Visual Analytic Design for Detecting Airborne Pollution Sources

VAST Award: Comprehensive MC2 Solution

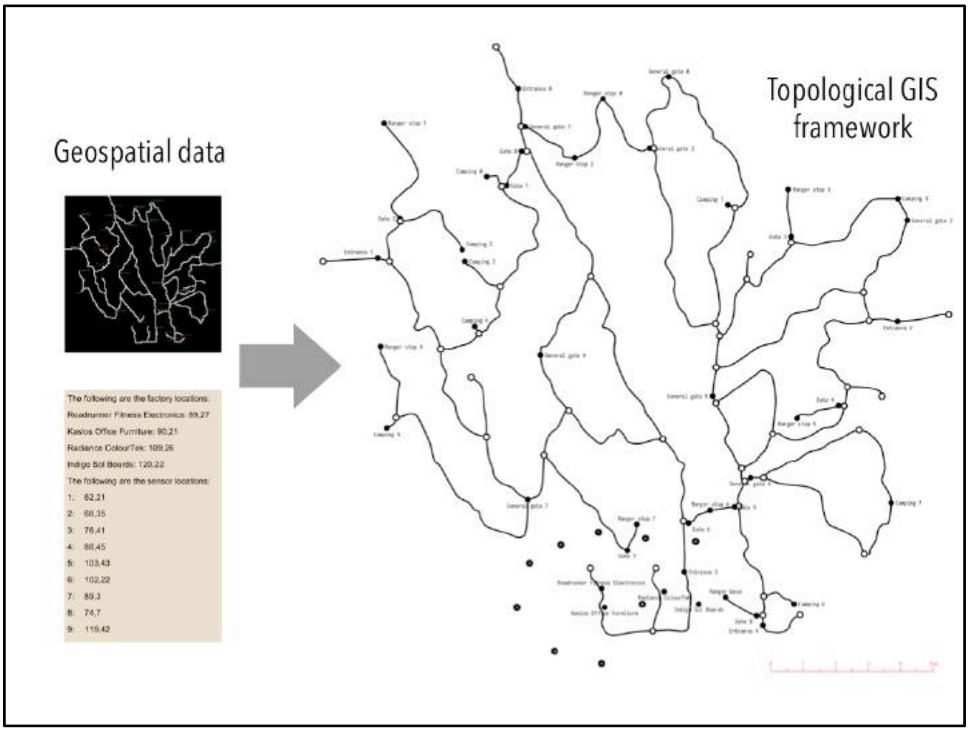
gicentre.net/vastchallenge2017

Johannes Liem and Jo Wood

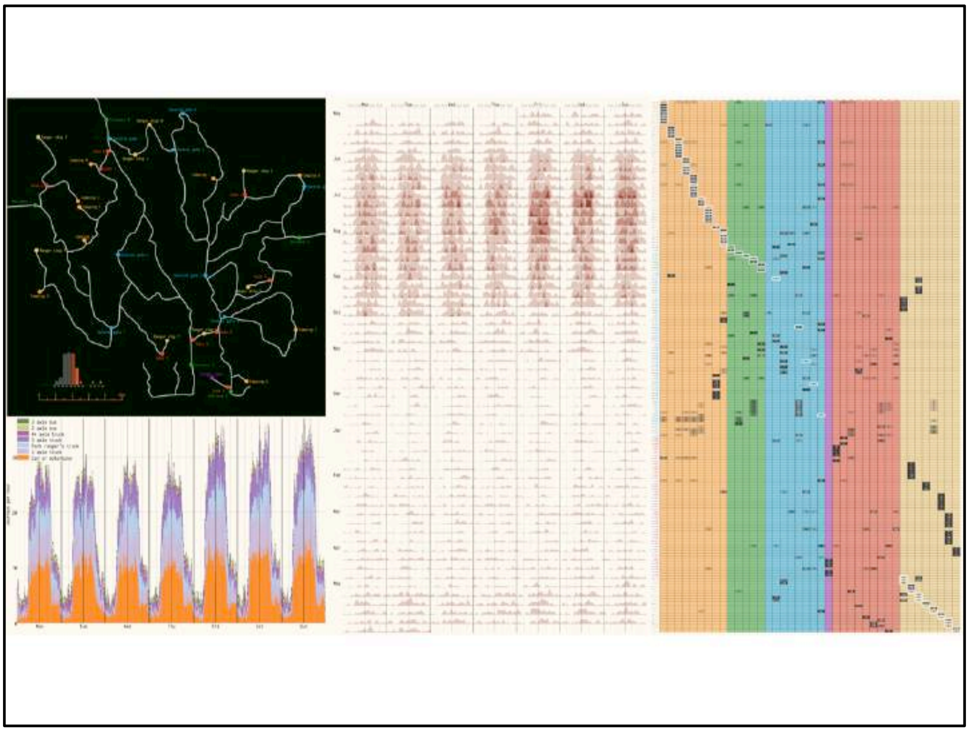
giCentre, City, University of London

johliem jwoLondon

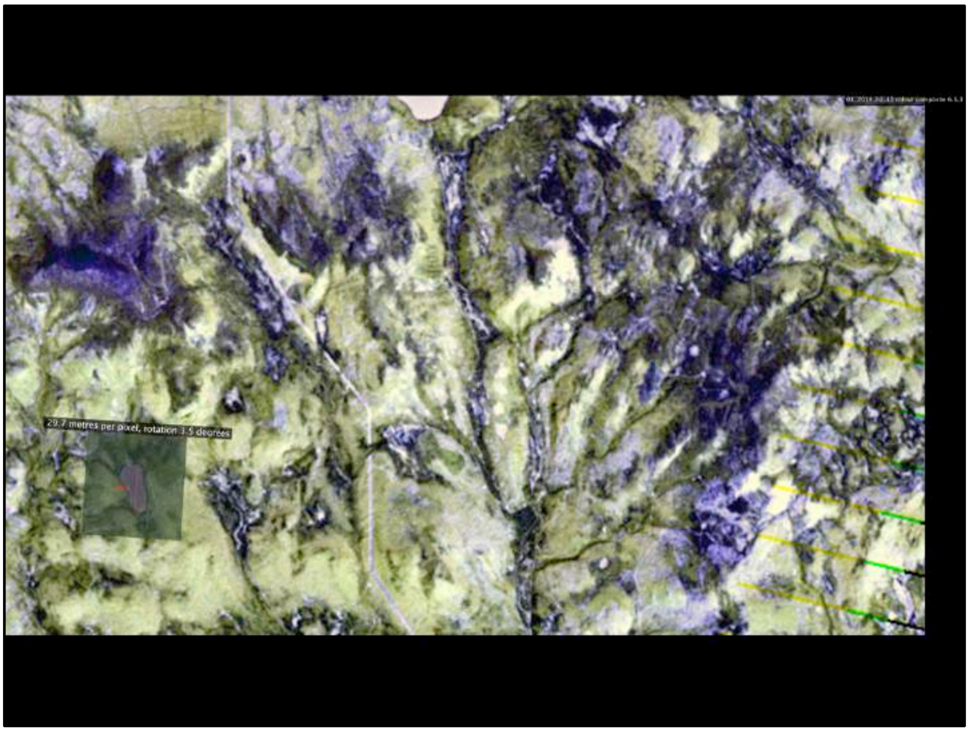
Data frame



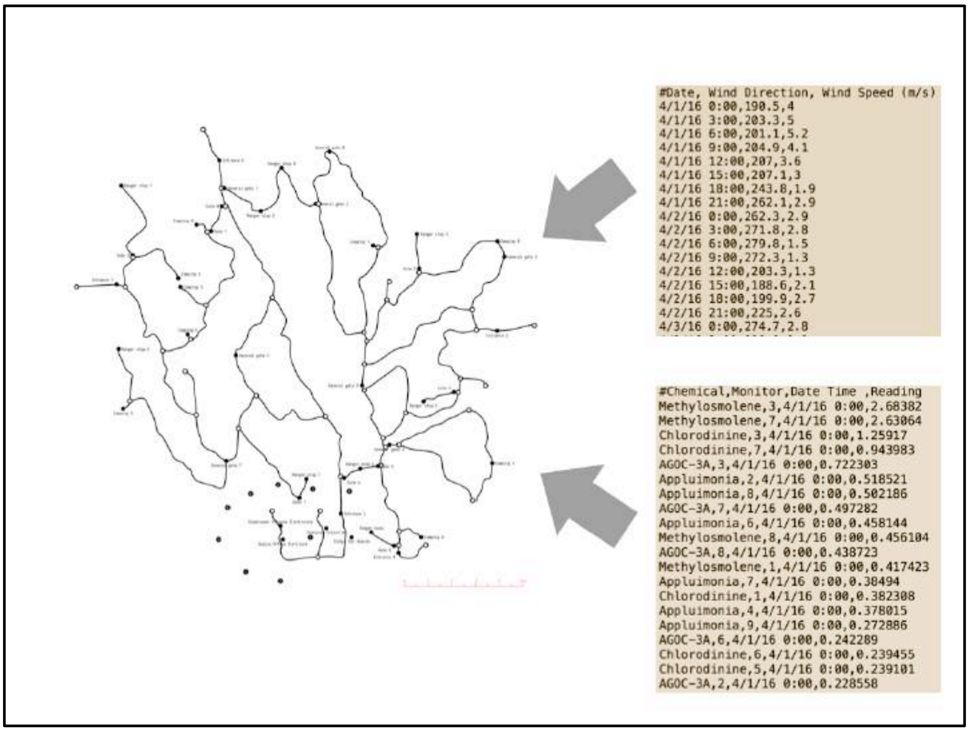
