-study questionnaire was split into sections with one section per visualization. Analysing the responses to the post-study and combining the responses with the observational data collected provided information on how effective the interface was.

Screen recording and audio recording were used to collect data about the usage of the system and participants were asked to think aloud. Data was collected on the use of the system because participants may not articulate all aspects of how they use the system, and there may be differences in the way that participants position windows or perform drag and

drop actions which cause some participants to complete the tasks more easily than other participants. These differences may only be visible on the screen recording. The audio recording ensured that responses given at the time of the system usage were not missed. As a participant may not write down everything that they comment on. The audio recording was used to keep a log of anything that was said while the study was being conducted. The screen recording footage was used to make inferences about the use of AtmoVis and how the participants interacted with the visualizations. The note-taking contained times that were of importance on the screen recording footage, comments made by the participant and comments on the participants use of the interface.

4.5 Measuring Effectiveness

The effectiveness of AtmoVis was measured according to the following criteria based on the definition of usability by Nielsen [43]. When a system is usable the number of “errors” produced by the user will be small and the user will have a high level of “satisfaction” [43]. The user study was marked to measure the number of errors participants made, along with observations about the way that participants perform tasks which were used to analyze their use of the system. The user should also have a good experience in using the visualization.

1 Is the user’s response an accurate response?

In Nielsen’s definition of usability, a system is usable when a participant makes few errors [43]. The number of errors can be measured partially by the accuracy of the response, though other errors in the use of the visualization could be picked up by the participant and corrected.

2 Did the user require assistance when performing a given task?

In the event that assistance is required by the user then this does impact the usability of the system as it indicates that the system is difficult to use.

3 Did the user have a good experience using the visualization system?

Nielsen's definition of usability requires that the user should be "satisfied" with their use of the system [43]. The user's response to a visualization was measured by performing a user study and collecting information from the participant's feedback. An effective visualization should produce good feedback from the participant.

4.5.1 Analysis of the Data

The participant responses were analysed with statistical techniques to determine which components of the interface were effective for answering the questions. The ranked questions in each section were added together to produce an aggregate score, ensuring that the higher score was always better. A t-test was applied to the aggregate scores for testing response differences between different groups. There were t-tests performed comparing participants with experience analyzing air quality data to participants without experience, and comparing participants with correct responses to study questions to participants with incorrect responses. Welch's t-test was chosen as the samples are small and the variances are not necessarily the same between the groups being tested.

In the post-study, quantitative questions for each section were ranked, added and averaged to get an overall result for each visualization section. A rank needed to be assigned to the response of each quantitative question to ensure that higher numbers are more positive responses. The scales on assistance with the visualizations were reversed so that little assistance produced a higher score and the questions on text size were removed for consistency. There were also qualitative questions asking for feedback on

the visualizations. The general comments section contained questions on the usability of the design as a whole and which situations it would be suitable for. Qualitative questions are important because they can find new information that the researcher did not know about the participants satisfaction with their experience that may not be discovered from observation [43]. A short discussion was conducted with each participant at the end.

Responses from the audio recordings and open-ended study questions were analysed both qualitatively and quantitatively. Qualitative responses were categorised by topic so that the number of responses addressing particular aspects of a visualization were counted. The results of the user study are presented in Chapter 5.

Chapter 5

Results

In this chapter, the results of the pilot study and the main user study are presented. Section 5.1 describes the changes that were made to AtmoVis as a result of the pilot study. Changes were applied between pilot tests and contributed towards the iterative development of AtmoVis before the main user study was conducted. In Section 5.2 the results of the main user study are presented. The results include quantitative statistics from the pre-study, study tasks, and post-study questionnaires as well as qualitative feedback from the participants.

5.1 Pilot Study

The objective of the pilot study (Section 4.2.2) was to improve the protocol for the main user study, including the questionnaires and documentation, and to find usability issues with AtmoVis before conducting the main user study. The pilot study was conducted incrementally and changes to AtmoVis were applied between studies. This section discusses some of the key changes that were made while the pilot study was being conducted. Each of the 11 participants recruited for the study was presented with a pre-study questionnaire, a post-study questionnaire, study tasks and instructional material. The participants were allowed to ask questions and

their questions were answered and recorded in a log book along with other observations about the participants. Adjustments were made to AtmoVis based on the qualitative feedback from participants and the observations recorded by the researcher in the log book. Table 5.1 shows background information from the participants recruited for the pilot study. The information was provided by the participants as part of the pre-study questionnaire.

Table 5.1: Pilot study participant experience and visualization tools.

PID	Occupation	Air Quality Experience	Visualization software	Spreadsheet
1	Student	no	R	cumulative graphs, trends, timelines
2	Researcher	no	D3, R	yes
3	Student	no		bar, line, pie, histogram
4	Researcher	no	R	lineplot, bar chart
5	Accountant	no		charts
6	Researcher	no	Matlab, latex, R, mathematica	line graphs
7	Researcher	no	R	chart, lineplot, graphs
8	Researcher	no	tableau	histogram, line graph
9	Researcher	no	Pie-charts, tracegraph, R	bargraph, pie
10	Data Analyst Researcher	no	tables, R	line, barcharts
11	Researcher	no		bar, histogram, scatterplots

R is a frequently used visualization tool and 8/11 participants identified themselves as researchers. None of the participants had experience analysing air quality data and spreadsheets are used by all participants for producing visualizations. Line plots are the most frequently reported

visualization with 6/11 participants reporting their use.

5.1.1 Heat Calendar

The heat calendar (Figure 3.4), was adjusted during the pilot study in order to improve the presentation of the heat map by labelling the pollutant displayed as daily averages on the heat calendar.

5.1.2 Line Plot

The line plot (Figure 3.5), was adjusted to improve the readability of the visualization by implementing a zoom function, tweaking the formatting of the options panel text and adjusting font sizes and colours.

Some line plot tasks were difficult for participants due to the range of the data being visualized and the scale presented by the visualization so a zoom was added to the line plot. The zoom was only implemented in the y-direction and the plot still showed the same number of hours on the time axis. The zoom only scales the y-axis so that the time scales for different line plots are consistent and easy to compare. A different time scale would make the scroll speed different for each line plot when the play button is pressed. Zooming is part of the information seeking mantra [52] and allows data to be viewed more easily once zoomed in.

5.1.3 Site View

The site view (Figure 3.7), was adjusted during the pilot study to improve usability by assisting the way that options are found, improving the map so that the surrounding landmarks are visible, and re-formatting text displayed.

Originally the site view would only allow points with data to be dragged and dropped and did not show any information about a site when there was no data for that site present. The site view was adjusted to show the

site name for every site regardless of whether there was data present and to allow sites with no data to be dragged and dropped.

5.1.4 Monthly Averages

The monthly averages visualization (Figure 3.8), was adjusted to make the options panel list more readable and to show the units on each of the averages shown. The options panel used for the monthly averages was adjusted to remove invalid variables. The changes to the monthly averages visualization allowed participants to find which variables were present at a particular site by dragging and dropping a point with no data onto a window containing the visualization.

5.1.5 Data Comparison

The data comparison (Figure 3.9), was improved over the course of the pilot study. Functionality was added to zoom the range on each axis with the mouse, the time range was limited to 12 hours to make the visualization more readable and non-numerical variables were removed from the variable list so that they could not be selected by accident.

Participants were observed loading invalid variables using the options panel. The options panel used for the data comparison was adjusted to remove invalid variables from the list and to prevent the data comparison plot from being viewed between the apply button being pressed and the data being retrieved. The removal of invalid variables was intended to prevent a possible cause of confusion when a non-numerical variable was selected for an axis.

A zoom was added to each axis of the data comparison as some tasks were difficult due to the range of the data being visualized and the scale presented by the visualization. The data comparison visualization also only scaled in the y-direction as the plot could be expanded in size to increase the distance between the axes. The inclusion of the zoom was

intended to make the line plot more readable when very small changes to pollutants were happening over a short time frame as each axis can be zoomed and dragged individually to show a different range of values.

5.1.6 Windowing

The AtmoVis user interface was designed around a collection of floating windows (Figure 3.1). Each visualization occupies a window and participants could click and drag on the window border to move them around. The windowing layout was substantially improved as a result of the pilot study. There were improvements to window movement, window placement, the options panel, play button layout, and the way that visualizations were added to the screen. The improvements were intended to make the floating windows easier to interact with.

Positioning and resizing The windows snapped to a grid. Difficulties with window placement were observed during the pilot study because the window grid was too large. Participants were observed having difficulty placing the windows accurately, so the grid size was reduced making the window movement more precise. Originally each visualization had the same default size and the participants were able to resize the windows during the pilot study to any size required. The window sizes were adjusted so that a separate default size was given to each visualization to reduce the amount of time that participants spend resizing and repositioning the windows. Some restrictions on the window sizes were also found to be necessary in order to prevent buttons from becoming lost or unusable when the windows were resized too small during the pilot study.

Visualizations were inserted onto the screen by drag and drop from a menu (Figure 3.3) onto a screen area and a notification was displayed if the menu was clicked instead. Participants were observed clicking on the menu entries as well as dragging the icons, so the menu functionality was

extended by allowing the menu entry for the visualization to be clicked in order to insert a visualization onto the screen.

During the pilot study, a participant was observed moving one visualization underneath another visualization. The context switching between windows was difficult because the visualizations did not automatically switch to the front when clicked. The windowing was changed so that clicking on a window frame moved that visualization to the top perspective.

The options panel The window frame contained two buttons. One button was for the options panel, the other button was for closing a visualization. The options panel button was styled with a gear icon (Figure 5.1).



Figure 5.1: The window frame border for the visualizations and the options panel. The top image is the frame for the visualization with a highlighted gear icon which the mouse has hovered over. The bottom image is the frame for the options panel with a back button instead of a gear icon.

During the pilot study participants were observed having difficulty changing the option panel back to the site view visualization, so Atmo-Vis was adjusted so that the gear icon was switched to a back button when the options panel was visible. The back button was intended to remind the user how to get back to the visualization and prevent users from having difficulty navigating away from the options panel. The look of the

two buttons was adjusted so that the button icon changes colour on hover. The icon size was decreased to leave more space between the icon and the border of the window frame improving the visual appearance of the icon.

Labels and user interface style changes There was some distance between the time selector and the play button as the time selection was in the right bottom corner and the play button was next to the menu button. Participants were observed moving the mouse a large distance across the screen, so the layout of the play button was adjusted to place the time selection next to the play button.

The window frame only contained the close button and the options panel button and some comments suggested that participants became confused about the data which was being visualized in each window. Additional labels were added to the window borders to distinguish different visualizations which were added to the canvas. The labelling on the window borders was intended to prevent users from becoming confused by misreading the visualization or forgetting the data which the visualization had been applied to.

The options panel used for both the line plot and the site view was adjusted to display the information as a table rather than as a list to improve the readability of the data being presented.

The font sizes were enlarged and colours were adjusted so that the line plot, data comparison, and site view would be more readable. Labels were adjusted on the line plot, data comparison, site view, and heat calendar in order to make the presentation of the data clearer.

5.1.7 Summary

During the pilot study, differences were observed in the way that participants interacted with AtmoVis. There were some different usage styles for the windowing system. It was necessary to design AtmoVis so that there was more than one way to interact with the windowing, so that the

interface suits different usage styles. For example, a participant was observed tiling the windows on the screen so that windows do not overlap. Another participant was observed placing the windows on top of each other and needed to switch the window on top to access the visualizations underneath. A third participant used visualizations as insets to the map, positioning line plots close to the relevant stations.

The pilot study also detected parts of AtmoVis that could be misunderstood by participants due to documentation or the way information was presented by the interface. The labelling of components on the interface was improved to make the interface more usable and zooming was added to both the data comparison and the line plot to make the visualizations more readable by the participants.

Additional documentation was added and updated based on the comments made by participants. The documentation was provided in the form of videos, and a slide show on a web site. There was a number of participants who did not watch the videos first. Documentation provided in a slide show form was found to be more digestible, as these slide shows were provided to the participant through a web browser tab. The documentation used screenshots and pictures of visualizations labelled with the functionality of the components. So that a participant could choose to launch into the study without reading or watching footage then use the slide show as a reference. The test setup was standardized to use a particular laptop to allow screen recording, to avoid test setup issues with equipment and to ensure that the display size and mouse were always the same. When all the adjustments had been made to the interface, documentation, and study protocol the main user study was conducted.

5.2 Main User Study

The main user study was conducted on a different group of participants to the pilot study so that the results of the study would be more relevant to the target group of AtmoVis.

5.2.1 Participants

Table 5.2: Participant dataset experience and visualization tools.

PID	Occupation	Atmospheric Science			Visualization Software	Spreadsheet
		Years	Air Quality Data Ranked 1-5 Scale			
			NIWA	Other		
1	Researcher	no	1	no	R, ArcGIS, ENVI, QGIS	histogram, line, bar, maps
2	GIS Student	no	no	no	R	
3	GIS Student	no	no	no	R, ArcGIS, QGIS	histogram, bar, pivot-chart, maps
4	Researcher	10	3	3	Instant atlas, Tableau, Excel, ArcGIS pro/online	
5	Scientist	15	2	4	R, openair, tmap, base	scatterplots, bar
6	Student, GIS analyst		no	4	R, ArcGIS pro, QGIS	bar, line, scatter
7	Researcher	20	5	5	Arcmap, R, surfer, Excel	scatter, histogram, box+whisker, bar
8	Researcher	25	2	no	R	scatter, column, line, pivot table
9	Data Analyst	4	4	4	plotly, info-gram, R, Tableau, Power BI	
10	Researcher	15	4	4	R, ArcGIS, Quantum, Surfer	line, bar
11	Researcher	30	4	3	Excel	chart, line, histogram
12	GIS Student	no	no	no	R	histogram, scatterplot
13	Researcher	8	2	no	R, Visual cortex	charts, line, wind roses
14	Data Analyst, Researcher	3	5	5	ArcMap, Python, R	line, scatter, bar, histogram
15	Researcher	22	5	5	Matlab, R, Excel, Python, Tableau	pie, bar
16	Researcher	no	no	no	R	Coloured matrices
17	Researcher	25	3	3	R	
18	GIS Analyst	no	no	no	Tableau, ArcGIS, ArcScene	line, pie, time series plots
19	Researcher	4	no	3	R, ArcGIS, AGIS, Excel	histogram, pie, wind rose
20	Data Analyst, Researcher	26	5	5	R, openair	line, column, bar

The 20 participants were chosen so that each of the persona categories, (Section 3.1.1), was present within the sample, and included research students in this area, researchers and air quality scientists from a range of different age categories and experience levels in data analysis. Table 5.2 lists the background of the participants from the main user study. The occupation, experience with atmospheric science, experience with air quality data and visualization software was recorded. The experience with the air quality data was ranked on a 1 to 5 Likert scale by the participant and the number of years experience in atmospheric science was also recorded. The number of years of experience in atmospheric science includes experience with air quality and other atmospheric science. R is also used frequently with 17/20 participants identifying that they had used it. ArcGIS and the related tools ArcMap and ArcScene were also frequently used with 7/20 participants identifying ArcGIS, 1/20 participants using ArcScene and 2/20 participant using ArcMap. The presence of GIS tools such as ArcGIS identifies that more participants were using mapping visualizations than in the pilot study. Line plots were frequently used with 9/20 participants identifying the use of the line plot, 6/20 participants used scatter plots which were proportionally higher than the pilot study and 8/20 participants used bar plots. 13/20 participants reported experience in atmospheric science.

5.3 Visualization Evaluation

This section analyses the feedback given by the participants and observations about how the participants performed during the study. The participants use of AtmoVis was then analysed to evaluate the effectiveness of the visualizations used in the system and to provide suggestions of ways that AtmoVis could be improved to make it more suitable for the day to day tasks of environmental scientists. The analysis of the user study contributes towards the research questions for this thesis (Section 1.1).

5.3.1 Visualization Effectiveness

Figure 5.2 of the box plot shows the distribution of averages in the Likert scale scores for each of the different visualization task sections of the post-study questionnaire (Appendix D). The post-study questionnaire contains different Likert scales in sections with one section for each visualization and a higher mean over that section indicates that a better result was achieved for the overall effectiveness of the visualization as measured by the Likert scale scores.

