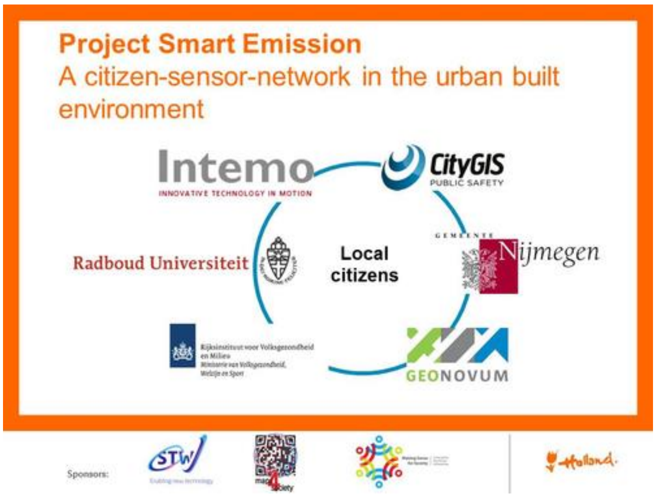
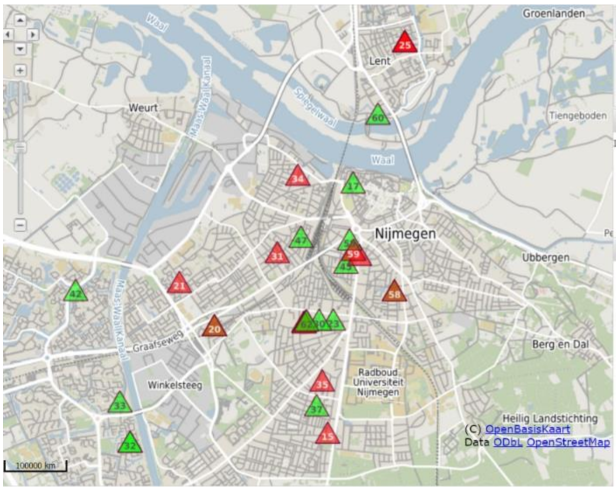
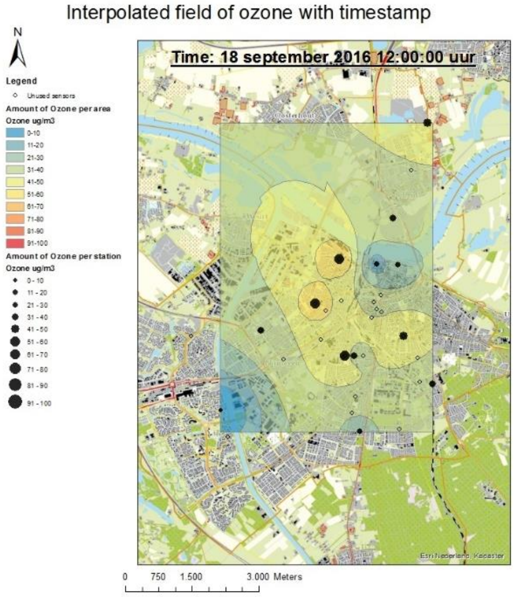
Visualizations in time and space with citizen sensor network data

As part of the Pilot study “smart emission” from the Radboud University Nijmegen





Interpolated field of ozone with timestamp



Name: Rick Jankowski
Student number: 2555127 (VU)
Internship: Smart Emission project
Link to the project: http://smartemission.ruhosting.nl/

Summary

1. Introduction

1.1 The research

The Smart Emission project is conducted at the Radboud University Nijmegen and is executed by a consortium of Dutch knowledge institutes, government and ICT-companies together with the citizens of the city of Nijmegen led by Dr. Ir. Linda Carton of the Radboud University Nijmegen. The goal of this project is to monitor, visualize and communicate a real-time “environmental footprint” of the city. To accomplish this a Open Geo Data infrastructure has been developed in combination with low-cost outdoor sensors. In this whole project the citizens play a participatory role as they can collaborate with the consortium professionals to make sense of all the data. (Carton, 2015) The data that is produced by the sensors are available through different viewers. In this research the Heron Viewer made by Just van den Broecke is used.