First stage, relevant to all mini challenges, was to create an integrated geospatial database including road network topology and geolocation of all point locations (road junctions, road sensors, air sensors, factory locations).



This allowed the integration of the data into all three mini challenges. Here an example from MiniChallenge 1

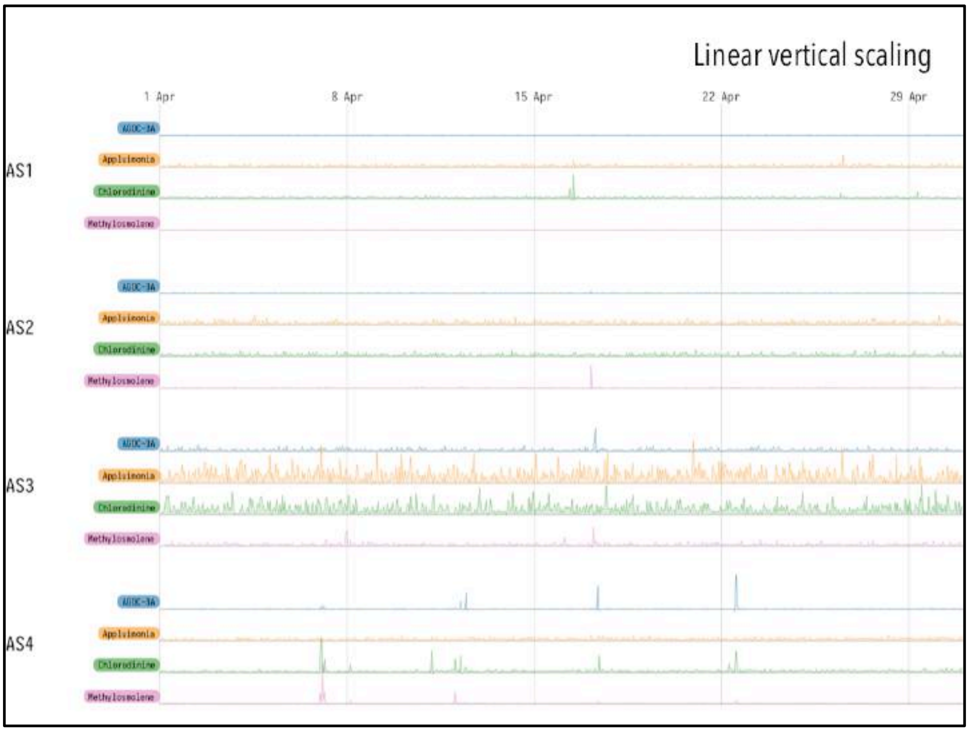


And Mini-Challenge 3 (remote sensing)