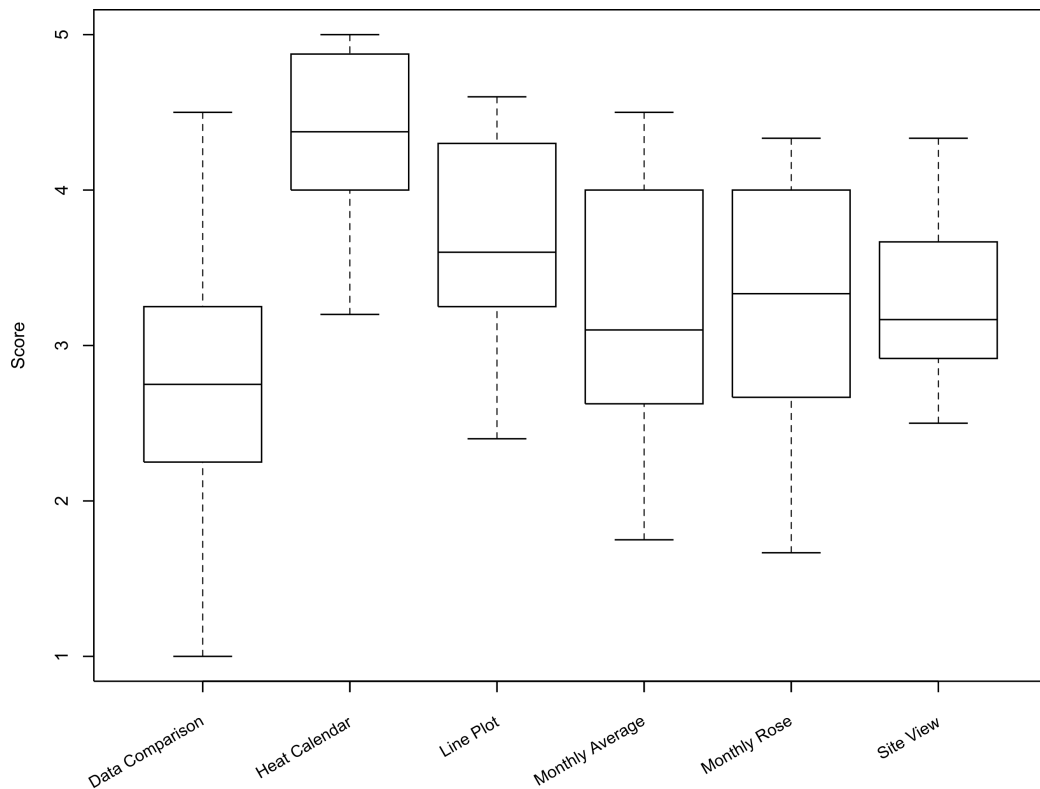


Figure 5.2: Effectiveness of visualization measured by the post-study questionnaire Likert scales aggregated by the method described in Section 5.3.1

The question on the size of each visualization was removed from the Likert scale data to ensure consistency. The scales on assistance with the

visualizations were reversed so that little assistance produced a higher score. Reversing the scale ensures that higher scores for the Likert scales produce higher mean results.

Every visualization except for the data comparison had a good median score, greater than 3, for the mean of the Likert scale scores for each visualization section. The order of visualizations from best median response to the worst is first the heat calendar (4.38), line plot (3.60), monthly rose (3.33), site view (3.17), monthly average (3.10) and finally the data comparison (2.75). The range for the heat calendar (1.80) and the site view (1.83) were smallest respectively showing that there was more consistency in the responses for those visualizations. The interquartile range of the heat calendar was (0.88) and the interquartile range of the site view was (0.75). The data comparison and the monthly average were the most contentious with the widest range of response averages. The interquartile range and the range shown on the box plot diagram displays the range of mean Likert scale responses for participants completing the post-study questionnaire. A larger IQR or range shows that there was a wider range of responses between the participants about the effectiveness of the visualization as described by the post-study questionnaire scales. The results show that the heat calendar was ranked the most effective by participants with a high median score and that all the visualizations except the data comparison were considered effective visualizations.

5.3.2 Heat Calendar

The heat calendar (Figure 3.4) was used to show daily average pollution for each day in the year which is currently on the time selector and it was also used as a navigational device. Clicking on a date on the calendar moves the time selector to that day.

User study tasks The participants were asked which month of 2012 contained the highest daily average level of PM_{10} measured with TEOM at a

temperature of 50°C in Woolston. All 20 participants answered the question correctly using the heat calendar, demonstrating that the heat calendar was effective at displaying the average level of a pollutant over a given day.

Participant Likert scale post-study responses The heat calendar Likert scales were answered by all 20 participants, and the results were positive for all of the questions (Figure 5.3). All participants responded that the calendar was useful for identifying high pollution areas of interest (Figure 5.3a) and 19/20 participants responded that the colour coding was useful in the heat calendar view (Figure 5.3b). 17/20 participants responded that the heat calendar was useful for time navigation (Figure 5.3c). The most contentious question was on text size where most participants responded that the text size was correct but 6 participants responded that the text size was neither effective or ineffective and 4 participants responded that the text size was ineffective (Figure 5.3d). 18/20 participants responded that they did not need much assistance (Figure 5.3e). In the heat calendar, the text scales according to the size of the calendar, so a participant can increase the text size by making the calendar bigger.

Statistical analysis The mean Likert scale score in the heat calendar section of the post-study questionnaire was calculated for each participant based on the method described by Section 5.3.1. Table 5.3 groups the mean Likert scale test scores based on the amount of experience the participants have 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 determined that the two groups "no air quality experience" and "air quality experience" have approximately the same mean ($t(17.54) = 1.04$, $p = .311$, $d = 0.45$). The t-test means that the participant's statement of their experience with air quality data did not significantly alter the participant's Likert scale rankings of their experience

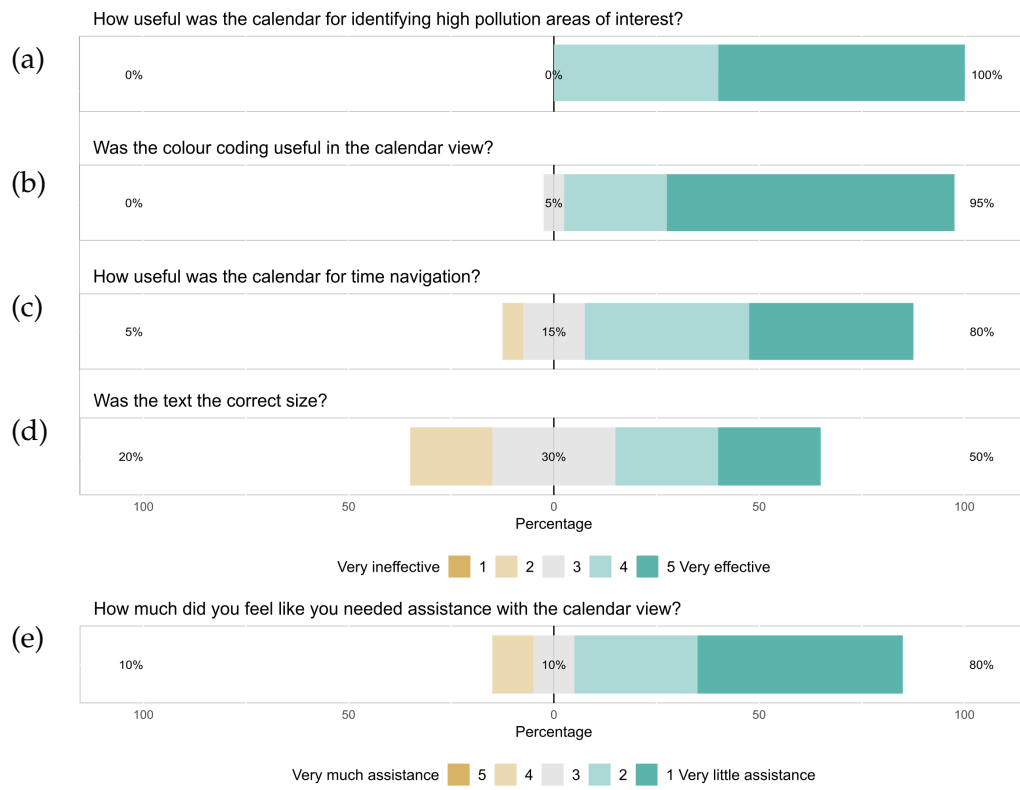


Figure 5.3: Heat Calendar: Likert scales from post-study questionnaire.

with the heat calendar. The heat calendar was evaluated in the post-study based on how effectively the heat calendar was for identifying high pollution, how effective the colour was, how effective the time navigation was and how much assistance was required.

Table 5.3: Summary statistics for the mean heat calendar section Likert score in the post-study questionnaire grouped by experience.

Air Quality Experience					
No Experience			Experience		
Mean	SD	n	Mean	SD	n
4.5313	0.4519	8	4.2875	0.5905	12

Effectiveness of the user experience The heat calendar received some positive participant comments about the colour and the way that the calendar interacts with other visualizations. The heat calendar was used to show trends in the value of a pollutant over a one year time frame (Aggregate data Q1, Q2). The heat calendar was also used as a navigational tool, for example, Aggregate data Q4 part b required the participant to click on the calendar to change the day shown.

Colour The heat calendar colour scheme was based on the highest and lowest pollution for the year of the selected pollutant, and participants were not required to compare the colours of different heat calendars. There were some positive comments relating to the way that the heat calendar shows yearly trends for a variable.

“... the heat calendar is very clear in terms of understanding some important trends there.” - PID 7

The colour scale could be improved by allowing the participant to change what the colours are so that the calendar colour could be set consistently with the monthly rose and other visualizations making the heat calendar easier to compare.

Interaction There were positive responses to the way that the heat calendar allows different days to be selected and interacts with other visualizations.

“I do like this calendar view that’s really nice.” - PID 7
“It automatically responds when you click on the day, I got ya, that’s pretty cool.” - PID 20

The most frequent concern about the heat calendar was that it did not provide additional information about the month currently selected. The

participant, were required to read the time selector to find the current month. 3/18 participants requested some form of a visual cue for the current month.

“... [I] would also welcome indication of what month I am looking at on the monthly views cause when I click on the calendar and move the cursor away that there’s no marker on the calendar left of what’s the month I am looking at.” - PID 15

5.3.3 Line Plot

The line plot (Figure 3.5) allowed a fixed 72-hour time frame to be visualized for several different sites with the same air pollutant or meteorological variable. The 72-hours were marked on the x-axis starting from the origin point at 0 hours. The line plot can show several sites at once on different coloured lines though only one variable can be displayed at once. Hovering over the points on the line plot showed the date and time, the value of the variable, and the site name.

User study tasks The participants were asked to find the date and time of the highest O_3 pollution visible on the line plot (mapping the data question 2a, Appendix C). 19/20 answered the question with a result from the line plot demonstrating that they were able to click and drag the point from the site view onto the line plot to produce a response. 11/20 participants responded with the highest pollution level on the visible section of the line plot, 3/20 participants found the correct day only, 5/20 participants looked forward to a higher pollution level using the play button.

The participants were asked to comment on the relationship between 2 different sites measuring O_3 in Auckland. 18/20 participants answered the question correctly describing the relationship between the O_3 measurements at both sites (mapping the data question 2b, Appendix C) demonstrating that the line plot is effective for trend comparison.

The participants were asked to comment on the relationship between O_3 and Solar Radiation at a single site. The O_3 measurements were on one line plot and the solar radiation was on another line plot. 16/20 participants answered the question correctly and stated that there was a weak relationship or that the peaks were at similar times (mapping the data question 3b, Appendix C). 4/20 participants stated that there was not a clear relationship.

Participant Likert scale post-study responses Likert responses for the line plot (Figure 5.4) were generally positive and demonstrate that the line plot was considered effective temporal data analysis. All 20 participants completed the Likert scales for the line plot. 13/20 participants responded that the line plot was effective for finding temporal patterns (Figure 5.4a). 11/20 participants responded that the play button was effective for finding temporal patterns in the data (Figure 5.4b) and 11/20 participants responded that the use of colour in the line plot was effective for interpreting the data (Figure 5.4c).

The mouse interaction also received generally positive responses with 13/20 participants responding that the mouse was easy or very easy to navigate with (Figure 5.4d) also 12/20 participants stated that they did not require assistance with the line plot (Figure 5.4e). The text size was generally considered appropriate with 13/20 suggesting no change. However, a larger proportion considered the text small than large (Figure 5.4f).

Statistical analysis The mean Likert scale score in the line plot section of the post-study questionnaire was calculated for each participant based on the method described by Section 5.3.1. Table 5.4 groups the mean Likert scale test scores based on the amount of experience the participants have analysing air quality data. The table describes the mean, standard deviation and the number of participants in each group.

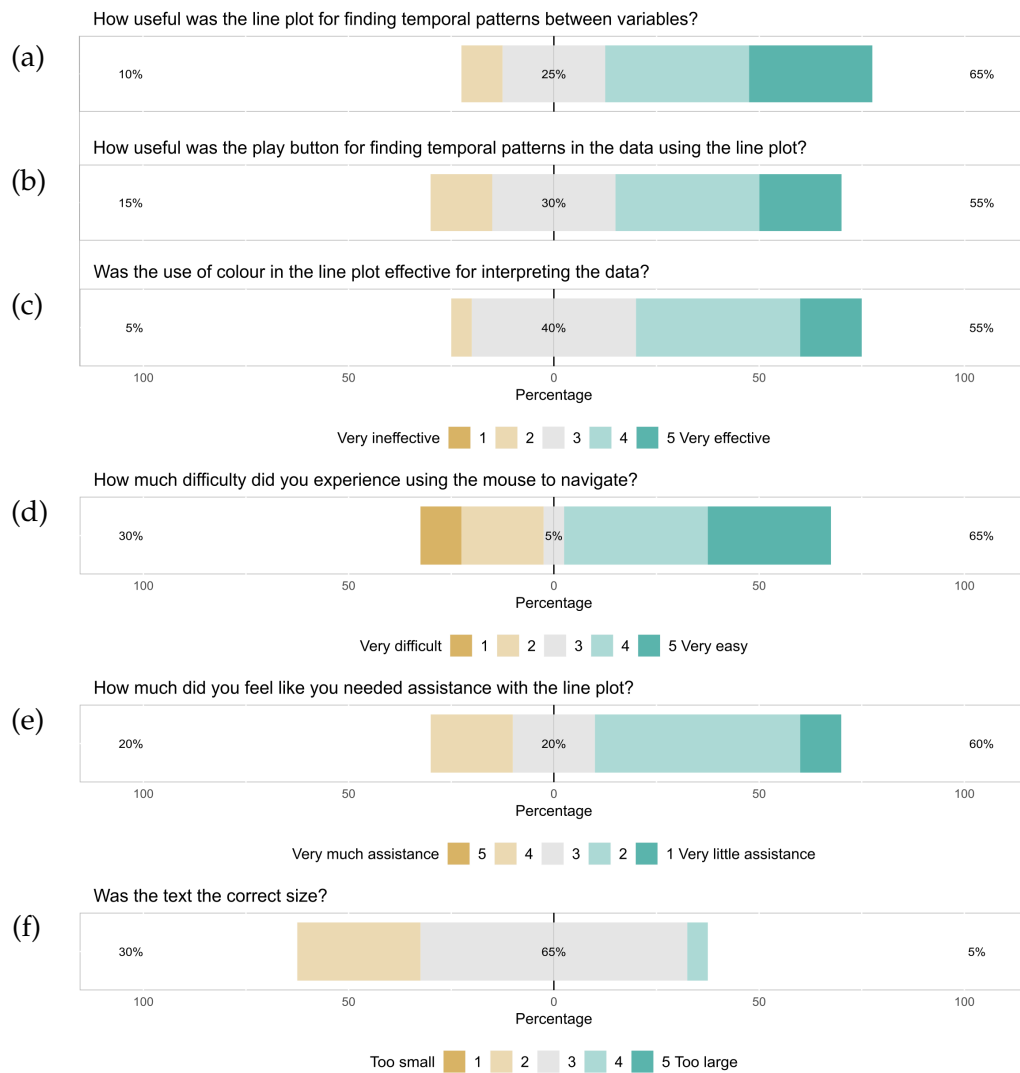


Figure 5.4: Line Plot: Likert scales from post-study questionnaire.

Table 5.4: Summary statistics for the mean line plot section Likert score in the post-study questionnaire grouped by experience.

Air Quality Experience					
No Experience			Experience		
Mean	SD	n	Mean	SD	n
3.9458	0.6794	8	3.4	0.6407	12

Welch's two-sample t-test determined that the two groups "no air quality experience" and "air quality experience" have approximately the same mean ($t(14.51) = 1.80, p = .093, d = 0.83$). The t-test means that 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 line plot. The line plot was evaluated in the post-study based on the effectiveness of the line plot for finding temporal patterns between variables, the effectiveness of using the play button with the line plot for finding temporal patterns, the use of colour, the difficulty of using the mouse with the line plot and how much assistance was required.

Effectiveness of the user experience The most frequently commented aspects of the line plot were about adding additional variables, scaling the y-axis and understanding the starting time. The starting time was confusing to 4 participants because the time at 0 was started from the current position on the time selector and the participants needed to use the interface to notice that the views were all linked to the time selector. Though it was also possible to find the start time by hovering over the first point on the line plot. Two participants suggested adding grid lines and 3 participants tried to access additional summary statistics to read the maximum value from the line plot, which was not available.

Additional Variables The design of the line plot opted to show only 1 variable at a time to avoid comparisons between pollutants with different units but 6/20 participant suggested adding functionality to support more than 1 variable on a line plot.

"If you could put ozone and NO₂ together then play them like that, that would be useful." - PID 11

Interaction 13 participants ranked the mouse usability as positive, showing between 3 and 5 on the Likert scale. The mouse could be used to zoom and shift the y-axis or to hover over individual points to get summary information about the site and the value on the plot.