This particular report concerns itself with the question if the data collected by the sensors can actually be used to make useful visualisations and if the results can help the citizens and be used for the collective sense-making the project seeks. In particular this experimental approach uses the data of Ozone. This research is done in context of a research assignment that is part of the National GI-minor of the Vrije Universiteit of Amsterdam.

1.2 The study area

The study area is the city of Nijmegen. A collection of low-cost (Jose) sensors are distributed throughout the city with some clusters around points of interest. The sensors are placed in the front or backyard of the participants, this is already part of the participatory part of the project.

fig. 1.1 Placement of the city of Nijmegen on a map of the Netherlands. (source: www.turkey-visit.com)

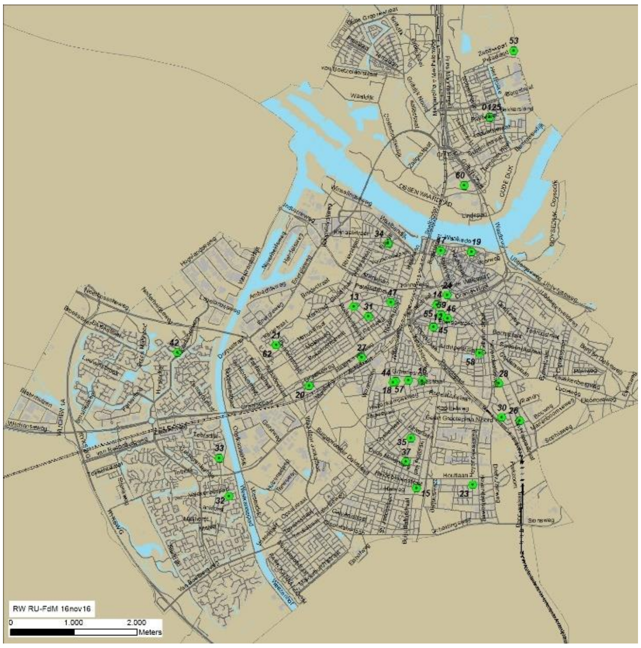


fig. 1.2 Distribution of sensors around the city of Nijmegen. (made by R. Wunderink)

1.3 The objective

The objective of this research is to see if it is to make useful visualisations with the data that can be extracted from the sensors throughout the city and to see if this can lead, together with the citizens, for a better understanding of the dynamic of the city, with an emphasis on the night/day cycle of pollutants. How this will be accomplished will be discussed further in this report.

1.4 The structure

First there will be a small literature review and the plan is discussed. Then the types and quality of the data will be dealt with, after that the data will be analysed and the results will be addressed. The end will be the discussion and conclusions where the shortcomings will be defined and the results summarized.

2. Literature review & plan

In this chapter the relevance of the research will come forward through a literature review. With the review in mind a plan is made that is relevant for the current research.

2.1 Literature

In literature, there already has been extensive covering of the use of GIS, however “maps and other analyses simply compare a limited number of particular moments or intervals rather than take advantage of the full structure of time” (Andrienko et. al. 2010). GIS seems to lack on certain fronts to visualize the temporal dimension. According to Monmonier (1990) there are three ways to show a temporal aspect on a map. There are “dance maps”, “chess maps” and “change maps”.