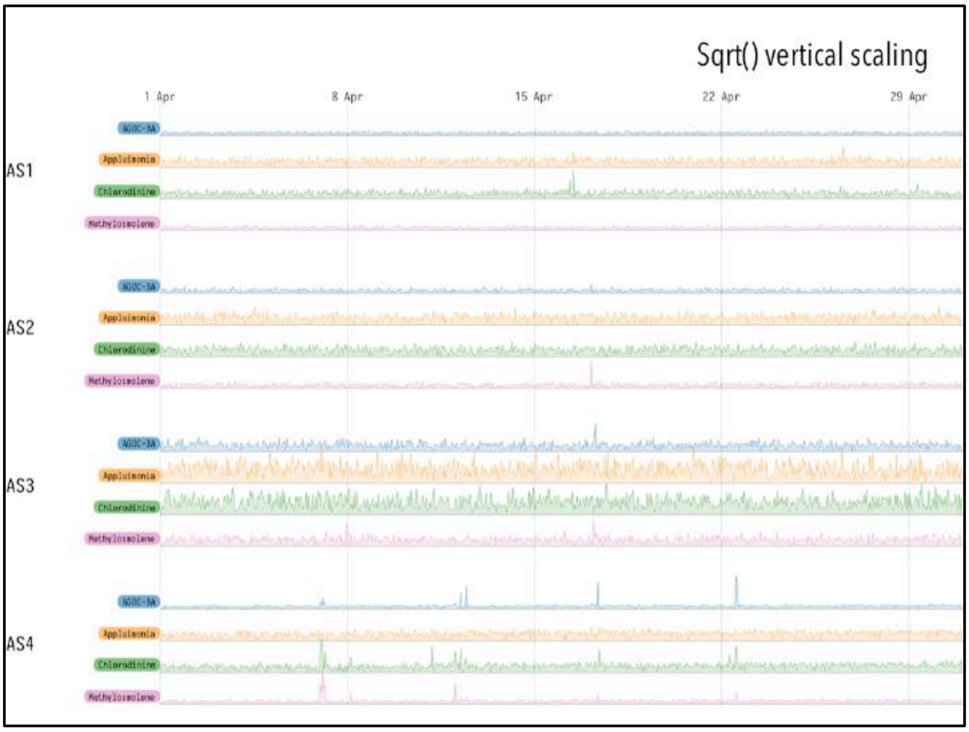


Data specific to mini challenge 2 (wind velocities and sensor readings) were then integrated in this framework

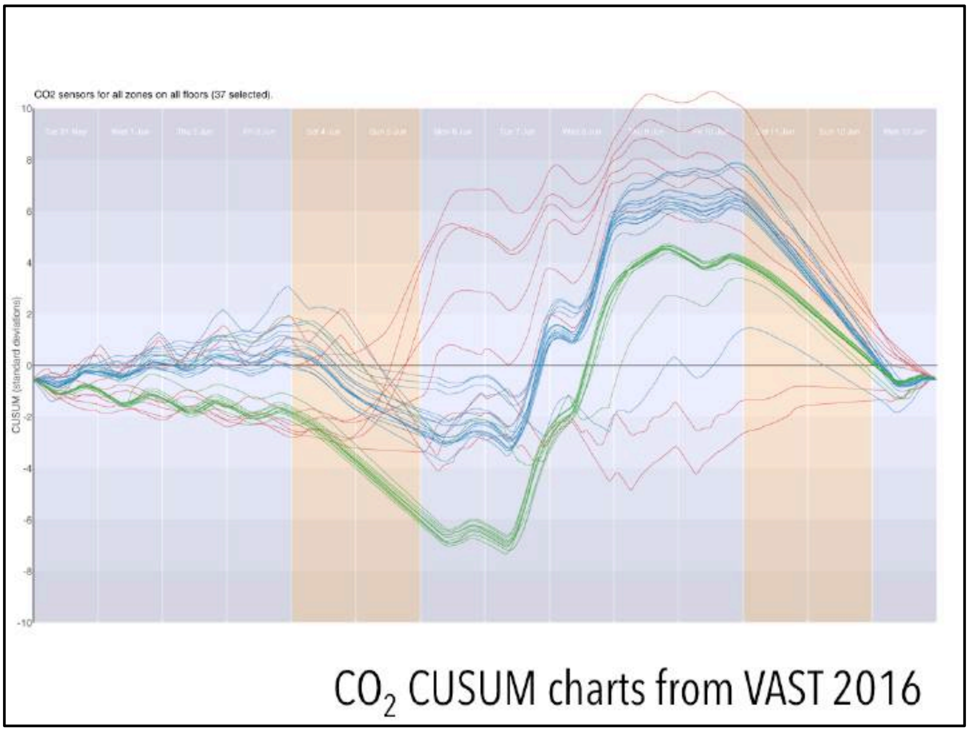
Sensor data cleaning and inspection



Raw sensor data with vertical scaling to maximum reading in parts per million.



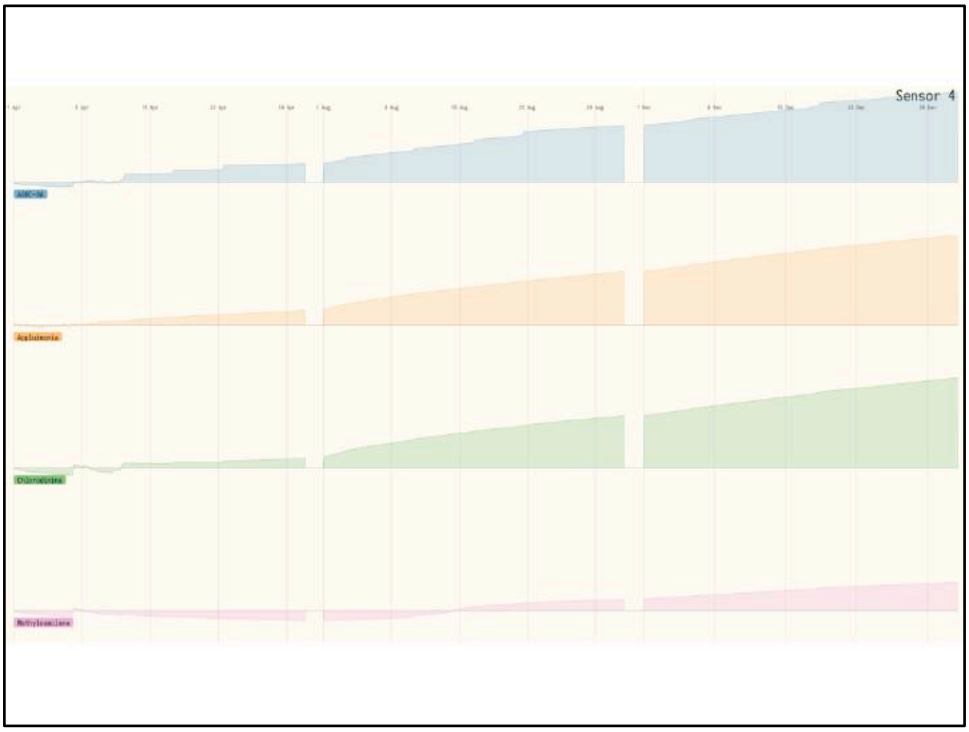
Sqrt of sensor data reduces visual impact of extreme events making visual detection of errors and trends easier.