When a line plot is produced the scale is automatically calculated from the maximum value of the entire dataset, but three participants asked if the line plot could be rescaled. The line plot could be zoomed by using the mouse wheel scroll, the zoom functionality could be improved by adding a visual cue so that it would be easier to see that the line plot is zoomable and also adding buttons to zoom in and out instead of relying on the scroll wheel. One participant commented that the scale made the line plot difficult.

“If the line axis is larger, sometimes the plot is scaled down, and difficult to observe the data.” - PID 14

Though it was possible to find out what the coloured lines on the line plot were by hovering over points, it is not easy to see at a glance if there are a lot of different monitoring sites on the line plot. Hovering over points on the line plot is not always possible because only one point can be hovered over at once so adding an additional key below the axis would improve the readability at a glance and make the line plot more effective.

“Its useful that when you roll over a data point it does tell you the location.” - PID 10

Time Scale Since the date was only displayed on mouse hover 2 participants voiced confusion about the time scale. Changing the scale on the time axis would make the line plot clearer by showing the current date and month with tick marks for hours rather than the time in hours from the current time on the time slide. A visual indicator of night and day would make diurnal patterns easier to infer.

“Well the thing is I don’t know what this bottom scale represents, time hours, I don’t know what this [time is] unless I look at what the start date is up here.” - PID 5

The line plot could also be improved by allowing different time frames to be applied so that longer trends can be shown on the same line plot.

5.3.4 Monthly Rose

During the user study, the participants were required to comment on the relationship between air pollutants and wind direction using the monthly rose visualization (Figure 3.6). The monthly rose uses a compass to graph a measured meteorological variable or pollutant (colour variable) against the wind direction for measurements with the distance from the centre of the compass representing the number of measurements in that direction and the colour of the arc segment representing the mean value of the variable. During the user tasks, the participants were required to read a monthly rose to find trends about the direction and concentration of a single pollutant and to read two monthly roses to compare the concentration and wind direction of a variable.

User study tasks 19/20 participants answered the user task question on wind direction in Woolston (aggregate data section question 2, Appendix C). 5 of these participants required assistance to comment on the monthly roses but were able to read the visualizations after receiving assistance. 15/20 participants found some trend in the wind direction and 9/20 found the relationship between concentration and wind direction correctly.

19/20 participants answered the question about the comparison between Woolston and St. Albans for SO_2 pollution at a given time in aggregate data question 3a. 5/20 participants answered with a correct comparison of both wind direction and pollution level at both locations, 14/20 participants answered the question with partial correct solutions.

“Oh right so north south east west, so this is like the directionality, frequency of counts by wind direction so mostly like east and west got that.” - PID 18

Participant Likert scale post-study responses The Likert scales from the post-study questionnaire demonstrate positive results for the effectiveness of the monthly rose (Figure 5.5) with 14/20 participants stating that the monthly rose was useful for identifying pollutants (Figure 5.5a), and 12/20 participants stating that the monthly rose was effective for finding relationships between data variables (Figure 5.5b). 11/20 participants stated that they needed little or very little assistance with the monthly rose (Figure 5.5c).

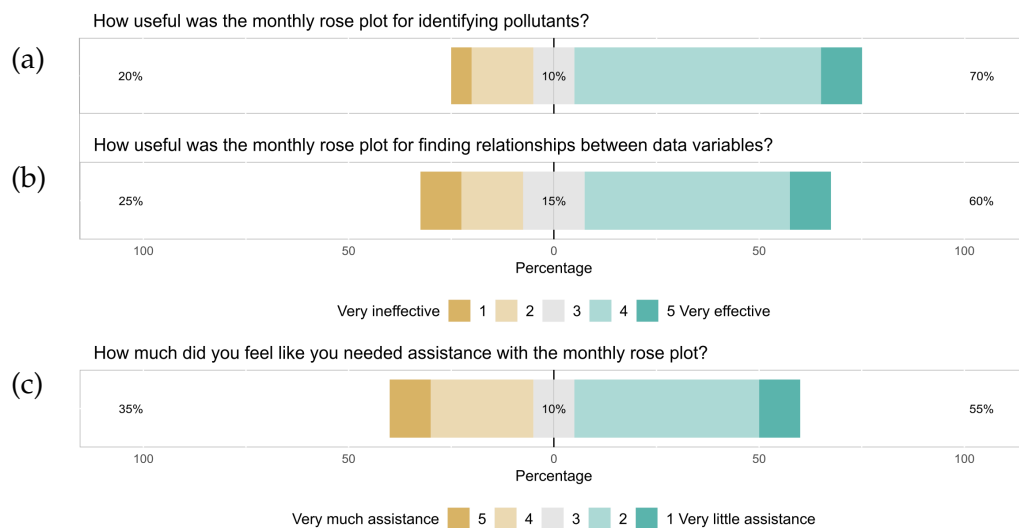


Figure 5.5: Monthly Rose: Likert scales from post-study questionnaire.

Statistical analysis The mean Likert scale score in the monthly rose section of the post-study questionnaire was calculated for each participant based on the method described by Section 5.3.1. Table 5.5 groups the mean Likert scale test scores based on the amount of experience the participants

have analysing air quality data. The table describes the mean, standard deviation and the number of participants in each group.

Table 5.5: Summary statistics for the mean monthly rose section Likert score in the post-study questionnaire grouped by experience.

Air Quality Experience					
No Experience			Experience		
Mean	SD	n	Mean	SD	n
3.3750	0.8626	8	3.1389	0.7972	12

Welch's two-sample t-test determined that the two groups "no air quality experience" and "air quality experience" have approximately the same mean ($t(14.29) = 0.62$, $p = .546$, $d = 0.29$). The t-test means that 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 monthly rose. The monthly rose was evaluated in the post-study based on how effectively pollutants could be identified, how effectively relationships between data variables were found, and how much assistance was needed.

Table 5.6 groups the mean Likert scale test scores based on the correctness of the response to the aggregate data question 2. The table describes the mean, standard deviation and the number of participants in each group.

Table 5.6: Summary statistics for the mean monthly rose section Likert score in the post-study questionnaire grouped by response accuracy.

Question Accuracy					
Incorrect			Correct		
Mean	SD	n	Mean	SD	n
3.733333	0.5477226	5	3.066667	0.8280787	15

Welch’s two-sample t-test determined that the two groups “incorrect” and “correct” have approximately the same mean ($t(10.65) = -2.05, p = .066, d = -0.86$). Participants answering aggregate data question 2 correctly did not give measurably different feedback on the post-study questionnaire Likert scales compared to participants answering the question incorrectly.

Effectiveness of the user experience

The comments recorded about the monthly rose refer to the scale categories, the colour scheme, the time frames chosen for the analysis and the interaction with the mouse. 4/20 participants commented that the monthly rose was useful for analysing the data and 2 participants stated that they would like to use AtmoVis to access the monthly rose. Some participants were not able to correctly read the monthly rose plot concentrations due to their unfamiliarity with the other wind rose visualizations and difficulties with the scale. 4/20 participants reported being unsure of what information the monthly rose was being displayed to them them.

Scale Categories When completing the monthly rose at Woolston the colour scale generated by default was effective for a single monthly rose and participants managed to complete the question. 2/20 participants did comment about the interval between the scales for a single rose but other

participants were able to complete the question without difficulty. Based on the participant responses for the marking of the monthly rose question on Woolston the monthly rose was still effective for describing the trend for a single variable in wind direction and concentration.

When completing the monthly rose questions which required more than one rose to be compared there were some difficulties associated with the scale. The default scale for the colour variable did not place the same distance between categories in the same visualization. 6/20 participants commented on the categories used for the monthly rose colour scale. The distance between colour categories in different visualizations did not necessarily match as the scales were calculated independently, which made the visualizations more difficult to read and compare.

“So the number of categories in each are different so you can’t do a direct colour comparison.” - PID 1

The distance between the percentage lines from the centre was also confusing to some participants, who found the categories difficult to compare when trying to compare the percentage as well.

*“So they [percentage markings] are not the same, in the same place. ”
- PID 19*

The monthly rose could be improved by adding functionality for setting the categories on each monthly rose individually or for locking the categories together for different monthly roses.

Interaction The interaction between the monthly rose and the heat calendar received some positive responses. Using the heat calendar to change the pollutant on a monthly rose visualization is faster than using an R script. Two participants stated that they would like to use AtmoVis in order 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 script itself.” - PID 20

The monthly rose does not identify what month is currently selected on the time slider.

“I can’t remember what day I clicked on, I knew it was a day in February so it would be quite good if the wind rose could say what day it was, or what time or what month or what period so you know what you’re looking at.” - PID 5

Labelling the current month on the monthly rose would make the visualizations more presentable and effective if the functionality was added to extract visualizations for insertion into documents.

5.3.5 Site View

The site view displayed a map of New Zealand with the location of the sites on the map (Figure 3.7). Each site was displayed as coloured circles and tooltip information was provided to participants when they hovered over a site. The circles were filled with colour when air quality data was present and transparent when air quality data was not present. During the user study, participants were asked to use the site view to set the monitoring site for other visualizations by dragging and dropping the circles, and also to read information about the dataset from the site view options panel. The monthly averages can be used to change the pollutant as well as the options panel.

User study tasks The participants were required to read the site view to find how many locations had filled in colour when the ozone was first selected by the options panel (mapping the data question 1b, Appendix

C), the locations were not visible on the default scaling. 16/20 participants answered correctly demonstrating that interesting sites can be identified based on colour value. Though four participants gave responses that were based on the starting location and did not pan or zoom the site view in order to discover the other sites.

Participant Likert scale post-study responses Responses to the site view Likert scale questions were generally positive (Figure 5.6) with 14/20 participants indicating that the mouse was easy or very easy to use for navigation (Figure 5.6e), suggesting that participants were able to navigate and use the map to access sites.

“Right so you can change what the site is viewing based on the labels, ok cool, that’s cool.” - PID 18

The participant responses to the questions about colour change (Figure 5.6a, 5.6c) and the options panel (Figure 5.6b), indicate that the site view can be effective for temporal analysis at a single site. 9/20 participants reported that the information displayed in the options panel was effective (Figure 5.6b). 18/20 participants responded to the Likert scale question asking whether the colour change was meaningful for identifying temporal patterns and 8/18 indicated effective or very effective (Figure 5.6a). 9/20 participants indicated that the use of colour on the map was effective for representing the data collected at each site (Figure 5.6c). Though the site view was effective for temporal analysis, 9/20 participants identified that the site view was ineffective or very ineffective for identifying spatial relationships of variables displayed on the map (Figure 5.6d). Discussions with one participant indicated that the air quality monitoring sites displayed on the site view were too far apart to be comparable which would explain why the site view was believed to be ineffective for identifying the spatial relationships of the pollutant by participant 11.

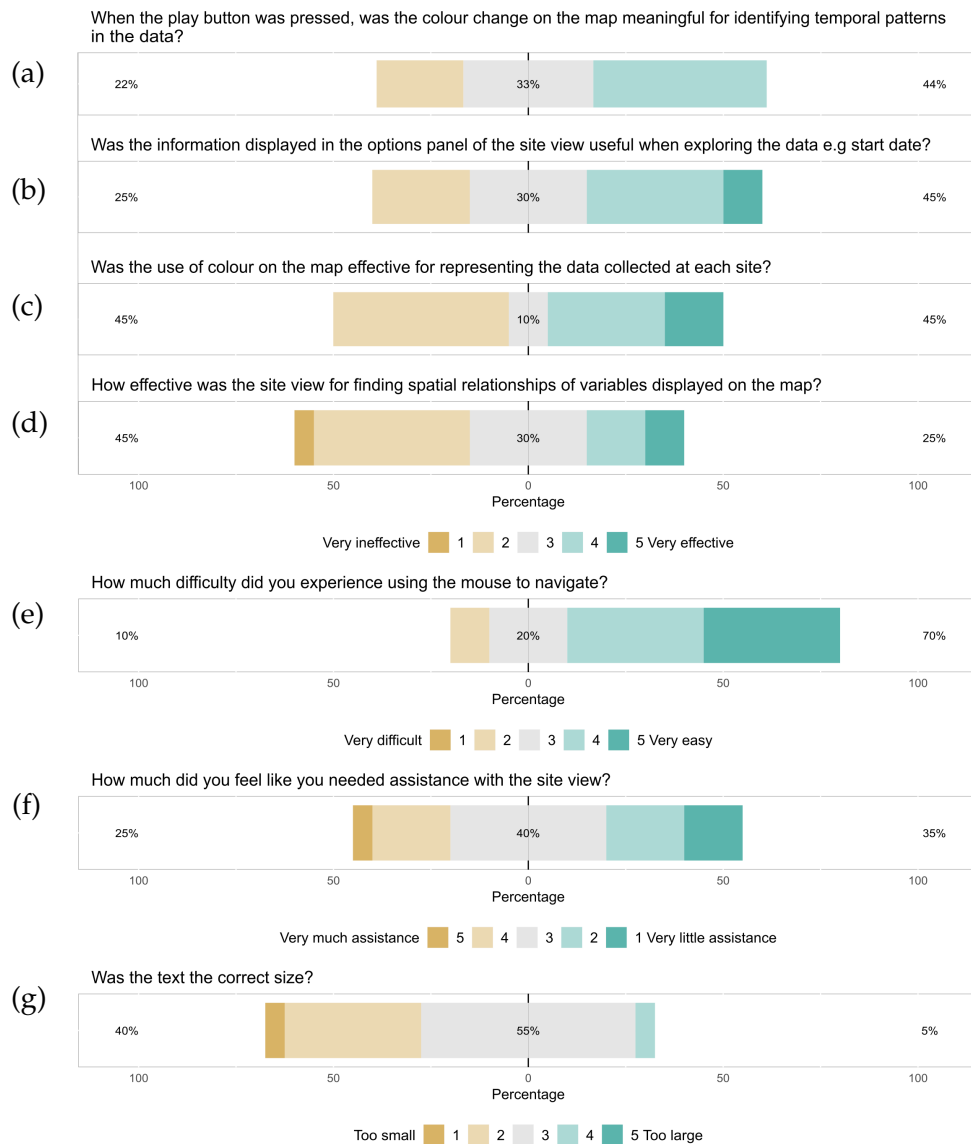


Figure 5.6: Site View: Likert scales from post-study questionnaire.

“The thing about air quality in general and especially in New Zealand is that it’s very highly localized so these places are 20 km apart so the air quality at the two sites is completely unrelated to each other so you wouldn’t expect to see anything ... With this kind of data set is what you’ve got is little islands of data which are not necessarily joined up.” - PID 11

Participants generally felt that they did not need much assistance, 7/20 participants felt that they did not need assistance with the site view (Figure 5.6f) while 8/20 felt that they neither needed little assistance nor much assistance. The use of colour was generally effective, 8/20 indicated that the text size was small (Figure 5.6g). The colour and the text on the site view could be improved given more time and further user studies.

Statistical analysis The mean Likert scale score in the site view section of the post-study questionnaire was calculated for each participant based on the method described by Section 5.3.1. Table 5.7 groups the mean Likert scale test scores based on the amount of experience the participants have analysing air quality data. The table describes the mean, standard deviation, and the number of participants in each group.

Table 5.7: Summary statistics for the mean site view section Likert score in the post-study questionnaire grouped by experience.

Air Quality Experience					
No Experience			Experience		
Mean	SD	n	Mean	SD	n
3.3333	0.5419	8	3.2063	0.5610	12

Welch's two-sample t-test determined that the two groups "no air quality experience" and "air quality experience" have approximately the same mean ($t(15.53) = 0.51, p = .620, d = 0.23$). The t-test means that 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 site view. The site view was evaluated in the post-study based on colour changes, the information displayed in the options panel, the colour

on the map, the spatial relationships displayed on the map, the use of the mouse to navigate, and the amount of assistance required.

Effectiveness of the user experience The comments recorded about the site view refer to the colour and the difficulty of finding site locations. The results demonstrate that the sites could be found easily when a site location was known (Figure 5.6) but when a site location was not known, navigation was more difficult. Also the colour could be hard to read when pollution levels are low because the pale colour is very similar to no colour at all which was used to indicate no value at the site.

Colour 6/20 participants discussed the colour of the site view. The colour scheme ranged from white to red and was scaled based on the maximum recorded measurement for the pollutant selected for the site view. One suggestion was to make the colour better by removing white from the colour scheme, as the base map was very pale in colour so it was sometimes difficult to distinguish the colourless sites from the white sites.

“There were some parts where those station colours were quite similar in shade to the base map and that made it hard to know in some cases if the station actually had data or whether that was absent.” - PID 3

“First of all this keeps changing I understand why it is, because it’s diurnal, ok, but it goes white at night time, ok so it looks like any other site, so it needs to go black or grey or something like that.” - PID 8

Finding site locations 5/20 participants discussed the difficulty of finding sites which were specified by the task instructions. One solution suggested was to add a search feature so that the site could be typed into a search box and the map can be zoomed to that site.

"Yeah I was going to say a search would be kind of good." - PID 18

The site view was good at providing access to the sites by mouse navigation providing that the site location was known.

"You have to know your geography." - PID 5

5.3.6 Monthly Averages

The monthly averages visualization (Figure 3.8) was used to set the pollutant on the site view. The monthly averages visualization could be used to show the averages of each pollutant available during the month on the time selector for the monitoring site being visualized. The monthly averages visualization also allowed the labels on the bars to be dragged and dropped onto other visualizations to set the variables displayed on those visualizations.