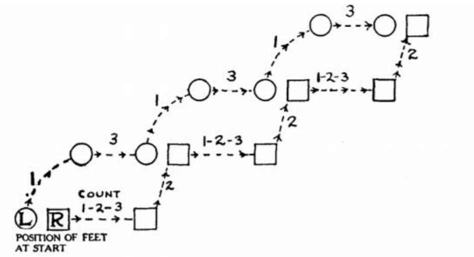


FIGURE 16. *A prototypic dance map showing the woman's steps for the Hesitation Waltz. [Source: Walker, Caroline, 'The Modern Dances: How to Dance Them,' Chicago, Saul Brothers, 1914, p. 46.]*

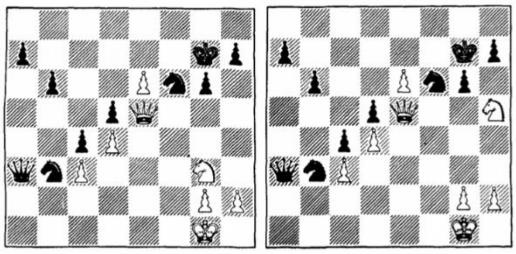


FIGURE 17. *Chess maps are two or more geographic-space displays justaposed so that the viewer can compare spatial patterns for different times.*

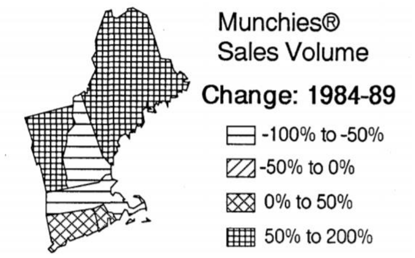


FIGURE 18. *A change map showing rate of change by state.*

fig. 2.1 (upper left): a dance map showing a time series in one image. (Monmonier, 1990)

fig 2.2 (upper right): a chess map showing two maps side by side to show the different times. (Monmonier, 1990)

fig 2.3 (lower left): a change map showing the change in unit during a certain time interval. (Monmonier, 1990)

These maps however do not use the time-space aspect in its full potential. There are two ways to make full use of the temporal dimension in the data. This can be done by a temporal scroll bar or by a map animation. With the temporal scroll bar, you can enable the user to scroll through the whole time series, this way a dynamic map is created. Because of the fact that the participant can control the time by himself will make him think more and will keep the attention longer than a static map (Monmonier, 2010). The other strategy to show the temporal dimension as well as possible is with animation graphics. With this technique, several static maps will be put together as separate frames so they will show as one flow on screen.

For a good visualisation of air pollutants with GIS there are three main options, simple locational, interpolation and dynamic modelling. The group of simple locational methods contains functions such as buffer or point in polygon. With these functions, there is an even area created around a point with a value. For air pollution, these methods are far from useful. The air (and with it air quality) is a continuous phenomenon so this cannot really be visualized by evenly distributed buffers.

To fix that shortcoming the interpolation method comes into play. This method uses an interpolation method to make a continuous field that takes into account the different values of different points. Depending on the type of interpolation there are of course other shortcomings. Air pollution is more complex than the assumptions made by the interpolation algorithm but at best they describe some spatial variation in a more continuous map.

The third option is dynamic modelling. Dynamic modelling is an advanced type of interpolation where there are functions added which ‘predict’ the movement of a pollutant in this case. These methods can be effective when there are different sources of the pollutant and the dispersion of the pollutants are influenced by buildings or other obstacles. The complexity of these models is also what makes them time consuming and for a general understanding not always necessary. (Briggs, 2005)

2.2 Plan

Looking at the literature, to make best use of the temporal data that is available, the optimal solution would be to make a time slider or an animation graphic, this way the citizens will look more closely and they can come together to a collective sense making.

The animation will be of a time-series of a week to make optimal use of the (possible) daily patterns of the measured gas. In the animation or temporal scroll bar an interpolation method will be used to incorporate the different measurement station throughout the city to create a continuous field.

3. Data & Method

3.1 Data source

The measurements are collected by a Jose sensor (fig 3.1), the sensor measures every 10 seconds. In most of the viewers the measurements come out as an hourly average. The output of an interval is shown in table 3.1. This table (table 3.1) is the output of the Whale viewer (fig 3.2). This is the easiest way to collect large amounts of data at once. The whale viewer allows the user to set an interval for which the data is needed, this has to be done per sensor and also per pollutant.