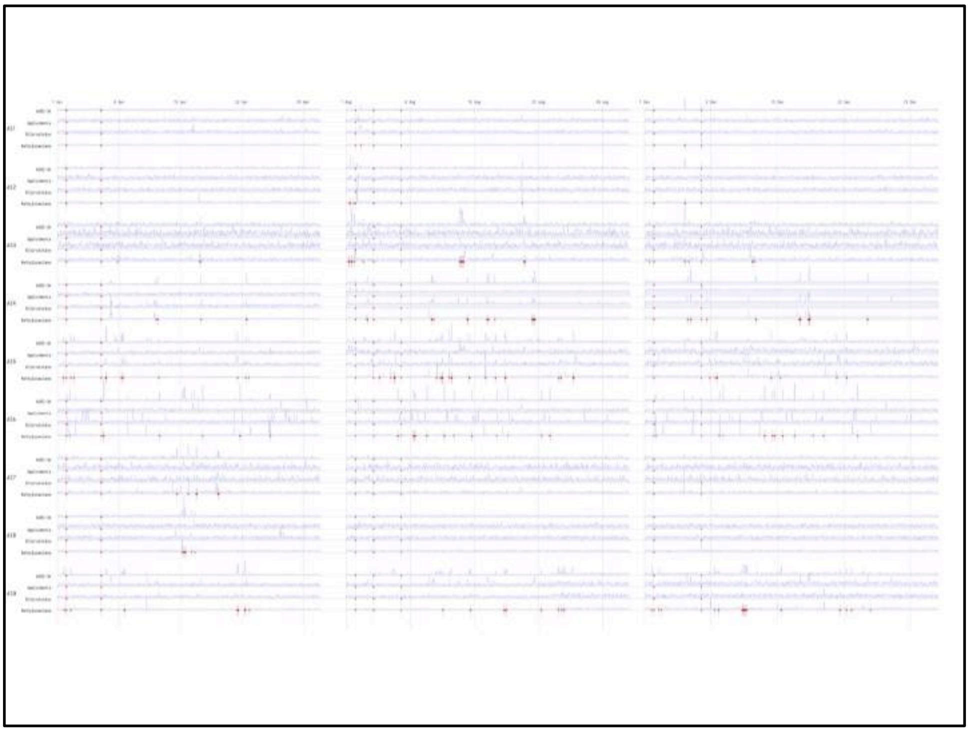
Cumulative Sum plots (CUSUM) allow early detection of deviations away from some expected cumulative average. This both smooths noise and shows systematic trends away from normal behaviour and so is suited to early detection of problems. Should be well suited to temporal visual analytics.

Used a similar approach in last year's VAST Challenge (http://gicentre.net/vastchallenge2016) for detecting anomalous trends in building sensors.

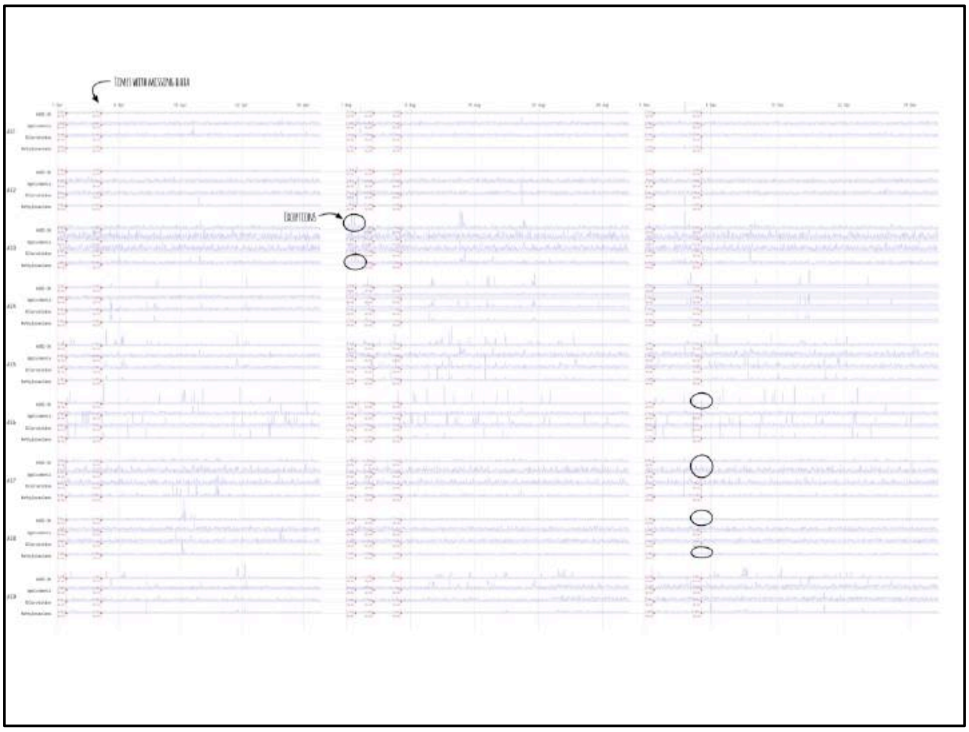
For this year’s challenge, CUSUM charts were particularly useful in detecting a gradual increases in average concentration over time and systematic errors in Sensor 4 which appeared to show a linear increase in time in parts per million readings for all 4 chemicals.