User study tasks 17 participants were able to use the monthly averages to drag and drop a pollutant onto a monthly rose (aggregate data 4a, Appendix C) so the monthly averages was effective for displaying the pollutants at a given site.

Participant Likert scale post-study responses The participant responses to the monthly averages visualization were generally positive (Figure 5.7). 12/19 participants responded that the monthly averages visualization was effective for changing variables on the site view (Figure 5.7a) and 9/20 participants responded that the monthly average plot was effective for identifying pollutants of interest (Figure 5.7b). 9/20 participants responded that the text size was effective (Figure 5.7c). 10/19 participants responded that the monthly averages was ineffective for finding relationships between pollutants (Figure 5.7d).

The monthly average plot was easy to use, 11/20 participants, so responded that they required little or very little assistance with the monthly averages (Figure 5.7e).

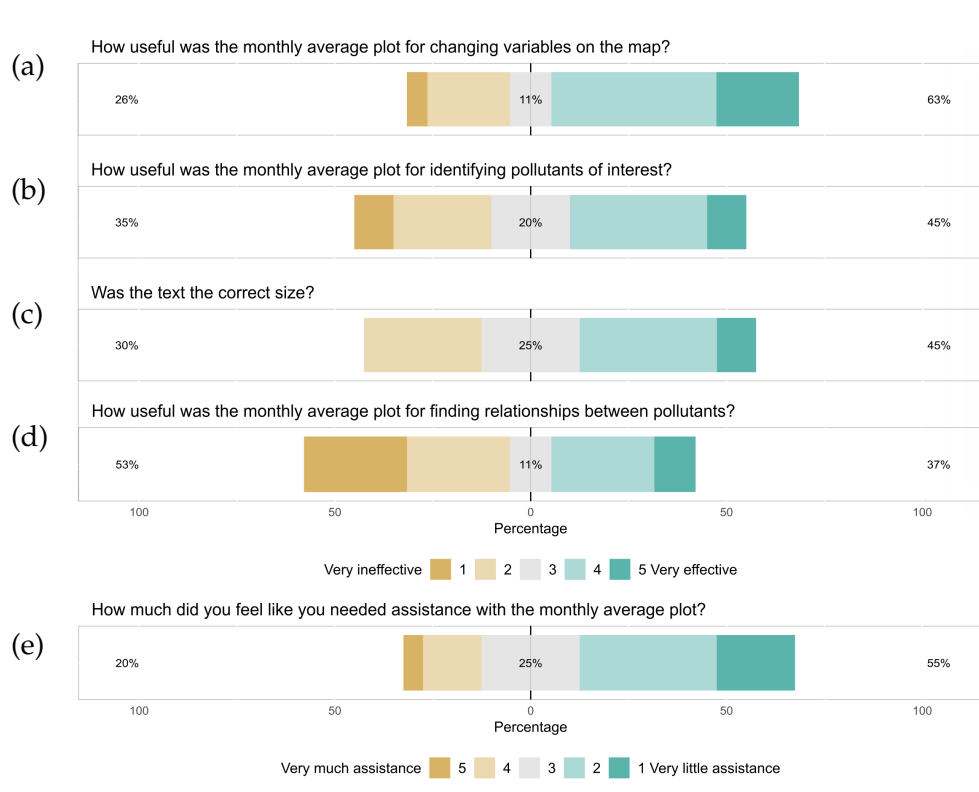


Figure 5.7: Monthly Averages: Likert scales from post-study questionnaire.

Statistical analysis The mean Likert scale score in the monthly average section of the post-study questionnaire was calculated for each participant based on the method described by Section 5.3.1. Table 5.8 groups the mean Likert scale test scores based on the amount of experience the participants have analysing air quality data. The table describes the mean, standard deviation, and the number of participants in each group.

Table 5.8: Summary statistics for the mean monthly average section Likert score in the post-study questionnaire grouped by experience.

Air Quality Experience					
No Experience			Experience		
Mean	SD	n	Mean	SD	n
3.8750	0.6682	8	2.7250	0.6995	12

Welch's two-sample t-test determined that the two groups "no air quality experience" and "air quality experience" do not have the same mean ($t(15.65) = 3.70, p = .002, d = 1.67$). The t-test means that the participant's statement of their experience with air quality data did significantly alter the participant's Likert scale rankings of their experience with the monthly average plot and participants with no experience in air quality data analysis were more positive towards the use of the visualization. The monthly average visualization was evaluated in the post-study based on the effectiveness of the visualization for changing variables on the map, the effectiveness of the visualization for identifying pollutants of interest, the effectiveness of the visualization for finding relationships between pollutants, and how much assistance was needed.

Effectiveness of the user experience The monthly averages visualization received comments on the way that the user can interact with the visualization, and the way that the scales are calculated and presented. A strength of the monthly averages was the way that the visualization could be used to interact with the site view by dragging and dropping a variable label.

Interaction The monthly averages visualization could be used to display the mean value of each pollutant at a given site over the month which contains the date on the time selector and the monthly averages visualization provides an interactive way to change the variable being displayed by other visualizations and to list what variables were available at a given site. There were some positive responses to the interactive way that the monthly averages could be used to change the variable displayed by other visualizations. One participant responded with the following comment when they dragged and dropped a label onto the site view when completing question 4a) of the aggregate data section in the study tasks (Appendix C).

“Right so you can change what the site is viewing based on the labels, ok cool, that’s cool.” - PID 18

Two participants reported being confused by the text “click and drag bar label onto other plots” which was displayed on the monthly averages visualization and they attempted to drag the bar instead, though the Likert test score shows that most participants did not require assistance. Observations from the video show that there were other participants who attempted to drag the bar and managed to correct their usage of the interface.

“Sorry what does it mean by it says drag and drop the from the label?” - PID 12

Scales The scale for the visualization was recalculated for each month which was displayed. One participant found the scaling tricky.

“I noticed on different days when the maximum value was different then it would change the scale of what your looking at, and for a sort of general getting kind of information out of it that’s not such an

issue but if your in a bit of half daze or whatever, you could look at something with a lower value and recognize it as being higher than what you just looked at ...” - PID 3

The monthly averages visualization could be improved by using stacked bars to show the average for the entire year rather than the average for a single month which would make different months more comparable.

5.3.7 Data Comparison

The data comparison visualization (Figure 3.9) was used to draw parallel coordinates for several pollutants at different sites, and the most frequently commented aspects of the visualization were the scale, colour, zoom and legend. The data comparison was reported least positive out of all the visualizations, based on the Likert scales. One previous user study measuring the effectiveness of a parallel coordinate plot found that the parallel coordinate plot was difficult for participants to use due to the amount of information visible on the parallel coordinate plot [36].

User study tasks 11/19 participants answered the parallel coordinate data comparison question 1a with a correct response. The question was about the relationship between PM_{10} recorded with a BAM and NO_2 . The data comparison was difficult for the participants to interpret.

Participant Likert scale post-study responses The data comparison Likert scale scores (Figure 5.8) were completed by 19/20 participants. The response to the data comparison was generally negative with 10 participants stating that the data comparison was ineffective or very ineffective for finding relationships among pollutants (Figure 5.8a), and 10 participants stating that the data comparison was ineffective or very ineffective for identifying pollutants of interest (Figure 5.8b). The text size was generally acceptable, 13 participants responded that the text size wasn't too

small or too large, though 6 of the participants stated that the text size was too small (Figure 5.8c). The text issue was also commented in the pilot study and the text was repositioned and scaled to fix the labelling, however, the changes made to the text in the pilot study were not sufficient for 6 participants though the majority of participants thought that the text size was correct. A better way of solving text issues would be to put an option in the options panel to adjust and scale the text.

Statistical analysis The mean Likert scale score in the data comparison section of the post-study questionnaire was calculated for each participant based on the method described by section 5.3.1. Table 5.9 groups the mean Likert scale test scores based on the amount of experience the participants have analysing air quality data. The table describes the mean, standard deviation and the number of participants in each group.

Table 5.9: Summary statistics for the data comparison section mean Likert score in the post-study questionnaire grouped by experience.

Air Quality Experience						
No Experience			Experience			
Mean	SD	n	Mean	SD	n	
3.2375	0.9661	8	2.3727	0.6424	11	

Welch's two-sample t-test was used to determine that the two groups "no air quality experience" and "air quality experience" do not have the same mean. ($t(11.40) = 2.20, p = .049, d = 1.09$). The t-test means that the participant's statement of their experience with air quality data did significantly alter the participant's Likert scale rankings of their experience with the data comparison and participants without air quality experience were

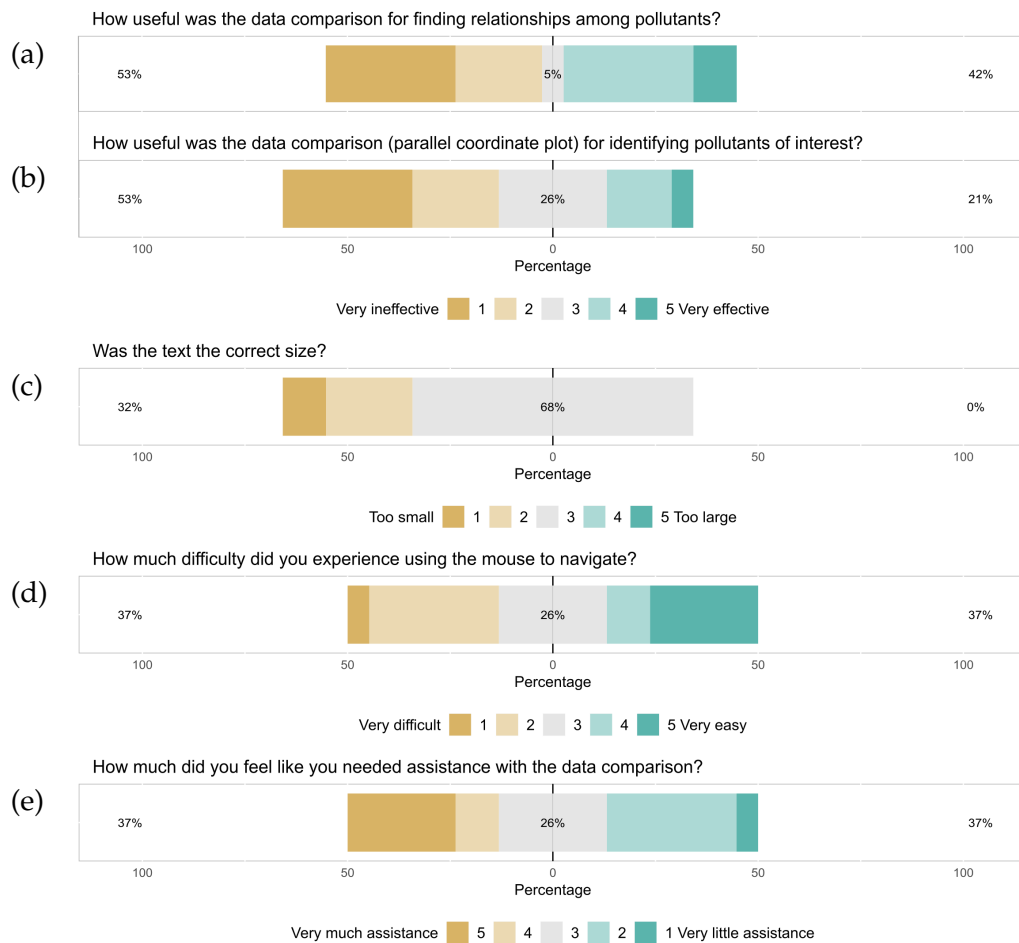


Figure 5.8: Data Comparison: Likert scales from post-study questionnaire.

more positive towards the visualization. The data comparison was evaluated in the post-study based on the effectiveness of the data comparison for finding relationships among pollutants, the effectiveness of the data comparison for identifying the pollutant of interest, the difficulty of using the mouse to navigate and the amount of assistance required.

Table 5.10 groups the mean Likert scale test scores based on the correctness of the response to mapping the data section question 4 (Appendix C). The table describes the mean, standard deviation, and the number of participants in each group. The mapping the data section question 4 was cho-

sen for an analysis based on accuracy as the number of incorrect solutions was sufficient for a t-test.

Table 5.10: Summary statistics for the data comparison section mean Likert score in the post-study questionnaire grouped by response accuracy.

Question Accuracy					
Incorrect			Correct		
Mean	SD	n	Mean	SD	n
2.008333	0.604497	6	3.0375	0.8227241	12

Welch’s two-sample t-test was used to determine that the two groups “incorrect” and “correct” do not have the same mean. ($t(13.35) = 3.00$, $p = .010$, $d = 1.35$). Participants answering data section question 4 correctly did give measurably different feedback on the post-study questionnaire Likert scales compared to participants answering the question incorrectly.

Effectiveness of the user experience The parallel coordinate plot was first used in mapping the data question 4 of the study tasks (Appendix C). In this question, the $PM_{2.5}$ measured with BAM and O_3 are compared for a single station in Woolston. There is also a parallel coordinate data comparison section question 1a where three pollutants, PM_{10} measured with BAM, and NO_2 are compared for a single station. Question 1b which asks the participant to compare results for two different site locations Newmarket and Henderson_Lincoln Rd.

Scale 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 through the scales could be changed by dragging and zooming the axis using the mouse.

The following statement was made after completing parallel coordinate data comparison question 1b.

“One useful scale is the current standards. You should default to that.” - PID 15

One participant was concerned by the gradients that can be produced by dragging bars on the parallel coordinate plot. Gradients between axis can be produced on any parallel coordinate plot, not just the data comparison provided to the participant. The gradients do not meaningfully show information in the dataset.

The following statement was made when comparing two sites on the data comparison visualization as part of the data comparison section question 1.

“It’s kind of weird ‘cause you can create like artificial gradients ‘cause those rows don’t line up.” - PID 13

Zoom The zoom allowed the range of the y-axis to be changed by scrolling the mouse. Each axis could be zoomed separately. 4 participants commented on the axis zoom. 2 participants asked whether it was possible to zoom the axis.

2 participants commented that the zoom was difficult to use due to unexpected mouse behaviour. These participants were observed positioning the mouse incorrectly before attempting to zoom.

“Maybe put in a couple of visual cues on the plot so you know where to point.” - PID 10

The zoom functionality was documented in a handout provided to the participants, however, a visual cue to zoom the axis would make the functionality more understandable. A visible box around the axis when the mouse hovers over would make the zoomable area clearer.

Colour Colour was used to distinguish between measurements which were taken at different sites and four participants commented on the choice of colour used by the data comparison. One participant commented that orange and green could appear on the same visualization. Making the colours customizable would allow the colours to be chosen consistently with other visualizations and would also avoid issues colourblindness which can affect the ability of a person to distinguish between different colours [58].

5.3.8 Windowing

Participants were able to drag, drop and move windows containing visualizations (Figure 3.1). The windowing system grid received a positive response from one participant.

“The snapping is really nice, I like the way it snaps and stuff to each other, its a good feature.” - PID 18

When a visualization was created by clicking on the menu the visualization was placed into the top left grid position. Three participants were confused as clicking on a visualization twice produced two visualizations in exactly the same layout position, though there was no visual cue to show that two different visualizations had been produced.

Three participants suggested additional tiling methods to manage windows and prevent clutter.

“It would be great if the tiles that I put them in get [were] not overlapping each other ... having some [of] the windows appear wherever there’s empty space with a smallish size.” - PID 15

A participant commented that the icons on the window frames were small and that the close button and the gear icon on the window frames could be difficult to read or click on. The gear icon was not understood by

all participants, though the gear icon was chosen for its familiarity and use on other settings. Four participants required assistance to find the options panel.

A better strategy would be to automatically tile visualizations on creation so that they do not overlap. This would ensure that a participant would be able to see whenever a visualization was produced.

In future saved layouts could be applied to position windows on the screen to reduce the amount of time spent on rearranging windows, also sessions could be saved and loaded present an overview.

5.4 Discussion

AtmoVis allows atmospheric air quality data to be explored and visualized using a selection of different connected visualizations. The objective of the user study was to evaluate the effectiveness of different visualizations for air quality data exploration (Section 1.1). In order to describe the effectiveness of the visualizations the results from the Likert scales, qualitative responses, study task responses and observations presented in Section 5.2 will be discussed.

The main user study produced some positive feedback for the heat calendar, site view, monthly averages, monthly rose and the way that AtmoVis allows the user to interact with the different visualizations. There were also aspects of the interface which could be improved and expanded to make AtmoVis more suitable for the target audience. This section discusses the results of the main user study for each visualization, possible improvements for AtmoVis and directions for future research.

5.4.1 Accuracy of AtmoVis for Persona Scenarios

Personas were used to assist in the design of AtmoVis (Section 3.1.1). Three personas, Oliver, Mathew, and Kath, were produced representing different

users in order to build an understanding of the users' requirements for AtmoVis. Each persona had goals and two example scenarios of where they would like to use the system. In the following section the persona scenarios are discussed in the context of the tasks performed in the user study and the accuracy of the response is reported. The accuracy is used to comment on the suitability of AtmoVis for the scenario tested.