The objective is to make it understandable for the citizens so if they want they could do it by themselves, that is why this will be the only data source discussed.

gid raw	insert time	device id	name	label	unit	time	day	hour	value min	value max	value raw	value	sample count	altitude
87945	2016-09-10T00:05:02.575	42	o3	O3	ug/m3	2016-09-10T00:00:00	20160909	22	33	39	7870	36	13	0
87965	2016-09-10T01:05:02.638	42	o3	O3	ug/m3	2016-09-10T01:00:00	20160909	23	37	39	6477	38	12	0
87985	2016-09-10T02:05:02.897	42	o3	O3	ug/m3	2016-09-10T02:00:00	20160910	0	36	39	5488	38	19	fs
88011	2016-09-10T03:20:03.022	42	o3	O3	ug/m3	2016-09-10T03:00:00	20160910	1	44	57	6736	52	78	0
88033	2016-09-10T04:20:02.231	42	o3	O3	ug/m3	2016-09-10T04:00:00	20160910	2	47	56	7004	52	71	0
88053	2016-09-10T05:20:02.434	42	o3	O3	ug/m3	2016-09-10T05:00:00	20160910	3	23	38	10234	34	52	0
88073	2016-09-10T06:20:02.928	42	o3	O3	ug/m3	2016-09-10T06:00:00	20160910	4	51	55	7527	53	32	0
88166	2016-09-10T11:05:02.676	42	o3	O3	ug/m3	2016-09-10T11:00:00	20160910	9	67	67	59003	67	11	0
88194	2016-09-10T12:20:02.675	42	o3	O3	ug/m3	2016-09-10T12:00:00	20160910	10	65	67	67642	66	36	0

Table 3.1 Example of the output for ozone from the Whale Viewer of the interval 00:00-12:00 of sensor 42 on 10-09-2016

fig. 3.1 A Jose Sensor attached to a roof.
(source: http://www.leefmilieu.nl/zelf-meten)

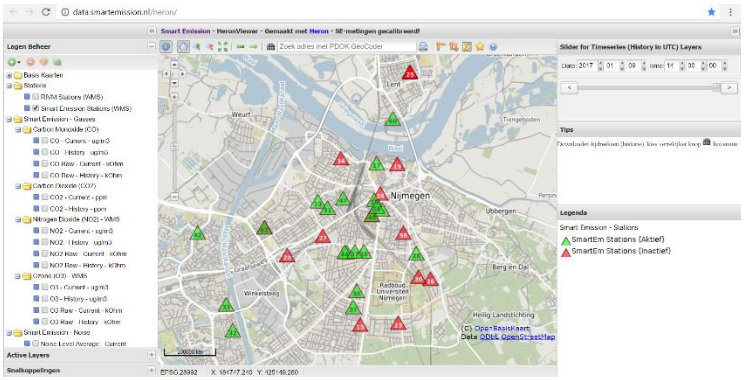


fig. 3.2 The desktop view of the Heron Viewer.
(source: Data.smartemission.nl/heron/)

3.2 Method

To be able to assess the data quality the method has to be clear first, this way you can discuss if the data is fit for purpose.

The objective is to make a visualisation of the dynamic of the city that can speak to the citizens so it is easier to understand. To do this an animation will be made of a week for the gas ozone. Ozone is chosen mainly because the measurements of ozone are well calibrated, the other pollutants are not yet calibrated so these would give unreliable measurements. For the animation, a map has to be made per hour and these maps will then be joined to make an animation of a whole week. This is partly done by using a model in the model builder function of ArcGIS.

