Assessing soil data from sites in London to identify links to key indicators for characterising soils

MSc Thesis Summary by Aisha Gloudon: September 2009

Supervisors: Dr Nick Voulvoulis & Dr Martin Head

Collaborator: Ms Dee Flight (British Geological Survey)

This project assessed soil data from sites in London generated through the OPAL soil and earthworm survey in comparison with an existing BGS soil observation dataset and identified key indicators which could enable the broad characterisation of soils using public participation surveys. The key objectives of this project can be summarized as:

1. Assess the robustness of the OPAL survey dataset generated in comparison with the BGS dataset (using directly and indirectly comparable observations from both surveys)
   a. Determine and evaluate the spatial distribution of OPAL and BGS datasets
   b. Determine a methodology for comparative analysis of the OPAL results relative to the BGS dataset
2. Derive key indicators, through the collation of OPAL survey observations, to enable the broad characterisation of soils using multiple observations
3. Test the effectiveness of the key indicators by assessing the correlation of the grouped results with an existing pollution dataset (London Mercury concentration)
4. Assess the ways in which the results from this project can optimize the design and structure of future OPAL environmental surveys for data quality and the effectiveness of public participation in soil surveying.

Soil is an important medium, providing a habitat for flora and fauna, a gas exchange surface between the atmosphere and the soil surface, a sink and source for water and many pollutants and often a building material or recreational resource for people. Similarly, Earthworms are sensitive to their surrounding habitat and act as bioindicators of the various soil characteristics such as heavy metal concentrations, moisture levels and amount of soil compaction. For these reasons the monitoring of soil character is valuable and justifies the pursuit of environmental surveys such as the OPAL Soil and Earthworm survey.

The OPAL Soil and Earthworm survey involves the public in recording soil properties and Earthworm abundance/speciation with 3100 surveys to date completed by schools, community groups and individuals. As well as the numerous benefits of increasing public involvement with their local environment, these observations provides a large body of information on environmental conditions and biodiversity, including from private sites (e.g. residences) which may not normally be accessible in the course of a survey by a single organisation.

The importance of evaluating soil quality and the potential effectiveness of environmental surveys via public participation provides the rationale for this
research project. There is little previous quantitative work on the accuracy of public participation data in environmental surveying, particularly with regard to soil surveys. Consequently, this project was important in forming the basis for a quantitative comparison of the accuracy/reliability of the OPAL soil survey through the development of key indicators and comparison to the quality controlled BGS survey data. The empirical data generated by the IC OPAL survey was analysed and represented spatially. The BGS and OPAL soil observational data was then collated and standardised in order to be comparable to each other. The IC OPAL data was compared against the BGS data and the results generated by the IC OPAL survey were evaluated in comparison to the well constrained BGS dataset.

Finally, the IC OPAL and BGS observations were grouped under three key indicators of soil character, namely 1. Ability to Support Life, 2. Geography, Geology & Climate, 3. Anthropogenic Impact/Pollution

The BGS dataset consisted of soil characteristic observational data and data on mercury concentrations at locations in London, gathered in 2008. The BGS dataset is rigorously quality controlled from collection to laboratory analysis and so provided a good comparator for the OPAL dataset, gathered in 2009 by members of the MSc OPAL research group (referred to as the IC OPAL dataset).

**Generation of datasets**

The relevant London soil observational data was extracted from the UK-wide BGS dataset and collated together with the Mercury concentration data also provided by the BGS. The IC OPAL soil observational data was obtained through fieldwork by members of the MSc OPAL research group using the OPAL survey methodology.

The two areas of London selected for sampling and analysis are shown in the two adjacent maps. Green dots represent exact matches between OPAL and BGS land use categories (see below), and red dots represent non-matches. All the dots were sampled during this research project.