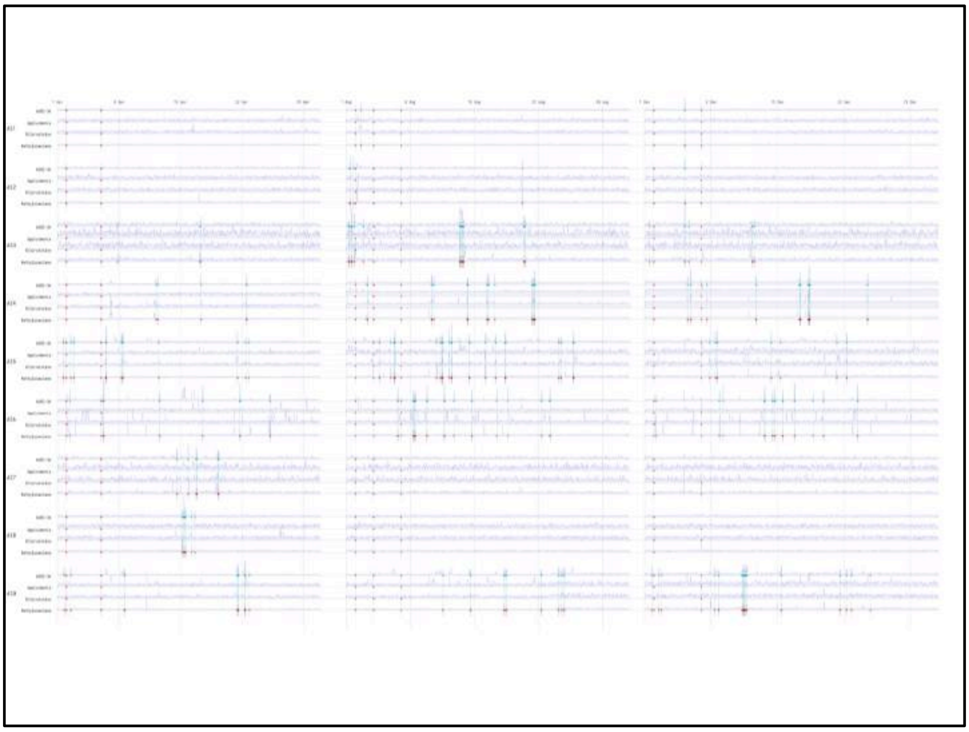
Systematic errors in Sensor 4 which appeared to show a linear increase in time in parts per million readings for all 4 chemicals.



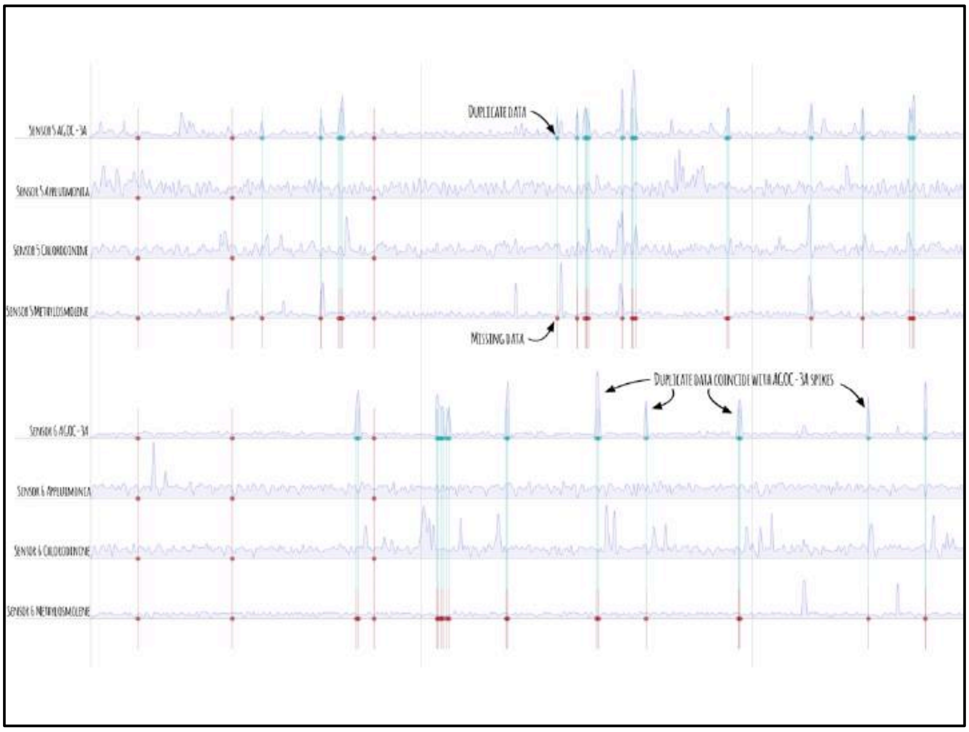
Layout of sensor readings over a common timeline allowed systematic data errors to be detected. Automatic symbolisation of missing values and duplicate values symbolised to show they occur often at specific (common) dates/times.



Layout of sensor readings over a common timeline allowed systematic data errors to be detected. Automatic symbolisation of missing values and duplicate values symbolised to show they occur often at specific (common) dates/times.

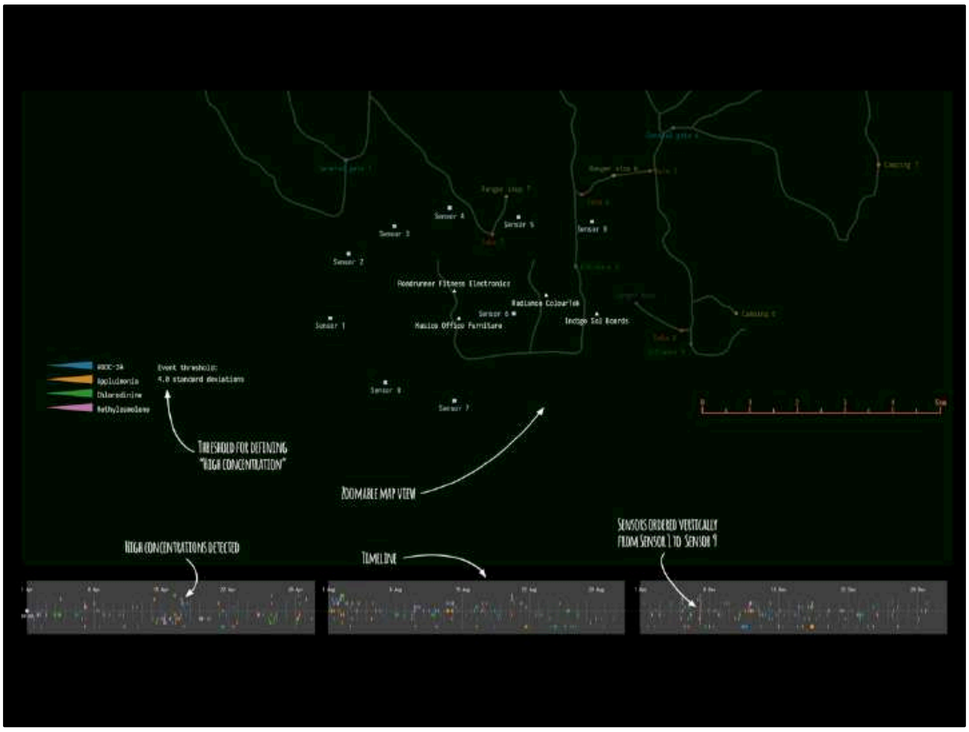


Duplicate readings (2 readings for a sensor of same chemical at same time) shown as green discs. These all coincide with missing data for the same sensor but for different chemical.