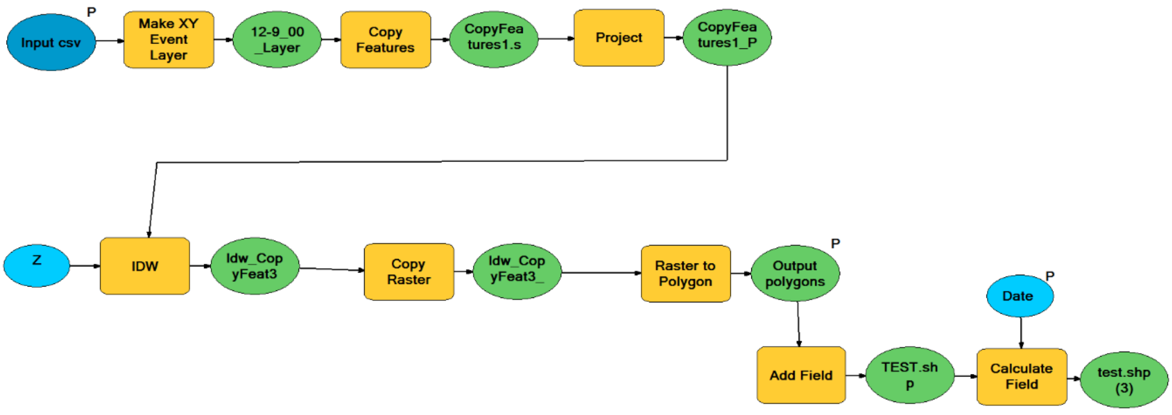


fig. 3.3 Model to go from a CSV-file to a interpolated map using the available sensors.

For this model to work the input data has to be of a certain format. The value of every measurement station on one particular hour has to be in an individual CSV-file. To be able to be able to put all the individual hourly maps together a timestamp has to be created, this is done by adding a field with a unique time code for every map.

3.3 Data quality

The data that is used is the ozone measurement of every sensor. Ozone has been thoroughly calibrated and therefore the most reliable. The other pollutants will not give a reliable result. There are however several points which could need attention, or can make the result look rather questionable:

- The number of sensors that measure at any given time.
  In the city of Nijmegen there are 33 sensors active, however the number of sensors that give a measurement at any given time is much lower, around 11. This is partly due to the WiFi connection the sensor is relying on.
- The sensors do not hang in the same place at every location. Some sensors are standing in the front- or back garden and others are hanging on the wall of the house. This can also affect the measurements every sensor gives, at least when a comparison has to be done.
- The Whale Viewer only gives values per hour, this way the individual piques are not time stamped and cannot be identified.
- The Whale Viewer output layout is not optimal. It is time-consuming to change the layout of the viewer output to the layout that the model needs as input. The input needs to be per hour all sensors in one file.

The only data needed for this application is the sensor number, the time and an hour mean value. The other values can be deleted. So, for this application the viewer output could be reduced. The pilot study has been approved to continue so with that the improvements and calibrations to the sensors. This way the measurements will become more reliable as the project continues.

4. Analysis & results

The whole interpolation has been made for week 37 of 2016, this is the 12th of september till the 18th of september. This was a week where there were relatively high temperatures, this results in higher ozone values. Because the animation cannot be shown here a chess map will be shown of a small interval to give a sense of what the animation looks like. For the whole animation the youtube channel of Smart Emission can be visited (link: https://youtu.be/cAUf_mca7As)