**Basic processing of datasets**

The observational data from both the BGS and IC OPAL surveys was selected, collated and analysed using descriptive statistics and some aspects of the data such as mercury concentrations were represented spatially using ArcGIS.
Reclassification of BGS and IC OPAL datasets

Although the BGS and IC OPAL datasets surveyed many of the same soil characteristics, the way in which these observations were recorded and/or categorised varied slightly between the two surveys. Therefore, in order to compare the BGS and IC OPAL datasets the recording categories needed to be standardised between the two and reclassified whilst maintaining as much of the original detail as possible in order to minimise decay of the data quality.

Comparison of results

Direct Comparisons

The categories of land use, soil texture, soil colour and the presence of objects within the soil, were present in both the BGS and IC OPAL surveys and so the results of these four categories were compared directly between the two surveys. Overall, the direct comparisons between the IC OPAL and BGS survey data for land use, soil objects, soil texture and soil colour to be reasonably well correlated. Using these parameters the IC OPAL soil observation dataset is shown to be very robust when directly compared against BGS soil observations.

The results of the direct comparisons within each category were tested for spatial bias using statistical analysis tools within ArcGIS. This analysis showed that none of the results of the matches exhibited spatial bias and therefore suggested that the results were not influenced by the identity of the sampler and that the sampling methodology was generally robust throughout the survey.

In addition, for exact matches, several differences were identified between the classification methods of the BGS and the IC OPAL surveys which, although few, proved to be responsible for a significant proportion of the variation between BGS and IC OPAL surveys. For example, in the case of Land Usage, the
classifications “Parkland” and “Industrial Site” were significantly underused in the IC OPAL survey compared to in the BGS survey. Closer analysis identified that “Parkland” was identified accurately when it was used but that it was not reliably used. This suggested that there was a variation between the surveyors’ understanding of “Parkland”. However, this variation may also be an artefact of the reclassification process in order to compare the BGS and IC OPAL datasets. This trend was repeated for many all the direct comparison categories where the proportion of exact matches were low. For this reason, the use of the “moderate” or “broad” match parameters is well justified for evaluating the validity of IC OPAL data.

Indirect Comparisons

...were carried out on those categories across both surveys which although recording different observations could be related to each other to draw conclusions on various soil characteristics. Many categories had the potential for indirect comparison between the two sites. For the purpose of this research the relationship between BGS recorded organic soil content and IC OPAL soil smell and number of plant roots was selected to test the robustness of the indirect comparison methodology. The indirect comparison of BGS organic content and IC OPAL abundance of plant roots and smell produced what appeared to be excellent relationships. They showed that the relationship expected as a result of literature study was revealed and a higher number of plant roots and an earthy smell were associated with high organic soil content. The abundance of plant roots echoed this relationship most closely. However, there was no clear correlation between higher concentrations of mercury and higher organic content as reflected by the IC OPAL plant root data.

Mercury Comparison

The IC OPAL categories pH, soil texture and organic content (represented by number of plant roots) were identified through a review of literature as the most significant of the observed factors influencing mercury concentrations. These categories were therefore compared against BGS mercury concentrations to test these relationships.

Comparisons were carried out on a site by site basis and then the total number of matches between surveys for each category was generated. The relationships between matches and non-matches were then explored further using graphical methods and analysis of relationships as identified from relevant literature.

Conclusions

This research was successful in assessing the functionality of the OPAL survey and demonstrated that in some aspects the results generated from a public participation survey of this nature can be of a similar standard to those generated by a high-quality comparator survey (e.g. BGS soil surveys).
Overall the accuracy and reliability of soil characterisation from various sites in London using the OPAL survey was shown to be of a high quality when tested using direct comparators. This reliability decreased with the use of indirect comparators, showing that further work needs to be done to combine IC OPAL categories into indicators so that indirect comparisons will be possible. The results of the IC OPAL survey were generally well correlated with Mercury levels but this was also less successful using indirect comparators.

This research provides the methodology foundation for further work within this field including the grouping of OPAL observations to form soil character indicators and the reclassification of BGS observations for comparison with OPAL data. In this way it goes some way towards enabling the raw data submitted by the public to be quality controlled and better analysed in the future.