Discussion of AtmoVis in the context of Kath's scenarios Kath's persona describes an air quality scientist. Kath's first scenario describes the requirement of finding directional trends in air quality data. The monthly rose is provided to describe trends between the concentration of a variable and wind direction. In aggregate data question 2 of the study tasks the distribution of PM_{10} measured with TEOM at a temperature of 50°C needed to be compared with wind direction at Woolston. The results indicate that trends could be effectively determined for the concentration of one variable compared with wind direction demonstrating the AtmoVis interface was effective for determining directional trends for a single variable at a monitoring site. Since the AtmoVis system only allows monthly roses for one time frame to be drawn at once, comparing monthly roses would be more difficult and the system could be improved by allowing a different time frame to be compared more easily.

Kath's scenario 2 in Section 3.1.1 identifies a requirement for finding out how related two pollutants are. The data comparison visualization and the monthly averages were designed to allow two pollutants to be compared directly. The data comparison visualization was used to compare the relationship between PM_{10} measured with a BAM and NO_2 at Newmarket in data comparison question 1 a. 11 participants answered the question correctly. The line plot was also used to compare the relationship between O_3 and solar radiation in mapping the data Question 2b. 16 participants answered the question correctly demonstrating that the line plot was more suitable at identifying the relationship between two variables

even though two line plots were required with one variable displayed on each. The feedback from the study participants using the line plot indicates that being able to compare different pollutants at the same site is an important feature as 6/20 participants suggested adding functionality for displaying more than one variable. So building more comparative functionality into the other visualizations would improve the usability of the interface. AtmoVis is effective for finding the relationship between the two pollutants and the line plot would assist with Kath's second scenario more than the data comparison.

Discussion of AtmoVis in the context of Oliver's scenarios Oliver's persona described a data analyst and Oliver's first scenario indicated that peak pollution levels should be visible using AtmoVis. The results of aggregate data question 5a indicate that participants were able to find the day of the year where SO_2 pollution was the worst so it was possible to identify the peak average pollution level using the line plot. The results of the aggregate data question 5c also indicate that participants were able to effectively find the worst time on a 72 hour line plot, showing that the line plot is an effective visualization for tasks such as Oliver's first scenario. Daily averages were visualized with the heat calendar as well as hourly peaks.

Oliver's second scenario requires the AtmoVis system to make aggregate data available through the interface so that averages can be compared. The aggregate data question 3a requires that monthly roses are compared to compare the distribution of SO_2 with wind direction for two monitoring stations in Christchurch. The results of aggregate data question 3a indicate that 5/20 participants answered with a correct comparison of both wind direction and pollution level at both locations, 14/20 participants answered the question with partial correct solutions. Though 19/20 participants were able to complete part of the question, information produced by the comparison was limited by the ability to compare the differ-

ent scales for the monthly roses so comparing monthly roses for aggregate data comparison between sites was less effective than reading information from a monthly rose at a single site. The heat calendar also makes aggregate data available and all participants were able to correctly identify the month containing the highest daily mean level of PM_{10} measured with TEOM at a temperature of 50°C in Woolston. The use of the heat calendar demonstrates that AtmoVis does make aggregate data available for temporal visualization tasks which would assist with Oliver's second scenario by allowing daily averages to be compared between different months.

Discussion of AtmoVis in the context of Mathew's scenarios Mathew's persona described a university student with an interest in air quality monitoring. Mathew's first scenario focused on the exploratory nature of the interface for finding out what pollutants are available at a given site and for discovering a temporal trend. The monthly averages visualization provided a method for finding out which pollutants were available at a given site. In the aggregate data question 3b the participant was required to state what else is being measured at Woolston. 19/20 participants used the monthly averages to report some of the other measurements. There were a large number of measurements available at the site and participants did not list all of them. The monthly averages visualization would assist in situations such as Mathew's first scenario where the available pollutants need to be identified for a given location.

Mathew's second scenario required aggregate data to be compared to find out if there were temporal differences in the averages. In the temporal pattern question 1b, participants were asked to use the site view, monthly averages chart and heat calendar to examine air pollution in Masterton during 2012 and to comment on any seasonal trends or days with the highest pollution. 16/20 participants answered the question as the study was long and some participants chose not to complete. 15/16 answered with a correct trend discussing the higher levels of particulate

matter in winter. None of the participants quoted exact averages with the monthly averages though comments from the use of the visualization indicated that the scaling could have made a temporal comparison more difficult using the averages than the temporal comparison using the heat calendar so rough trends were reported. The heat calendar would assist with Mathew's second scenario as it was suitable for comparing trends over time for the mean daily level of air pollution at Masterton. The comparison of the monthly averages over time was less effective than the use of the heat calendar for answering the question. The monthly averages only showed the month for the time frame and comments from one participant suggested that the visualization system would be more effective for comparing averages if all of the averages for a year were visible. The temporal comparison of the averages was more effective through a visualization like the heat calendar rather than through a visualization like the monthly averages.

5.4.2 Heat Calendar

The heat calendar was effective for reading time series trends in the data and finding days with a high daily average, All twenty participants found the month with the highest daily average pollution for PM_{10} measured with TEOM at a 50°C temperature setting. A calendar heat map has previously been used to visualize air pollution across China [69] and their calendar heat map for visualizing temporal trends was used with a geospatial map for visualizing spatial trends [69]. In contrast to this research project the system presented by Zhou et. al. [69] placed the heat map calendars directly onto the geospatial map. In AtmoVis, the heat calendar was a separate visualization and data was inserted onto the heat calendar by drag and drop from the site view. The heat map calendar used by Zhou et. al. [69] was effective because it was used to find temporal trends, though, a formal user study was not conducted on the visualization. In contrast,

a formal user study was conducted as part of this research and the heat calendar received positive responses, the mean Likert scale score for the heat calendar section of the post-study questionnaire had a median of 4.38, calculated with the method described in Section 5.3.1. Differences in the Likert scale responses from the post-study between participants with experience in air quality data analysis and participants without experience were not able to be detected ($p = .311$). Experience with air quality analysis did not effect the participants response to the effectiveness of the visualization. Both air quality experts and non experts were able to read and understand the heat calendar visualization. If AtmoVis was extended in the future with other datasets the heat calendar could display information about estimated emissions in the area as well as the pollution in the atmosphere. The heat calendar is also generally applicable to data which is not directly related to air pollution.

5.4.3 Line Plot

The line plot was effective for reading time series trends in the data and for finding peak pollution levels over a short time frame based on the results of the study task question where 16 participants found the relationship between O_3 and Solar Radiation at a single site. The line plot had generally positive responses in the Likert scale questions with an average score of 3.60 (Figure 5.2). The Trelliscope visualization system [30] also uses a line plot to display time series trends and similarly this user study on Trelliscope demonstrated that the line plot was effective when used with a recommendation system for identifying generator trips. The Trelliscope line plots were generated using data from a recommendation system in contrast to this research project which did not use a recommendation system. Differences in the Likert scale responses from the post-study between participants with experience in air quality data analysis and participants without experience were not able to be detected ($p = .093$). Experience with

air quality analysis did not effect the participants response to the effectiveness of the visualization and both air quality experts and non experts were able to understand the information presented by the line plot.

The line plot could be extended by allowing the colour scheme to be chosen as a preference and inserting a legend for the colour instead of relying on the mouse hover to show the site. Some participants asked whether it was possible to put more than one variable onto the same scale. Adding a button for zooming the visualization instead of relying on the mouse scroll would make the visualization more intuitive as only some participants used the mouse to zoom the line plot when pollutant levels were low. The line plot could be used in the future to display information from other datasets such as traffic monitoring data or additional meterological information.

5.4.4 Monthly Rose

The monthly rose was effective for finding trends in wind direction and in the directional concentration of air pollution at a given station. 15/20 participants found some trend in the wind direction and 9 found the relationship between concentration and wind direction correctly. It was more difficult to answer the question about the comparison between Woolston and St. Albans for SO_2 pollution at a given time (Question 3). 19/20 participants answered with a valid solution, 5 of those participants answered with a correct comparison of both wind direction and pollution level at both locations and 14 participants answered the question with partial correct solutions so the comparison between different locations was less effective. wind rose plots have previously been used to present the effects of wind on background particle concentrations at freight truck terminals [29]. The wind rose plots were overlayed onto a geographical map so that the surrounding area of the sensor could be visualized. Statistical tests were performed in that study to find differences in concentration based on wind

direction. AtmoVis can be used to visualize monthly roses and the window based layout allows the monthly roses to be dragged and positioned. AtmoVis does not have the functionality to also perform statistical tests over a timeframe though this functionality could be added in the future to expand the possible uses for the monthly rose.

The monthly rose received positive responses, the mean Likert scale score for each participant had a median of 3.33, calculated with the method described in Section 5.3.1. Differences in the Likert scale responses from the post-study between participants with experience in air quality data analysis and participants without experience were not able to be detected ($p = .546$). Experience with air quality analysis did not effect the participants response to the effectiveness of the visualization.

5.4.5 Site View

The site view was effective for accessing the different sites that were available. 16/20 participants were able to locate sites with a particularly high value for a pollutant based on the colour, though there were some possible improvements which would make sites easier to find (Section 5.3.5). A previous user study has incorporated a geospatial map to visualize pollution in China [69]. The geospatial map was found to be effective for visualizing the site locations and a heat map was used to visualize a calendar showing temporal trends [69]. Unlike this research project, Zhou et. al. did not conduct a user study on the effectiveness of the visualizations. The visualizations were used to find trends in the air pollution in China and those results were reported.

19/20 participants managed to use the site view to interact with the line plot by dragging and dropping points demonstrating that the site view was effective for interacting with other visualizations (Section 5.3.3). The Likert scale averages show that the site view had positive responses overall with a median score of 3.17, calculated with the method described in

Section 5.3.1 (Figure 5.2), though improving the use of colour could make the site view clearer. The font was generally satisfactory though 6 participants considered the font size too small (Figure 5.6). Differences in the Likert scale responses from the post-study between participants with experience in air quality data analysis and participants without experience were not able to be detected ($p = .620$). Experience with air quality analysis did not effect the participants response to the effectiveness of the visualization, and both air quality experts and non experts were able to understand the information presented by the site view. Qualitative responses suggested that geographical knowledge of the area did help participants when finding locations.

The site view could be extended with a search feature to make it easier for participants to locate, some participants had difficulty locating places which were specified on the question sheets. The site view could also be extended by adding overlays for the regions or allowing more than one pollutant to be visualized at once. The geospatial map presented in the study by Zhou et.al. [69] did not provide a search feature and zoomed the map so that the entire map was visible. Enlarging the font on the hover menus would improve the usability of the site view. The use of the site view would generalize to other datasets, for example if traffic data was inserted then the site view could be used to show information about congestion through other types of sensors.

5.4.6 Monthly Averages

17/19 participants were able to use the monthly averages to drag and drop a pollutant onto a monthly rose so the monthly averages visualization was effective for displaying the variables available at a given site and allowing the variable on another visualization to be changed easily. The monthly averages visualization was a bar chart of the pollutant levels averaged over a given month and participants were able to see all the levels on a

single scale with the units displayed underneath the column.

Bar graphs have previously been used by the Voyager system to describe data. Different categorical variables can be graphed against a numerical variable. Though the way that the user interacts with the variables contrasts with AtmoVis. Voyager provides a list of fields which can be dragged and dropped onto an encoding section and different views are produced through a recommender engine [65]. The labels on the bar graph cannot be dragged and dropped from the visualization onto other visualizations, unlike AtmoVis which uses the monthly averages visualization to show what variables are available and provide interaction with other visualizations through drag and drop.

The monthly averages received positive responses from participants, the mean Likert scale score for each participant had a median of 3.10, calculated with the method described in Section 5.3.1. There was a statistically significant difference between responses for participants with previous experience in air quality data analysis and participants without experience in air quality data analysis ($p = .002$). The mean Likert scale score for participants without experience in air quality data analysis had a higher mean than participants with experience in air quality data analysis, the means were 3.88 and 2.73 respectively, showing that participants without experience in air quality data analysis were more positive about the effectiveness of the visualization. The monthly averages visualization could be extended in the future to show information from emission inventories so that the amount of air pollution in the atmosphere could be compared to the amount emitted from the surrounding area. Allowing more than one time frame to be shown on the monthly averages would make the visualization more effective for temporal comparison.

5.4.7 Data Comparison

11 participants answered the parallel coordinate data comparison question about the relationship between PM_{10} recorded with a BAM and NO_2 with a correct response. The data comparison was less effective for identifying relationships between variables than the line plot. Parallel coordinate plots have previously been evaluated in a user study evaluating the usability visualization methods in an exploratory geo-visualization environment [36]. Their study required participants to complete tasks and the tasks were graded zero or one depending on whether the solution was correct or not and the results of the study found that participants could become confused by the amount of data displayed by the interface. Also, the parallel coordinate plot was difficult to use for tasks such as identifying relationships between variables where a 55% success rate was recorded which is consistent with the results of the user study for AtmoVis where 55% of participants managed to identify the relationship based on the results for data comparison question 1a. Parallel coordinate plots are an effective visualization for showing the relationship between a very large number of variables as correlations can be visually identified when there is a large number of axis on the visualization and the variables are treated equally [56]. The poor accuracy could be due to the lack of previous experience with a parallel coordinate plot.

The data comparison received negative responses from participants compared to the line plot, the mean Likert scale score for each participant had a median of 2.75, calculated with the method described in Section 5.3.1. There was a statistically significant difference between the responses for participants with previous experience in air quality data analysis and participants without experience in air quality data analysis ($p = .049$). The mean Likert scale score for participants without experience in air quality data analysis had a higher mean than participants with experience in air quality data analysis, the means were 3.24 and 2.37 respectively, showing that participants without experience in air quality data analysis were

more positive about the effectiveness of the visualization.

5.4.8 Windowing

There were also differences detected in the way that window systems are used by different participants. Participants in the study were observed piling windows on top of each other, tiling them out neatly, using windows near maximized, or ‘splattering’ [59] windows around the screen. The observations about the use of windows in AtmoVis demonstrate that differences usage styles can emerge.

The drag and drop interaction is really powerful, however, it is limited because only some plots can have data dragged out of them. The windowing of the interface does encourage participants to have a large number of windows visible on the same interface. One study on window switching found that on a single monitor the median number of windows that a user has visible is 1.7 [59], though the video of the task completion demonstrated that the median of the number of windows visible for each user in mapping the data question 4 (Appendix C) was 5 windows visible. When the windows are used in a comparative connected way the number of windows visible when using AtmoVis will be higher than the median for a single monitor setup. The study on window switching also found that when the number of windows increases the amount of time switching also increases. So extending AtmoVis to use a window switching feature and supporting multiple monitors could reduce the amount of time spent window switching. A study on the use of a multi monitor high-resolution displays found that participants spent less time using file system navigation, minimizing and maximizing. Instead, they piled up windows on different monitors in categories as an extension of their memory [17].

The interface of AtmoVis could be extended to allow different groups of visualizations to occupy different time frames. Instead of grouping all of the plots to one play button, there could be groups with a play button

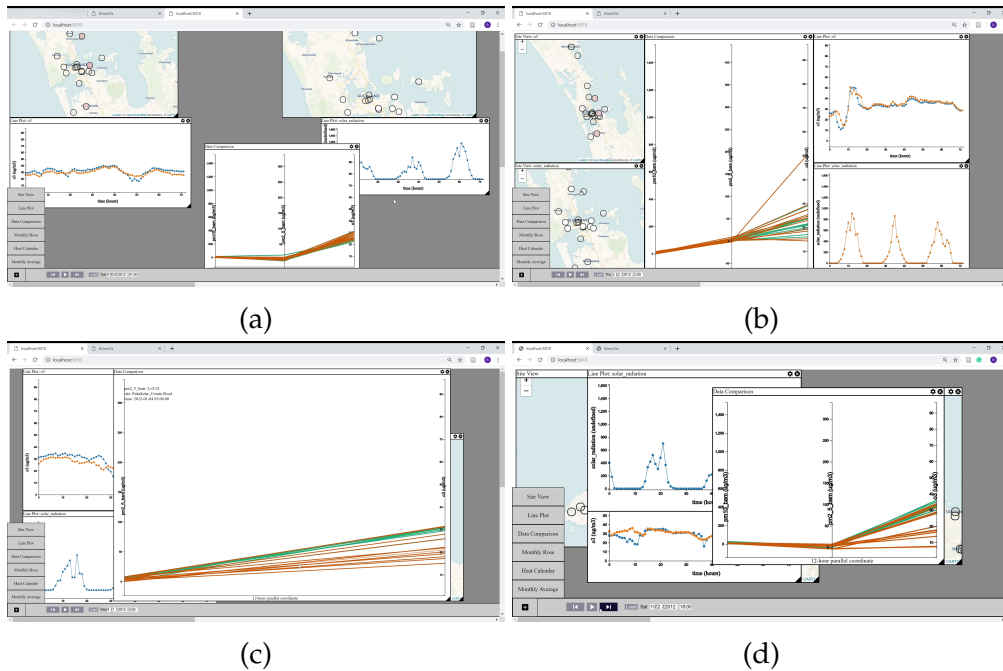


Figure 5.9: Four different usage styles in AtmoVis, Figure 5.9a shows window splattering, Figure 5.9b shows border to border careful tiling, Figure 5.9c shows a near maximized window and Figure 5.9d shows a stack of windows.

each. Drag and drop could be extended by allowing calendar days to be dragged to move data on that day, allowing more of the plot types to be filterable, e.g dragging a selection of data out of the line plot into another plot.