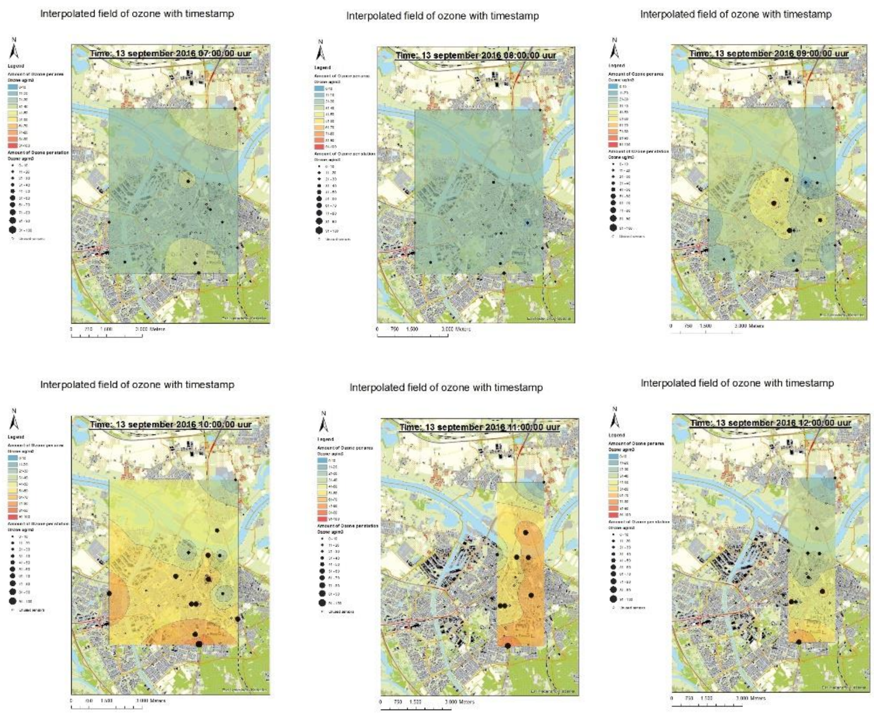


fig. 5.1 Chess map of the interpolated fields using the measurements of ozone for the interval 7:00 to 12:00 on September 13th.

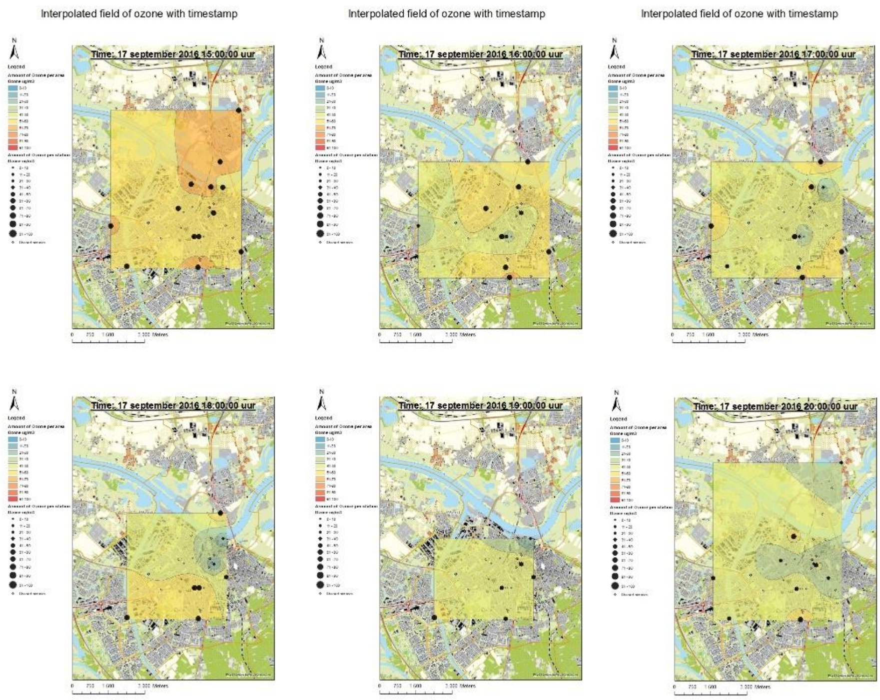


fig. 4.2 Chess map of the interpolated fields using the measurements of ozone for the interval 15:00 to 20:00 on September 17th .

).

In figure 4.1 the interval 07:00-12:00 is shown. The first transition from relatively lower values to higher values can be seen here around 9:00. This is the moment where the sun has already heated up the surface so the ozone starts forming more. The maps also show relatively higher value in the area in the south. These maps are mostly for the visualisation only, most of the interpretations will be left to the experts.