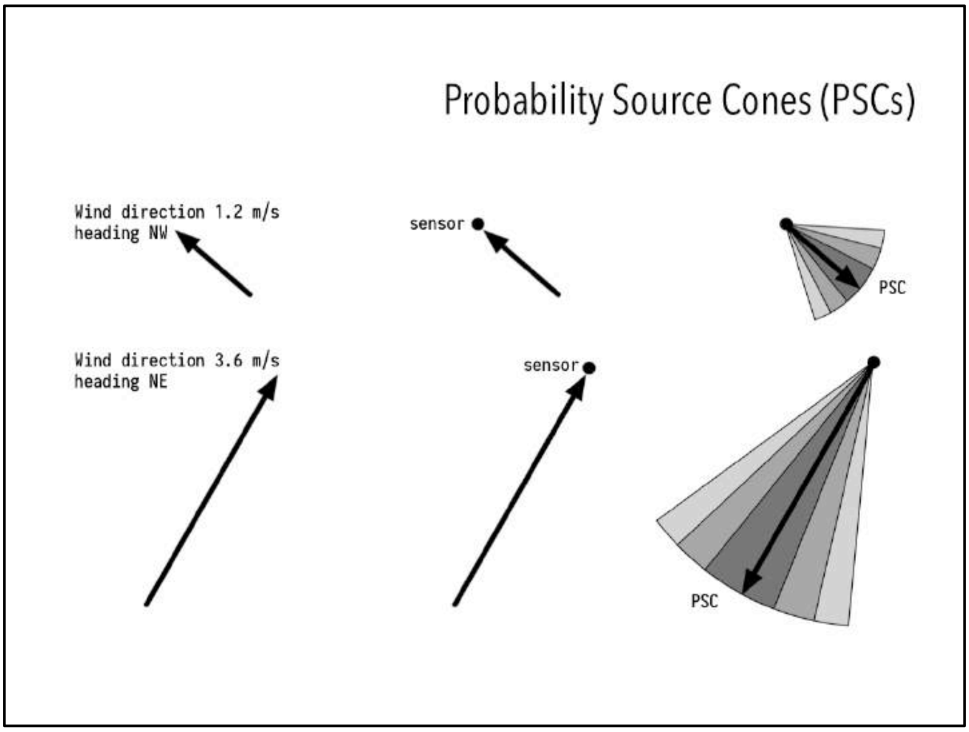


(Zoomed view). This allowed the correct chemical type to be assigned to the sensor reading for these cases (the value chosen to be transferred was the one that was more typical of the reading distributions of the two candidate chemical types)

Spatio-temporal analysis of pollutants

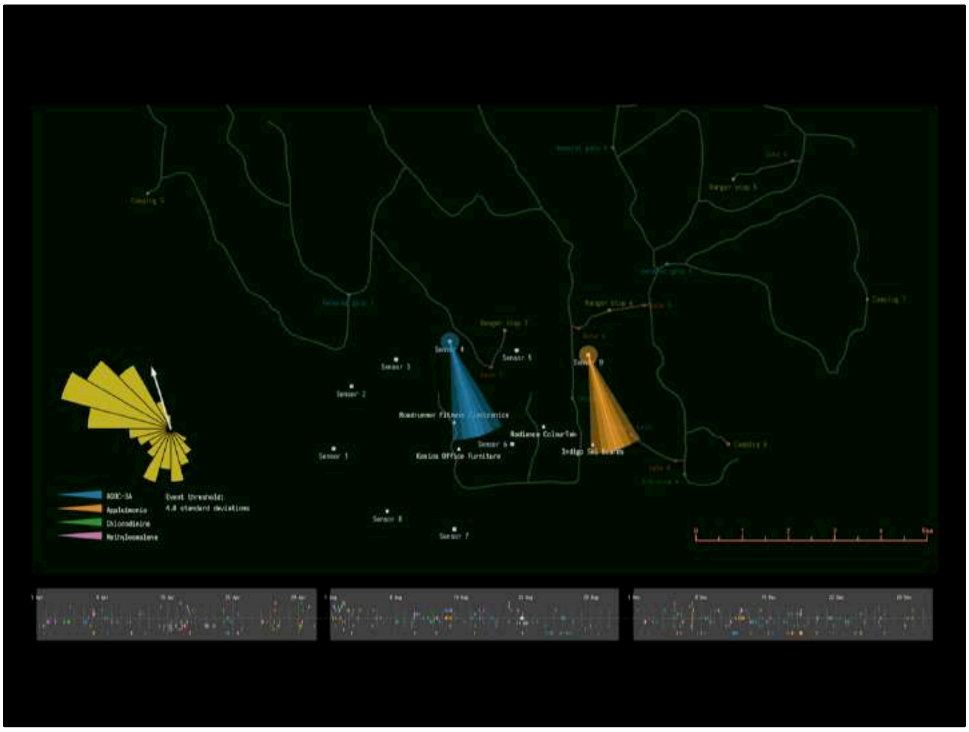


Map view layout shows location of sensors, possible sources and a timeline showing when pollution events occur.