5.5 Suggested Improvements

Further Testing In a “within subjects” test the same group of participants test different versions of a prototype [21]. Due to time constraints and cost constraints, only one system was tested in the main user study after the pilot study was performed on different AtmoVis prototypes. In the future, several iterations of user testing could be performed to further develop the prototype using the same participants. The laboratory style test was chosen as incorporating AtmoVis into the daily work of a

group of participants was not practicable at this stage in the projects' development. The user testing could be integrated into the daily use of the software tool for a group of professionals and a more detailed study of the usability of the tool could be performed over a longer time frame in a workplace environment.

Data extraction AtmoVis does not implement all of the functionality described in Shneiderman's taxonomy [52]. For example, there is no way to extract a selection of data out of the visualization system to save in a spreadsheet and perform further analysis.

"I can't get the monthly average data from this panel, so if there is a button I can automate that [exports data] into a CSV file or something so the user can use it directly otherwise they should write [it] down."
- PID 16

Metadata Some participants reported that they did not fully understand the nature of the data that they were looking at. AtmoVis could be expanded with a metadata viewer which can provide textual information about monitoring sites and pollutants measured as well as photographs of monitoring equipment installed.

"So my one question would be what is the resolution of this data is it like 1 min, 10 min, 1 hour in terms of the data source." - PID 5

Saving state The windowing system could be improved by allowing the window layout of the system to be saved and reloaded.

5.6 Summary

AtmoVis was evaluated in a pilot study and the main user study. The visualizations for AtmoVis were tested by the participants completing study

tasks and answering questions about the visualizations in a post-study questionnaire. The heat calendar received the highest median score for effectiveness (4.88), the line plot received the second highest median score (3.60), and the data comparison received the lowest median score for effectiveness (2.75) (Section 5.3.1). The heat calendar received positive responses from participants about the way that colour was used to show trends in the daily averages for pollutants and for the way that the heat calendar was used to change the date and time when a day was clicked on (Section 5.3.2). 13/20 participants responded that the line plot was effective for describing temporal patterns and 18/20 participants answered mapping the data question 2b correctly showing that the line plot was effective for comparing the same variable at different sites. 16/20 participants answered mapping the data question 3b correctly showing that two line plots were effective for comparing two different variables (Section 5.3.3). 14/20 participants stated that the monthly rose visualization was effective for identifying pollutants (Section 5.3.4). Welch's t-test found statistically significant differences in the median effectiveness reported by participants with experience in air quality data analysis and participants without experience for the monthly averages (Section 5.3.6) and the data comparison (Section 5.3.7). For both visualizations, the responses from participants with no experience in air quality analysis were more positive than the responses from participants with experience in air quality analysis. The user study identified that the windowing system was usable by the participants and four different usage styles were observed: Piling windows on top of each other, tiling neatly, splattering windows around the screen and nearly maximizing one window (Section 5.4.8). The use of the windowing system with the visualizations indicates that the window system was effective and participants were observed dragging, dropping and clicking window borders to organize the layout. There was some feedback from participants which suggested additional tiling functionality would be useful. AtmoVis was effective for the exploration of air quality data.

Chapter 6

Conclusions

Air pollution has a variety of adverse health effects including cardiovascular disease [18] and respiratory mortality [31]. In order to understand the effects of air pollution on health and on the environment at a particular location, data containing air pollution information needs to be collected and analysed from that location because air pollutants such as particulate matter disperse. Data visualization can assist with data analysis by making the collected data easier to interpret. In this research a novel web-based data visualization prototype, *AtmoVis*, was developed to visualize air pollution data. Experts in air quality monitoring such as environmental scientists were consulted during the development of *AtmoVis* to assist with the requirements gathering process. Personas and persona goals representative of the target audience were produced from the information gathered. The air quality experts had experience with analyzing air quality in New Zealand so an air quality dataset of New Zealand was integrated into *AtmoVis*. In New Zealand, there are some locations where air quality is poor, though the air quality is generally good in most places. *AtmoVis* was developed iteratively by using the results of pilot testing to inform the design process and study protocol before performing the main user study to evaluate the effectiveness of the system for visualizing air quality. The heat calendar, line plot, site view, monthly averages, and monthly rose were

found to be effective and the line plot was the most effective for comparing temporal trends in the air quality data. The following contributions were made in this thesis.

6.1 Contributions

- **A novel web-based prototype data visualization system was developed to visualize air quality data.**

AtmoVis has six visualizations: Site view, line plot, heat calendar, monthly averages, data comparison, and monthly rose. These visualizations are provided through an interactive windowing interface with a breadth-first design. The functional requirements for AtmoVis were developed using three personas, Oliver, Mathew and Kath. The three personas were used to identify the functionality required for AtmoVis and to produce the AtmoVis system goals (Section 3.1.5). The personas were useful in the construction of the user study tasks because good personas produce realistic system goals and study tasks which are suitable for measuring the effectiveness of the system (Section 3.1.5).

- **A user study was conducted to evaluate the effectiveness of AtmoVis, the visualizations contained in AtmoVis and to make recommendations for the improvement of the tool.**

The results of the user study were used to evaluate the effectiveness of AtmoVis and discuss the responses produced. The user study identified which visualizations were most effective and identified ways to improve AtmoVis and the visualizations.

The most effective visualizations in the context of the system goals are discussed below. The effectiveness was judged based on the accuracy of the visualization and feedback from experts (Section 4.4).

Goal 1 Allow pollutants measured in a region to be discovered.

The site view was effective for displaying locations of interest and 17/20 participants were able to use the monthly averages to find the pollutants at a given site (Section 5.3.6).

Goal 2 Allow regions measuring a pollutant to be discovered.

The site view was found to be effective for displaying the monitoring sites (Section 5.3.5). 14/20 participants indicated that the mouse was easy or very easy to use for navigation.

Goal 3 Allow temporal trends for a pollutant to be compared.

Both the heat calendar and the line plot were effective for comparing temporal trends between pollutants. 20/20 participants were able to find the highest daily average level of PM_{10} at Woolston using the heat calendar (Section 5.3.2). 16/20 participants were able to find the relationship between solar radiation and ozone at a single site using the line plot (Section 5.3.3). The data comparison was less effective with 11/20 answering a question comparing PM_{10} with NO_2 correctly (Section 5.3.7).

Goal 4 Allow spatial trends to be compared.

The monthly rose was effective for finding trends in the wind direction of a variable, though the site view was less effective for identifying spatial trends. The line plot and the data comparison also allowed trends between sites to be compared and 18/20 participants compared the levels of O_3 at two different sites correctly using the line plot (Section 5.3.3). The line plot was more effective for comparing relationships between variables than the data comparison (Section 5.3.3).

Goal 5 Encourage a breadth-first exploration of the data collected and to reduce the barrier for investigating air pollution.

AtmoVis is designed to encourage a “breadth-first” exploration and allows variables at different sites to be visualized using a window based system.

6.2 Research Questions

During the user study, the participants were required to use the line plot, monthly rose, monthly averages, site view and data comparison to explore the data and identify relationships between variables in the dataset. The line plot and the heat calendar were most effective for assessing the temporal trends in the visualizations.

RQ1 How effective are the visualization techniques for exploring air quality data?

The heat calendar had the highest score for effectiveness based on the post-study questionnaire responses. The heat calendar received positive feedback from participants about its navigational use. The visualization system received generally positive feedback in the post-study questionnaires and the responses also identified some parts of the visualization system which could be improved. The data comparison received a lower effectiveness score than the other visualizations and there was a statistically significant difference between users with experience analysing air quality data and users without. The users without experience gave more positive responses about the effectiveness of the visualization compared to users with experience. The heat calendar and the line plot required the least assistance. They were also the most effective in terms of the overall effectiveness of the visualization. The data comparison required the most assistance and received the lowest overall effectiveness score.

RQ2 How effective is the user experience of AtmoVis for exploring air quality data?

The window interface was evaluated by inspecting the participant's use of the interface. Participants were able to effectively use several windows at once to answer questions. The drag and drop interactivity between different visualizations received positive feedback from participants.

RQ3 How accurate are experts when using the visualizations?

The user study tasks were marked to find out whether the visualizations evaluated were being used accurately. The accuracy was discussed in the analysis of the system goals (Section 6.1, system goals 3 and 4). 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 effective because the participant was 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 effective for spatial analysis, and the site view was effective for navigation.

In conclusion, the results of the user study demonstrate that air quality data analysis in New Zealand could benefit from interactive visualization through a web-based interface. The heat calendar, line plot, site view, monthly averages and monthly rose were effective for analyzing air quality through AtmoVis and an interactive web based interface for data exploration with a window layout was an effective method for accessing these visualizations and inferring relationships among air quality variables at different monitoring sites.

6.3 Future Work

AtmoVis could be improved by future work in the following areas:

Windowing The windowing interface could be extended with additional functionality for minimization, maximization, window switching and multi monitor support in order to make the windowing system faster and easier to use when there are a large number of windows open. The evaluation of the windowing interface could be extended to evaluate the additional changes.

Monitors A limitation of AtmoVis is that it was designed for a single monitor. A multi monitor system was not tested. There are differences in window placement for application windows in single monitor and multi monitor systems so extending AtmoVis to allow a multi monitor layout would both improve the functionality of the tool and enable other usage styles to be investigated.

Evaluation The evaluation of AtmoVis could be extended to integrate AtmoVis into the daily work cycle after improvements have been made, then evaluate how participants use AtmoVis in a work environment rather than in a laboratory environment. An “in-the-wild” [48] study could be performed to measure the effect of the software on human behaviour when it is used in a particular way. Further research could be performed on the efficiency of the program by logging the amount of time taken to generate different visualizations and whether there are any new efficiency issues when the amount of data increases. More statistical within subjects testing could be performed after extending the system to determine whether changes to the interface would make the system better.

Additional Features Allowing data to be exported from spreadsheets would assist with the analysis of the data so the visualization system could be extended by adding an export button to visualizations to save a spreadsheet or other file formats. Loading and saving visual layouts would allow analysis to be paused and resumed later.

Additional datasets could be built into the system so that traffic data, water quality, land use, and other information could be compared to air quality. Inserting metadata about emissions sources in each region would allow more inferences to be made about the causes of the air pollution measured in each region.

The functionality of the system could be extended using a recommender engine to find areas that are particularly interesting or pollutants that are related. Recommender engines are used by other data visualization systems and can recommend data for analysis based on similarity.

A file export system to produce layouts of different visualizations using an R script would allow more bidirectional interaction between other systems. A plug-in extension system which adds interaction to visualizations produced by other systems would allow the interface to consolidate many visualization systems into one place and improve the utility of the interface.

AtmoVis could start with a help window visible to assist first time users. The help window would provide a reference describing the system functionality in an interactive way. AtmoVis could be made more intuitive by adding a help system which can be dragged and dropped from the menu like a visualization. The help system would open up a documentation viewer in a window.

AtmoVis allows visualizations from R to be used, but implementing a domain specific language could allow more complex plots to be produced by performing operations on the data, e.g plotting the difference between two variables. An algebra for generating tables was defined as part of the Polaris system [57] (Section 2.2.4) and a similar algebra for generating visualizations could be defined to compose together visualizations in AtmoVis.

Appendices

Appendix A

Information Sheet



VISUAL ANALYTICS FOR AIR QUALITY: INFORMATION SHEET FOR PARTICIPANTS

You are invited to take part in this research. Please read this information before deciding whether or not to take part. If you decide to participate, thank you. If you decide not to participate, thank you for considering this request.

Who am I?

My name is Benjamin Thomas Powley and I am a Masters student in Computer Science at Victoria University of Wellington. This research project is titled "Visual Analytics for Air Quality Data".

What is the aim of the project?

The "Visual Analytics for Air Quality Data" project is an M.Sc. research project aimed at designing and developing a prototype system for exploratory visual analysis for large, complex spatio-temporal datasets and measuring the effectiveness of the visualizations used within the system. This research has been approved by the Victoria University of Wellington Human Ethics Committee, Application reference number #0000026810.

How can you help?

You have been invited to participate because you are either a domain specific expert in air quality monitoring, or a domain specific expert in visualization systems. If you agree to take part I will interview you at Victoria University of Wellington. I will ask you to perform tasks using a visual interface for data exploration. Screen recording, key logging and audio recording software will be in use during the study to observe your interactions with the software. I will ask you questions about your usage of the interface and the question answers will be combined with information from the recordings taken. The interview and tasks will take about 60 minutes.

You can choose to not answer any question or stop the interview at any time, without giving a reason. You can withdraw from the study by contacting me at any time up to two weeks after the interview date. If you withdraw, the information you provided will be destroyed or returned to you.

What will happen to the information you give?

This research is confidential*. This means that the researchers named below will be aware of your identity but the research data will be combined and your identity will not be revealed.

* Confidentiality will be preserved except where you disclose something that causes me to be concerned about a risk of harm to yourself and/or others.



Only my supervisors, the transcriber (who will be required to sign a confidentiality agreement) and I will read the notes or transcript of the interview. The interview transcripts, summaries and any recordings will be kept securely.

What will the project produce?

The information from my research will be used in Masters and/or academic publications and conferences.

If you accept this invitation, what are your rights as a research participant?

You do not have to accept this invitation if you don't want to. If you do decide to participate, you have the right to:

- choose not to answer any question;
- ask for the recorder to be turned off at any time during the interview;
- withdraw from the study any time up to two weeks after the interview;
- ask any questions about the study at any time;
- be able to read any reports of this research by emailing the researcher to request a copy.

If you have any questions or problems, who can you contact?

If you have any questions, either now or in the future, please feel free to contact either:

Student:

Name: Benjamin Thomas Powley
Email address:
powleybenj@myvuw.ac.nz

Supervisor:

Name: David Pearce
Role: Senior Lecturer
School: Engineering and
Computer Science
Phone: 04 463 5833
Email:
david.pearce@ecs.vuw.ac.nz

Supervisor:

Name: Craig Anslow
Role: Lecturer
School: Engineering and
Computer Science
Phone: 04 463 6449
Email:
craig.anslow@ecs.vuw.ac.nz

Human Ethics Committee information

If you have any concerns about the ethical conduct of the research you may contact the Victoria University HEC Convenor: Dr Judith Loveridge. Email hec@vuw.ac.nz or telephone +64-4-463 6028.

Appendix B

Pre-Study Questionnaire

Visual Analytics for Air Quality Data : Pre-study questionnaire

General:

1. Age: <20 20 - <25 25 - <30 30 - <35
 35 - <40 40 - <45 45 - <50 50 - <55
 55 - <60 60+

2. Gender: Male Female Other

3. Occupation:
 data analyst researcher other

if you circled other please state your occupation:

Analysis tools:

Have you used any of the following tools for work or study?

3. Spreadsheet Software.

Yes No

Please specify your level of expertise:

Basic Features 1 2 3 4 Advanced Features 5

(a) Have you used spreadsheet formulas?

Infrequently 1 2 3 4 Frequently 5

(b) Have you used spreadsheets for visualizing data (e.g chart, line plot, histogram e.t.c)?

Yes

No

What sort of visualizations do you draw using spreadsheets for work or study?

(c) Have you written spreadsheet macros in a programming language?

Yes

No

If yes, please specify which programming language:

(d) Have you produced macros using keystroke recording?

Yes

No

Programming tools:

Have you used any of the following tools for work or study?

4. R

Yes

No

Please specify your expertise:

Basic features

1

2

3

4

Advanced features

5

5. Python

Yes

No

Please specify your proficiency:

Basic features
1

2

3

4

Advanced features
5

6. Other programming languages.

Yes

No

If yes, please specify which programming languages:

Please specify your expertise:

Basic features
1

2

3

4

Advanced features
5

3rd party visualization software:

Have you used any of the following for work or study?

7. Tableau

Yes

No

If yes, please specify your usage:

Please specify your proficiency:

Basic
1

2

3

4

Advanced
5

8. PowerBI

Yes

No

If yes, please specify your proficiency:

Basic
1

2

3

4

Advanced
5

Please specify your usage:

9. Other visualization software? Please specify:

Experience with Data:

1. Do you have prior experience in atmospheric science?

Yes

No

If yes then for how many years.

Please comment on your previous experience:

2. Do you have previous experience with data analysis?

Yes

No

If yes then how many years.

If yes then what sort of experience?

3. Do you have any experience using different kinds of data sets other than the air quality monitoring datasets e.g traffic congestion data?

Yes

No

If yes then discuss the data set and the type of data.

Data set knowledge:

1. Have you ever analyzed/used the NIWA air quality monitoring data set?

Yes

No

If yes, please specify your proficiency:

Basic
1

2

3

4

Advanced
5

2. Have you ever analyzed/used air quality monitoring data sets from a different source e.g Air quality data sets from LAWA?