In figure 4.2 the other transition can be made out. This is in the beginning of the evening when the sun gives less heat to the surface. The southern area of the map still has a relatively higher value so this could mean that there is an external source of ozone here.

In the interpolated fields the higher values can be seen due to the color scheme that has been applied.

This shows that the animation can indeed be used to show the day/night cycle of ozone. This can be seen even better in the full animation. In the ArcGis software there is a temporal scroll bar available that would enable the user to decide what time he wants to see, but this technique does not translate well to an app that the participants can fully use.

There are however several sidenotes to the interpolations.

- By using the Inverse distance weighing interpolation method the sensors became the source of the pollutant. This means that the sensor becomes a point-source instead of the real source that is close by. This can also mean that the value close by the sensor could be higher but is not measured this way.
- Like mentioned already discussing the data quality, not all sensors give a measurement at all times. This limits the precision of the interpolation. This can be seen in the last two images of fig. 4.1. The used values have a small spread, because there only will be interpolated, not interpolated, the continuous field is relatively small and does not cover the whole area.
- The sensor stations are calibrated but the measurements are not nationally verified. This means that conclusions made by the data is not admissable in court but can only be used as a advisory instrument.

5. Conclusion & discussion

5.1 Conclusion

As the results show, the day and night cycles of the city can be distinguished using a interpolation method in combination with an animation. Part of this animation was shown to some participants and they were happy with what they saw.

The objective of the project is achieved but there are still a lot of points which can be fixed. This approach for example is good to make visualizations to show to the citizens but it is rather time consuming and most of all not doable with the limited knowledge of GIS software of the citizens to do themselves.

The recommendation for further research is therefore to put the focus on what the citizens themselves can do in a form of open source GIS, like Qgis. This way you can empower the citizens to come up with their own solutions and conclusions so they can share their experience with each other.

5.2 Discussion

There are several shortcomings to this project. First of all, the data available can cause a problem. From the 33 sensors placed throughout the city, at most there are 13 sensors that give data for any given time, this varies however depending of the pollutant. This problem can be solved by placing more sensors throughout the city to get more measurements. Another problem can be the quality of the data in terms of calibration. It is already known that with this data there cannot be any formal decisions made because the measurements are not verified by an institute, it can only be of service as an advisory tool.

It would also be better if the output from the whale viewer would change. With the way it is now the pre-processing of the data before it can be analysed is too long, especially if the citizens themselves are going to do the analysis then the pre-processing should be brought back to almost zero.

References

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., ... & Tominski, C. (2010). Space, time and visual analytics. *International Journal of Geographical Information Science*, *24*(10), 1577-1600.

Briggs, D. (2005). The role of GIS: coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health, Part A*, *68*(13-14), 1243-1261.

Carton, L.J. Et Al (2015), FILLING THE FEEDBACK GAP OF PLACE-RELATED ‘EXTERNALITIES’ IN SMART CITIES: Empowering citizen-sensor-networks for participatory monitoring and planning for a responsible distribution of urban air quality

Elwood, S. (2008). Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal*, *72*(3–4), 173–183.

Logan, J. A. (1985). Tropospheric ozone: Seasonal behavior, trends, and anthropogenic influence. *Journal of Geophysical Research: Atmospheres*, *90*(D6), 10463–10482.

Monks, P. S. (2000). A review of the observations and origins of the spring ozone maximum. *Atmospheric Environment*, *34*(21), 3545–3561.

Monmonier, M. (1990). Strategies for the visualization of geographic time-series data, Cartographica, 27 (1)

Monmonier, M. (2010). Strategies for the Visualization of Geographic Time-Series Data. *Classics in Cartography: Reflections on Influential Articles from Cartographica*, 55-72.

Vingarzan, R. (2004). A review of surface ozone background levels and trends. *Atmospheric Environment*, *38*(21), 3431–3442.