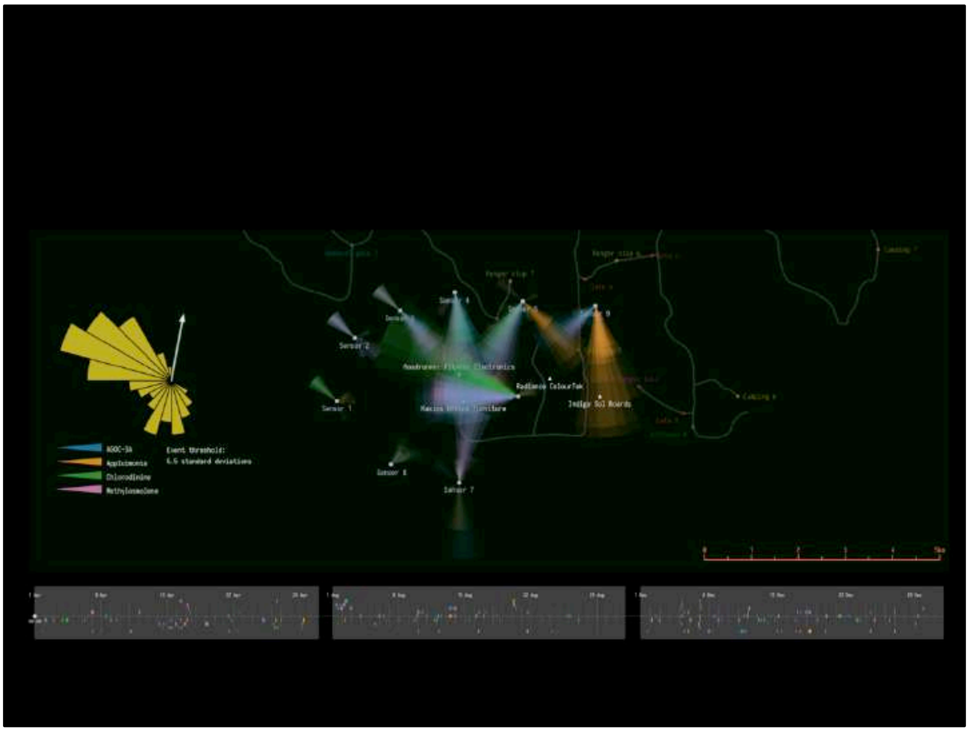


PSCs (probability source cones). Length of sector is proportional to wind speed. Orientated 180 degrees to wind direction. Transparency increases with angular width.

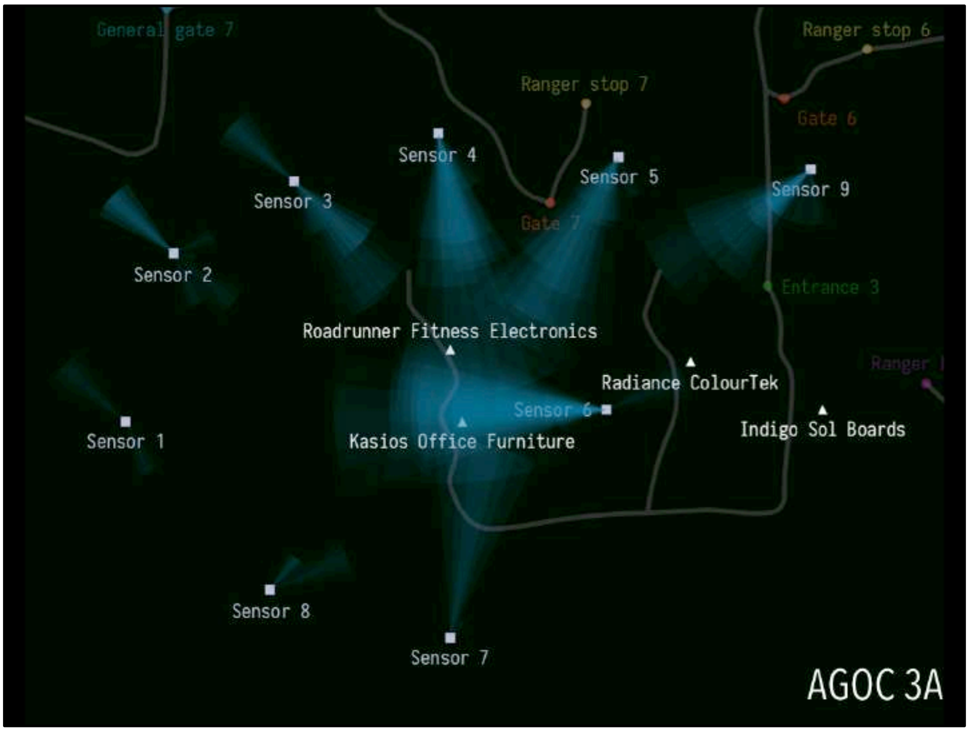
Provides a fuzzy glyph for uncertain vector representation with a known fixed endpoint (sensor).



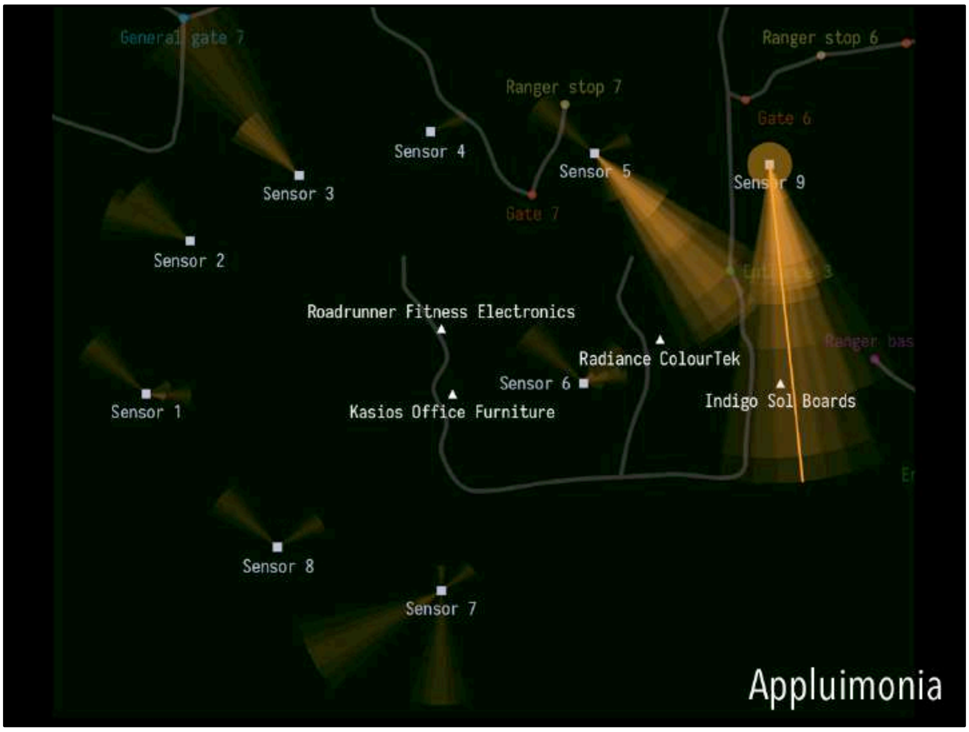
2 PSCs triggered by two simultaneous detection events, one for AGOC (blue) the other for Appluimonia (orange)