Yes

No

If yes, please specify your proficiency:

Basic
1

2

3

4

Advanced
5

If no for both questions 1 and 2 then skip to question 9

3. What sort of comparisons did you make between different air pollutants at the same monitoring station?

4. What sort of comparisons did you make between air pollutants on a larger scale spanning several monitoring sites?

5. Were you using aggregate data derived from the station monitoring data?

Yes

No

If your answer was yes what sort of derived data were you using?

6. Were you interested in time series trends for air quality monitoring data at a given station?

Yes

No

7. When analyzing/using air quality monitoring data sets what sort of analysis software did you use?

8. Were you comparing/ analyzing any other datasets alongside the air quality dataset to infer air quality information?

Yes

No

If yes please provide information on the nature of the other datasets:

9. Please state/describe any other relevant research interests or experiences working with data

Thank you for participating.

Appendix C

Study Instructions

Visual Analytics for Air Quality Data: User Study Tasks

Mapping the data:

1. Use the menu to insert a map of New Zealand by dragging the site view button onto the screen. Then use the options panel to select O_3 .

(a) What is the start date for the O_3 measurements displayed on the options panel?

(b) Enter the start date and time from the options panel into the time selector next to the play button, change back to map view and click load. Reading points from the map, how many recording stations are highlighted with filled in color on the start date?

2. Leaving the map in place, Insert a line plot. Drag the Howic_Music Point (in Auckland) O_3 site onto the line plot.

(a) At what date and time is the highest pollution level for O_3 visible on the line plot? What is the value?

(b) Insert Whangaparaoa_Shakespear Park (in Auckland) onto the same line plot, how are the levels of O_3 related?

3. Add a second map of New Zealand onto the screen and set the map to show solar radiation. Insert an additional line plot and drag the Howic_Music Point from the solar radiation map onto the line plot.

(a) How many monitoring stations are recording both O_3 and solar radiation at the date and time on the time selector (i.e. how many stations have filled in color for both O_3 and solar radiation)?

(b) Using both line plots can you see any relationship between O_3 and solar radiation at Howic_Music point? Describe the relationship.

4. Insert a parallel coordinate plot by dragging and dropping the data comparison button from the launcher menu, Use the options panel to set the pollutants to *PM_2_5_BAM*, *PM_10_BAM* and *O₃*. From the *O₃* map, drag and drop Pukekohe and Whangaparaoa (in Auckland) onto the parallel coordinate plot.

Click the play button and observe for about 12 hours. Is there any relationship between particulate matter *PM_2_5_BAM* and *O₃*? Describe the relationship.

Aggregate data

Reload the page before starting this section.

1. Insert a site view, set the pollutant to *PM10_TEOM50*. Insert a heat calendar and set the time selection to 01/01/2012 1:00. Drag Woolston (in Christchurch) onto the calendar.

What month contains the highest daily average level of *PM10_TEOM50* in 2012 for the Christchurch monitoring station?

2. Without removing the calendar, Insert a monthly rose plot and drop Woolston onto the plot. On the calendar, click on any day in the month with the worst *PM10_TEOM50* pollution.

Is there any relationship between wind direction and the level of *PM10_TEOM50* on the month with the highest *PM10_TEOM50* pollution in 2012? Describe the relationship.

Reload the page before starting this section

3. Insert a site view with the pollutant set to *SO₂*, load the data. Insert a monthly average chart and two monthly rose plots. Set the date to 01/01/2012 1:00 and drag Woolston onto the monthly average chart. Drag Woolston onto one of the monthly rose plots then drag St.Albans_Coles Pl (in Christchurch) onto the other rose.

(a) Comment on the distribution of *SO₂* with wind direction.

(b) What else is being measured at Woolston ?

4. Drag and drop the PM2_5_FDMS label from the monthly average chart onto the map view. Insert two more monthly rose plots. Then drag and drop a point from Woolston onto one of the roses and the point from St.Albans_Coles PI onto the other rose.

(a) Comment on any similarities or differences between PM2_5_FDMS and SO₂.

(b) Delete the monthly average chart and replace it with a calendar view. Drag and drop Woolston onto the calendar view. Click on a day in February and comment on the similarity or difference between PM2_5_FDMS and SO₂ for the month of February on the monthly roses. Is there a trend continued from January?

Reload the page before starting this section

5. Insert a site view with the pollutant set to SO₂. A monthly average chart, A calendar and one line plot. Set the date to 01/01/2012 1:00, Drag and drop Woolston onto the monthly average chart. Drop Woolston onto the calendar. Drop Woolston onto the line plot.

(a) Identify the day in the year where SO₂ pollution is the worst.

- (b) Looking at the calendar, is there a trend in the days where the SO_2 air pollution is the worst in Woolston? Describe the trend.

- (c) Click on the day before the worst day in January for SO_2 . At what time was the air pollution the worst on the 72 hour line chart for SO_2 and what was the maximal value for SO_2 ?

Parallel coordinate data comparison

Reload the page before starting this section.

1. Insert three map views. Set the map views to show CO, PM10_BAM and NO_2 . Set the time to 01/01/2012 1:00. Load the data, then insert a Data Comparison plot. Set the parallel coordinate plot to show CO, PM10_BAM and NO_2 , click apply. Then drag and drop Newmarket (In Auckland) onto the parallel coordinate plot. Also drag Henderson_Lincoln Rd (In Auckland) onto the plot.

- (a) Press the play button, observe for about 12 hours. Is there any relationship between PM10_BAM and NO_2 at the Newmarket station? Describe the relationship.

- (b) Is there any relationship between the results for Newmarket and Henderson_Lincoln Rd? Describe the relationship.

Temporal Pattern:

Reload the page before starting this section

1. What are the temporal patterns in air pollutants recorded, and can trends be identified?

- (a) Set the date to 01/01/2012 and choose any station, Use the **Line Plot**, **Site View**, **Heat Calendar** and **Monthly Average Chart** to observe temporal patterns among a selection of air pollutants and meteorological variables. Use the **play button** and describe the patterns on a line plot over time.

- (b) Use the **Site View**, **Monthly Average Chart** and **Heat Calendar** to examine PM10_BAM in Masterton during 2012. Comment on any seasonal trends or days with the highest pollution.

2. What spatial information can be identified about air pollutants?

- (a) Use the **Monthly Average Chart**, **Site View** and **Monthly Rose** to comment on the behavior of different pollutants for any site of your choosing, note that the monthly rose requires wind direction to be present. What site did you choose?

Appendix D

Post-Study Questionnaire

7 When the play button was pressed, was the colour change on the map meaningful for identifying temporal patterns in the data?

Very Ineffective 1 2 3 4 Very Effective 5

8 Would you like to see any functionality added to the site view?

Yes No

Suggestion:

Line Plot

9 Was the text the correct size?

Too Small 1 2 3 4 Too Large 5

10 How much difficulty did you experience using the mouse to navigate?

Very Difficult 1 2 3 4 Very Easy 5

11 Was the use of colour in the line plot effective for interpreting the data?

Very Ineffective 1 2 3 4 Very Effective 5

Heat Calendar

16. How useful was the calendar for identifying high pollution areas of interest?

Very Ineffective 1 2 3 4 Very Effective 5

17. Was the text the correct size?

Very Ineffective 1 2 3 4 Very Effective 5

18. How useful was the calendar for time navigation?

Very Ineffective 1 2 3 4 Very Effective 5

19. Would you like to see any functionality added to the calendar?

Yes No

Suggestion:

20. How much did you feel like you needed assistance with the calendar view?

Very little assistance 1 2 3 4 Very much assistance 5

26. Would you like to see any functionality added to the monthly average plot?

Yes No

Suggestion:

27. How much did you feel like you needed assistance with the monthly average plot?

Very little assistance 1 2 3 4 5 Very much assistance

Data Comparison

28. How useful was the data comparison (parallel coordinate plot) for identifying pollutants of interest?

Very Ineffective 1 2 3 4 5 Very Effective

29. Was the text the correct size?

Too Small 1 2 3 4 5 Too Large

30. How much difficulty did you experience using the mouse to navigate?

Very Difficult 1 2 3 4 5 Very Easy

39. Comment on any unusual aspect(s) of the dataset inspected.

40. Did you identify any temporal patterns in the air quality data set provided? If yes then give a brief description.

41. Did you feel you could discuss the data better after using the Atmovis tool?

Yes

No

42. What aspect of the Atmovis tool helped you the most with understanding the dataset?

43. Would the Atmovis tool be useful in a presentation to demonstrate air quality information?

Yes

No

If your answer was yes then where would you be likely to use the tool:

How could the tool be improved from a presentation perspective?

44. Did you feel more engaged with the task when using the Atmovis tool compared to using a spreadsheet?

Yes

No

45. Would you prefer to use the Atmovis interface over a spreadsheet for any data analysis tasks?

Yes

No

If yes then please specify what tasks:

46. Additional Suggestions or comments:

Thank you for participating

Appendix E

Instructional Slides






Introduction

The purpose of this research project is to design and build an effective prototype for visualizing spatial-temporal data from multiple sources related to air quality. The effectiveness of the prototype will be evaluated by user study. The prototype system will allow analysts to understand trends between different monitoring stations more effectively.

videos

The AtmoVis system consists of several different types of visualization which work together to visualize the data.

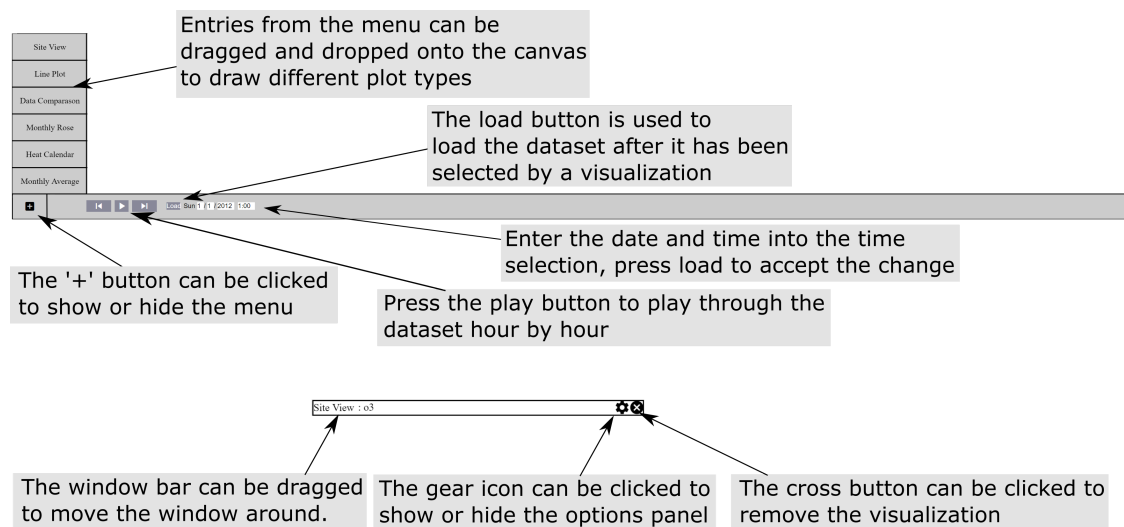
Here is a selection of tutorial videos to explain the different visualizations

-  Loading Data
-  Line Plot
-  Data Comparison
-  Monthly Rose
-  Calendar And Monthly Average

<
<

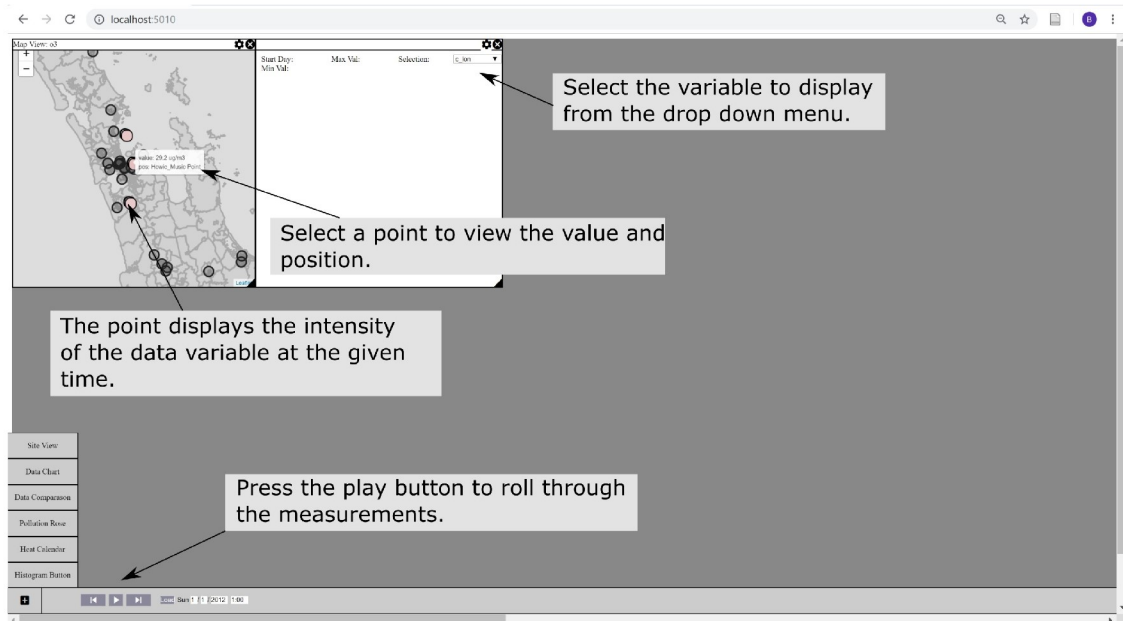
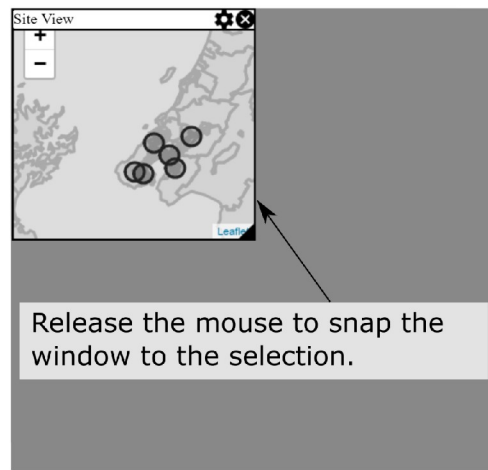
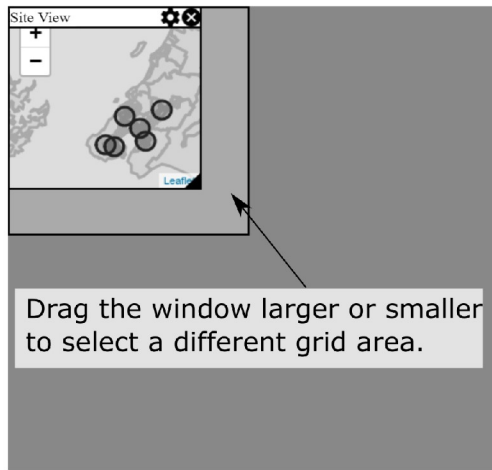
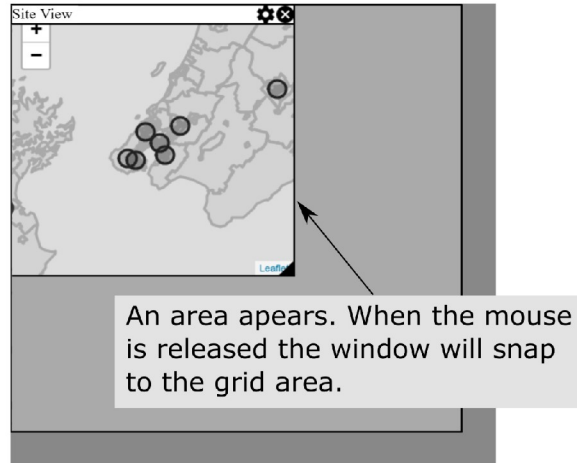
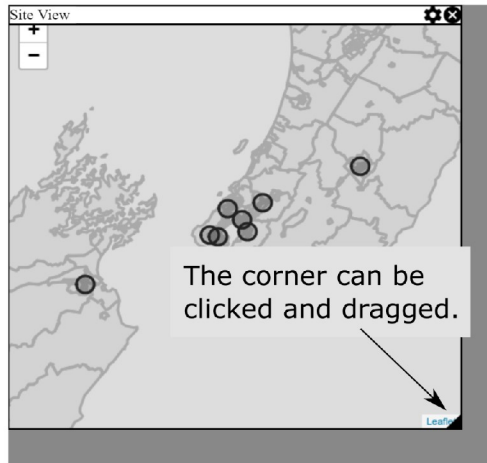
Instructions

The instructions for using AtmoVis are provided as a series of slides with each slide describing the usage of a visualization. The main layout consists of a menu launcher, a canvas and a play button with the date and time.

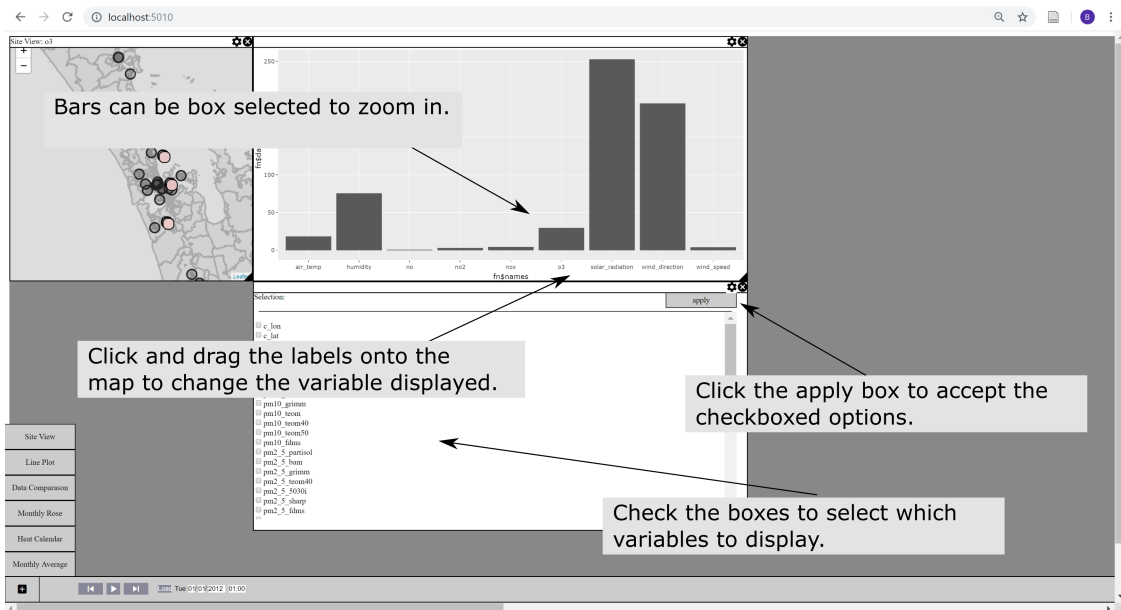
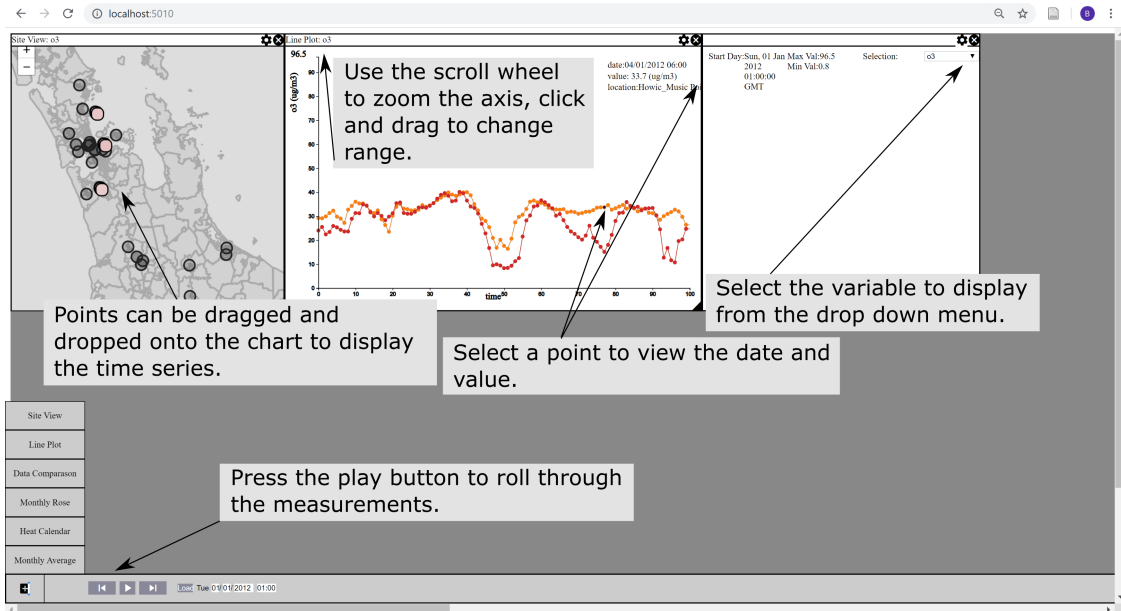


Window Resizing

The window system is snapped to a grid, windows can be moved around and scaled.



Map: Area Unit 2015_V1_00 Clipped [1], © Stats NZ Geographic Data Service used under license, Creative Commons Attribution 3.0 New Zealand [2]



The screenshot shows a web browser window at localhost:5010 displaying a 'Data Comparison' tool. The interface includes a line plot on the left, a list of variables in the center, and a zoomed-in view of the plot on the right. A sidebar on the bottom left contains navigation options: Site View, Line Plot, Data Comparison (selected), Monthly Rose, Heat Calendar, and Monthly Average. A status bar at the bottom shows 'Run 11 of 12012'.

Annotations:

- Click the apply button to accept the variable selection.** (Points to the 'apply' button in the top right of the variable list panel.)
- Check the boxes to select which variables to display on the axes.** (Points to the checkboxes next to the variable names in the list.)
- Use the scroll wheel close to the axis to zoom in, drag axis to move the range.** (Points to the axes of the top plot.)
- Points can be dragged and dropped from the map onto the chart to display the parallel coordinate plot.** (Points to the zoomed-in plot area.)
- Press the play button to roll through the measurements and draw lines on the plot.** (Points to the play button in the bottom status bar.)

Bibliography

- [1] Area unit 2015_v1_00 clipped, <https://datafinder.stats.govt.nz/layer/87752-area-unit-2015-v1-00-clipped/>. Accessed: 04/07/2019.
- [2] Creative commons attribution 3.0 new zealand, <https://creativecommons.org/licenses/by/3.0/nz/>. Accessed: 04/07/2019.
- [3] Gapminder Tools, <https://www.gapminder.org/>. Accessed: 04/07/2019.
- [4] Icons, <https://material.io/tools/icons/?style=baseline>. Accessed: 04/07/2019.
- [5] Leaflet — an open-source JavaScript library for interactive maps, <https://leafletjs.com/>. Accessed: 04/07/2019.
- [6] NIWA, <https://www.niwa.co.nz/>. Accessed: 04/07/2019.
- [7] Niwa Weather, <https://weather.niwa.co.nz/>. Accessed: 04/07/2019.
- [8] Openstreetmap, <https://www.openstreetmap.org/copyright>. Accessed: 04/07/2019.
- [9] plumber, <https://www.rplumber.io/>. Accessed: 04/07/2019.
- [10] pyproj, <https://pypi.org/project/pyproj/>. Accessed: 04/07/2019.

- [11] Python Data Analysis Library — pandas: Python Data Analysis Library, <https://pandas.pydata.org/>. Accessed: 04/07/2019.
- [12] Vega Specification, <https://vega.github.io/vega/docs/>. Accessed: 04/07/2019.
- [13] Welcome to Flask — Flask 1.0.2 documentation, <http://flask.pocoo.org/docs/1.0/>. Accessed: 04/07/2019.
- [14] New Zealand’s Environmental Reporting Series: Environment Aotearoa 2015. In *New Zealand’s Environmental Reporting Series: Environment Aotearoa 2015*. Ministry for the Environment & Statistics New Zealand, 2015, pp. 27–36.
- [15] 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.
- [16] AIKEN, A., CHEN, J., STONEBRAKER, M., AND WOODRUFF, A. Tioga-2: A Direct Manipulation Database Visualization Environment. In *Proceedings of the Twelfth International Conference on Data Engineering* (Washington, DC, USA, 1996), ICDE ’96, IEEE Computer Society, pp. 208–217.
- [17] ANDREWS, C., ENDERT, A., AND NORTH, C. Space to Think: Large High-resolution Displays for Sensemaking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2010), CHI ’10, ACM, pp. 55–64. event-place: Atlanta, Georgia, USA.
- [18] BARNETT, A. G., WILLIAMS, G. M., SCHWARTZ, J., BEST, T. L., NELLER, A. H., PETROESCHEVSKY, A. L., AND 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.

- [19] BOSTOCK, M., AND HEER, J. Protovis: A Graphical Toolkit for Visualization. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov. 2009), 1121–1128.
- [20] BOSTOCK, M., OGIEVETSKY, V., AND HEER, J. D3: Data-Driven Documents. *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)* (2011).
- [21] BUDIU, R. Between-Subjects vs. Within-Subjects Study Design, <https://www.nngroup.com/articles/between-within-subjects/>, May 2018. Accessed: 04/07/2019.
- [22] CALDWELL, J. Air quality monitoring report for Hamilton, Tokoroa, Taupo, Te Kuiti, Putaruru, Turangi, Cambridge and Te Awamutu-Kihikihi - 2014, Oct. 2015.
- [23] CARSLAW, D. C., AND ROPKINS, K. openair — An R package for air quality data analysis. *Environmental Modelling & Software* 27–28, 0 (2012), 52–61.
- [24] CARTO. The World’s Leading Location Intelligence Platform — CARTO, <https://carto.com/>. Accessed: 04/07/2019.
- [25] CHANG, W., CHENG, J., ALLAIRE, J. J., XIE, Y., AND MCPHERSON, J. *shiny: Web Application Framework for R*. 2018. Accessed: 04/07/2019.
- [26] CHANG, Y.-N., LIM, Y.-K., AND STOLTERMAN, E. Personas: From Theory to Practices. In *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges* (New York, NY, USA, 2008), NordiCHI ’08, ACM, pp. 439–442. event-place: Lund, Sweden.
- [27] CLARKSON, E., DESAI, K., AND FOLEY, J. ResultMaps: Visualization for Search Interfaces. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov. 2009), 1057–1064.

- [28] FEKETE, J.-D., VAN WIJK, J. J., STASKO, J. T., AND NORTH, C. The Value of Information Visualization. In *Information Visualization: Human-Centered Issues and Perspectives*, A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, Eds. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 1–18.
- [29] GARCIA, R., HART, J. E., DAVIS, M. E., REASER, P., NATKIN, J., LADEN, F., GARSHICK, E., AND SMITH, T. J. Effects of Wind on Background Particle Concentrations at Truck Freight Terminals. *Journal of occupational and environmental hygiene* 4, 1 (Jan. 2007), 36–48.
- [30] HAFEN, R., GOSINK, L., MCDERMOTT, J., RODLAND, K., DAM, K. K. V., AND 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.
- [31] HALES, S., SALMOND, C., TOWN, G. I., KJELLSTROM, T., AND 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.
- [32] HSU, Y.-C., DILLE, P., CROSS, J., DIAS, B., SARGENT, R., AND NOURBAKHSH, I. Community-Empowered Air Quality Monitoring System. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), CHI '17, ACM, pp. 1607–1619.
- [33] ISENBERG, T., ISENBERG, P., CHEN, J., SEDLMAIR, M., AND MÖLLER, T. A Systematic Review on the Practice of Evaluating Visualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2818–2827.

- [34] JAMES, T. State Of The Environment Ambient Air Quality Monitoring Programme Protocols & Monitoring Site Details. *Tasman Council* (Aug. 2008), 1–61.
- [35] KIRK, A. Data visualisation : a handbook for data driven design / Andy Kirk. In *Data visualisation : a handbook for data driven design / Andy Kirk*. Los Angeles : Sage Publications, 2016, p. 185.
- [36] KOUA, E. L., MACEACHREN, A., AND KRAAK, M. J. Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographical Information Science* 20, 4 (Apr. 2006), 425–448.
- [37] LI, H., FAN, H., AND 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).
- [38] LI, N., JIANG, Z., LIU, Z., AND MENG, X. A Method of Hierarchical Time-series Data Visualization. In *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction* (New York, NY, USA, 2013), VINCI '13, ACM, pp. 113–114.
- [39] LIU, J., LI, W., WU, J., AND LIU, Y. Visualizing the intercity correlation of PM2.5 time series in the Beijing-Tianjin-Hebei region using ground-based air quality monitoring data. *PLOS ONE* 13, 2 (2018), 1–14.
- [40] LIU, Y., BARLOWE, S., FENG, Y., YANG, J., AND JIANG, M. Evaluating exploratory visualization systems: A user study on how clustering-based visualization systems support information seeking from large document collections. *Information Visualization* 12, 1 (2013), 25–43.
- [41] MARCUS, A., COMORSKI, D., AND SERGEYEV, A. Supporting the evolution of a software visualization tool through usability studies. In

- 13th International Workshop on Program Comprehension (IWPC'05)* (May 2005), pp. 307–316.
- [42] MUNZNER, T. *Visualization Analysis and Design*. AK Peters Visualization Series. CRC Press, 2014.
- [43] NIELSEN, J. *Usability Engineering*. Academic Press, Inc., 1300 Boylston Street, Chestnut Hill, MA 02167, 1993.
- [44] NIELSEN, L. Persona Writing. In *Personas - User Focused Design*. Springer London, London, 2019, pp. 55–81.
- [45] OTHMAN, N., MATJAFRI, M. Z., LIM, H. S., AND ABDULLAH, K. Retrieval of Aerosol Optical Thickness (AOT) and its Relationship to Air Pollution Particulate Matter (PM10). In *2009 Sixth International Conference on Computer Graphics, Imaging and Visualization* (Aug. 2009), pp. 516–519.
- [46] PETERS, V. L., AND SONGER, N. B. Evaluating the Usability of an Interactive Map Activity for Climate Change Education. In *Proceedings of the 10th International Conference on Interaction Design and Children* (New York, NY, USA, 2011), IDC '11, ACM, pp. 197–200.
- [47] PRUITT, J., AND ADLIN, T. *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.
- [48] ROGERS, Y. HCI Theory: Classical, Modern, and Contemporary. *Synthesis Lectures on Human-Centered Informatics* 5, 2 (2012), 1–129.
- [49] ROTH, S. F., LUCAS, P., SENN, J. A., GOMBERG, C. C., BURKS, M. B., STROFFOLINO, P. J., KOLOJECHICK, A. J., AND 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.
- [50] SANTOS, B. S., DIAS, P., SILVA, S., FERREIRA, C., AND MADEIRA, J. Integrating User Studies into Computer Graphics-Related Courses. *IEEE Computer Graphics and Applications* 31, 5 (Sept. 2011), 14–17.
- [51] SATYANARAYAN, A., AND 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.
- [52] 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.
- [53] SHNEIDERMAN, B., AND 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.
- [54] SIEVERT, C. *plotly for R*, <https://plotly-book.cpsievert.me>. 2018. Accessed: 04/07/2019.
- [55] SMITH, G., CZERWINSKI, M., MEYERS, B., ROBBINS, D., ROBERTSON, G., AND TAN, D. S. FacetMap: A Scalable Search and Browse Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (Sept. 2006), 797–804.
- [56] SPENCE, R. *Information Visualization: Design for Interaction (2nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2007.

- [57] STOLTE, C., TANG, D., AND HANRAHAN, P. Polaris: A System for Query, Analysis, and Visualization of Multidimensional Databases. *Commun. ACM* 51, 11 (Nov. 2008), 75–84.
- [58] STONE, M. *Field Guide to Digital Color*. A. K. Peters, Ltd., Natick, MA, USA, 2003.
- [59] TAK, S. *Understanding and Supporting Window Switching*. Doctor of Philosophy, University of Canterbury, 2011.
- [60] THOMAS, J., AND COOK, K. Illuminating the path: the research and development agenda for visual analytics. *Los Alamitos, CA: IEEE Computer* (2005).
- [61] TORY, M., AND MOLLER, T. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics* 10, 1 (Jan. 2004), 72–84.
- [62] WEI, F., LIU, S., SONG, Y., PAN, S., ZHOU, M. X., QIAN, W., SHI, L., TAN, L., AND ZHANG, Q. Tiara: a visual exploratory text analytic system. In *In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '10* (2010), ACM, pp. 153–162.
- [63] WILKINSON, L., ANAND, A., AND GROSSMAN, R. High-Dimensional Visual Analytics: Interactive Exploration Guided by Pairwise Views of Point Distributions. *IEEE Transactions on Visualization and Computer Graphics* 12, 6 (Nov. 2006), 1363–1372.
- [64] WONG, D. W., YUAN, L., AND PERLIN, S. A. Comparison of spatial interpolation methods for the estimation of air quality data. *Journal Of Exposure Analysis And Environmental Epidemiology* 14 (Sept. 2004), 404–415.

- [65] WONGSUPHASAWAT, K., MORITZ, D., ANAND, A., MACKINLAY, J., HOWE, B., AND HEER, J. Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan. 2016), 649–658.
- [66] YI, J. S., KANG, Y. A., STASKO, J., AND JACKO, J. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov. 2007), 1224–1231.
- [67] ZHANG, J., AND MARCHIONINI, G. Evaluation and Evolution of a Browse and Search Interface: Relation Browser++. In *Proceedings of the 2005 National Conference on Digital Government Research* (Atlanta, Georgia, USA, 2005), dg.o '05, Digital Government Society of North America, pp. 179–188.
- [68] ZHENG, Y., LIU, F., AND 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.
- [69] ZHOU, M., WANG, R., MAI, S., AND TIAN, J. Spatial and temporal patterns of air quality in the three economic zones of China. *Journal of Maps* 12, sup1 (2016), 156–162.