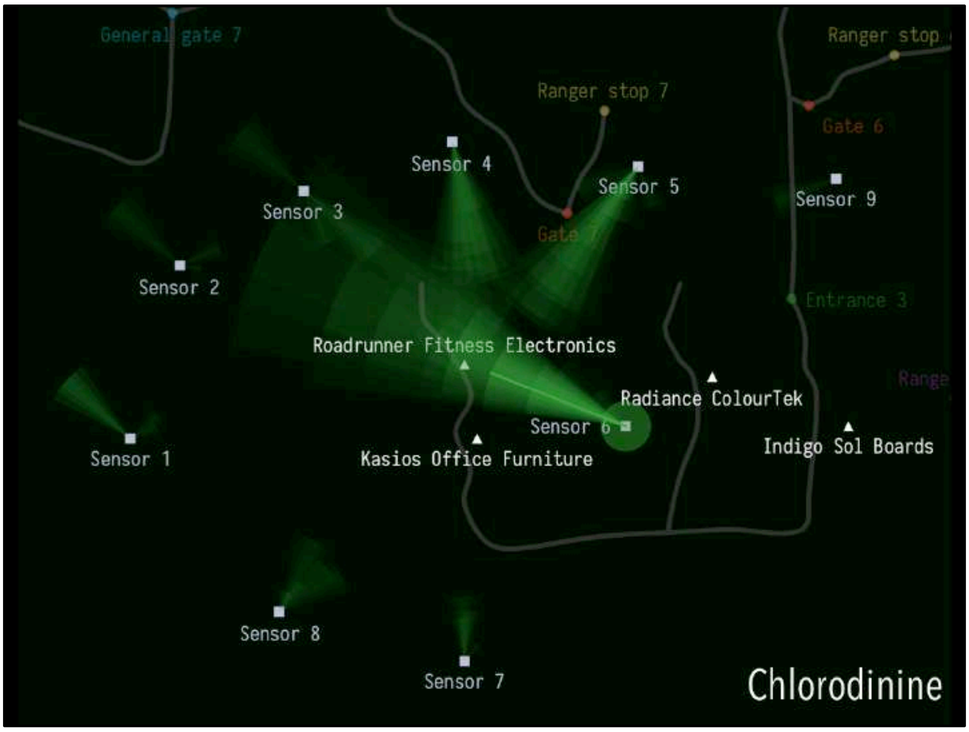
Overview of all PSCs. Threshold for magnitude of reading representing an “event” and therefore triggering a PSC can be changed interactively.



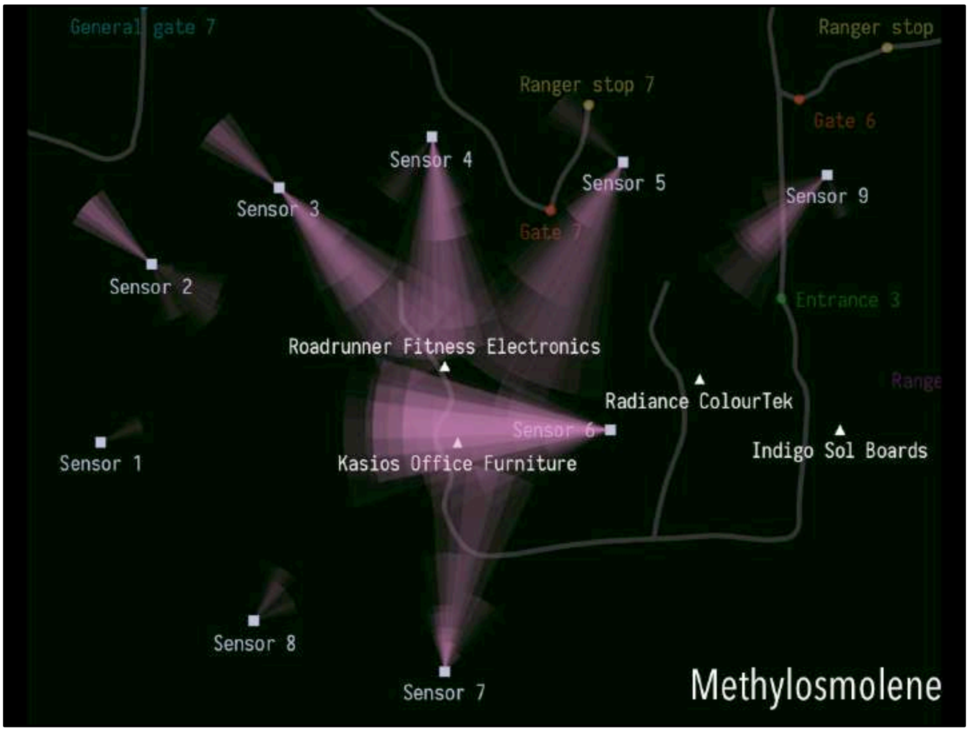
Detection of AGOC 3A with fuzzy triangulation suggests Kasios Office Furniture, not Roadrunner Fitness Electronics the source (especially with evidence from Sensor 6)



Detection of Appluimonia points strongly to Indigo Sol Boards as the source



Triangulation of Chlorodinine PSCs points strongly Roadrunner Fitness Electronics as the source



Triangulation of Methylosmolene PSCs points to Kasios Office Furniture and rules out Roadrunner Fitness.

Conclusions

Transferrable observations…

Existing GIS technologies remain insufficiently flexible for interactive and uncertain visual analytics; requires coding.

Adapting glyph symbolisation to represent uncertainty opens up new visual analytical techniques (fuzzy triangulation).

PSCs could be adapted to more sophisticated spatio-temporal modelling (e.g. plume dispersal models) or any objects with uncertain spatial footprints.

gicentre.net/vastchallenge2017

jwoLondon
johliem

Visual Analytic Design for Detecting Airborne Pollution Sources