The Development of Linked Databases and Environmental Modelling Systems for Decision-Making in London

Royse KR*, Bricker SH, Jackson CR, Kingdon A, and Hughes AG

British Geological Survey, Keyworth, Nottingham, United Kingdom, NG12 5GG

*Corresponding author: email: k.royse@bgs.ac.uk Tel: +44 (0)1159363456

Abstract

A basic requirement for a city's growth is the availability of land, raw material and water. For continued and sustainable development of today's cities we must be able to meet these basic requirements whilst being mindful of the environment and its relationship with anthropogenic activity. The heterogeneous and complex nature of urban systems where there are obvious environmental and anthropogenic inter-dependencies necessitates a more holistic approach to decision-making. New developments such as linked databases of environmental data and integrated environmental modelling systems provide new ways of organising cross-disciplinary information and a means to apply this to explain, explore and predict the urban systems response to environmental change. In this paper we show how, accessibility to linked databases, detailed understanding of the geology and integrated environmental modelling solutions has the potential to provide decision-makers and policy developers with the science based information needed to understand and address these challenges.

Keywords

Linked data, Environmental modelling, Decision support, London and the Thames Basin, Parametisation

1. Introduction

London is already one of the most densely populated cities in Europe; only Copenhagen, Brussels and Paris currently have higher densities (UNEP 2006). The London Plan (The Mayor of London, 2008) predicts that the total population of the city (currently 8.2 million) will rise by 900,000 by the year 2016, an increase greater than the current population of Leeds (UK) in less than a decade. Furthermore, projections for the years beyond 2016 indicate that London's population will continue to grow, intensifying the requirements for natural resources, especially drinking water, and space for living, working and recreation.

The demands for space in the urban environment have resulted in some cities, like New York, spiralling upwards by building massive skyscrapers, while others such as Montreal have exploited the sub-surface (Galipeau and Besner 2003). London depends on the sub-surface for a significant range of services, hosting utilities infrastructure such as energy, telecommunications, drinking water, storage (e.g. cellars and parking), sewerage, and transportation. The challenge, therefore, is not just in understanding what the impact of increased sub-surface infrastructure will have on the environment but how the environment will impact on the effectiveness of these systems. Two examples have been chosen to illustrate how these challenges can be overcome: Use of a linked database system to make data readily assessable to the users and a model composition to understand the impact of environmental change on the Sub-Surface hydrology of the Thames Basin.

In London, the development of linked database systems, which provide access to high quality coded borehole, physical and chemical property data, permits the development of the 3D geological understanding of the subsurface and better characterisation of potential unforeseen ground conditions (e.g. rockhead anomalies such as drift filled hollows and collapsed Pingo structures, faulting and sand channels within the Lambeth Group) in order to reduce project costs, and reduce risk (financial and/or Health and Safety). The enhanced geological understanding provides the foundation for numerical modelling, where for example in the London area the groundwater model being developed for the Environment Agency is being enhanced by including a better representation of the Chalk structure and hydraulic properties. The Chalk groundwater model of the London urban area and greater Thames Valley forms part of a linked groundwater modelling system, which accounts for the groundwater resource situation across the Thames catchment. This modelling approach brings together groundwater and surface water abstraction, in order that the water situation for London can be determined and it's resilience under climate change scenarios better assessed.

Regionally, London is located within the western portion of the Anglo-Brabant Massif, north of the Variscan Front and east of the Mesozoic North Sea rift system (Erratt *et al.*, 1999; Ellison *et al.* 2004). The London Basin (Figure 1) is traditionally considered to have formed in the Oligocene to mid-Miocene times during the main Alpine compressional event (Ellison *et al.* 2004). Formations in London (Figure 1) range from Cretaceous (144 to 65 Ma) to Quaternary (2 Ma to present day) in age. The Cretaceous Chalk is typically a fine grained white limestone. It has a total thickness of between 175 and 200 m and generally thins from west to east. Overlying the Chalk is the Palaeogene that comprises the Thanet Sand Formation, the Lambeth Group (consisting of the Upnor Formation overlain by a complex mix of various facies, attributed to the Reading and Woolwich formations) and the Thames Group, which consists of the Harwich and London Clay Formations. Quaternary deposits are encountered throughout the London Basin. These include evidence of ancient river systems and the development of the present-day River Thames valley. London has a history of *"anomalous ground conditions"*; for example, gravels have been encountered where clay might reasonably be expected, sudden changes have been found in the elevation and inclination of the Lambeth Group, bedding parallel shears are found in almost horizontally dipping London Clay, cones of depression for water levels in the Chalk can have a rectangular shape seen in plan and models for ground water flow in the Chalk repeatedly suggest that the notion of the Chalk as an hydraulic continuum is probably inappropriate; the list could go on. As many of these surprises have been found in isolation and in association with a particular investigation, and as the resources that revealed them usually only permit their presence to be noted, no further work follows; they remain as one more *"anomaly"* to add to the growing list of such anomalies within the Basin. These anomalous ground conditions can prove costly to new development projects, if not picked up in the initial site surveys, and may lead to project over-runs. However, by improving the accessibility to geological, and property databases and the development of integrated environmental modelling solutions will allow for a more complete understanding of the sub-surface environment.

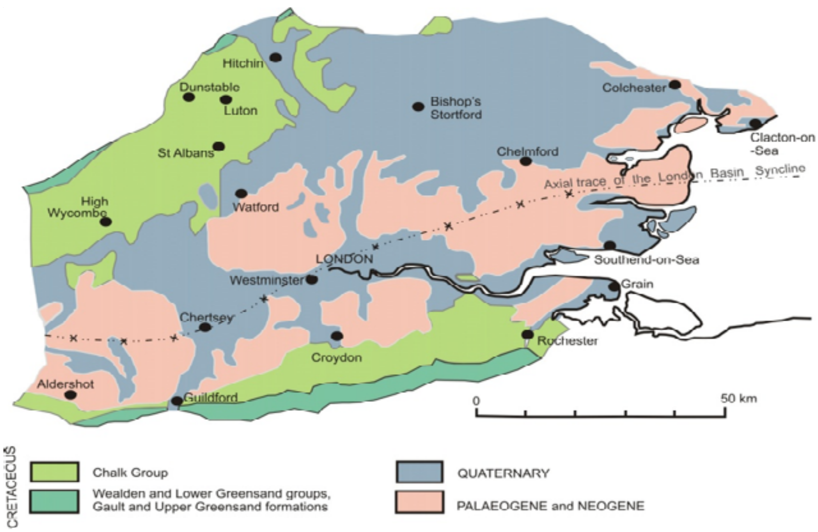

Figure 1: Location map of the Thames basin and London

2. Linked Databases

Unforeseen ground conditions are one of the most significant sources of delay in major building projects in London, and as a consequence increasing build costs. The costs of ground investigations represent a significant component of the total development costs; reduction in these risks has the potential to save on the costs of large construction and infrastructure project and thus stimulate the UK economy. There is a huge amount of geological property information for London held by the British Geological Survey (BGS) which has been collected over the years by the BGS, donated by research scientists, and by industry and government bodies under statutory requirements. These datasets are used within BGS to deliver an enhanced understanding of subsurface heterogeneity in London, to deliver physical property datasets and 3D voxelated geological models to potential customers/end-users for this information. A 3D model of the bedrock and superficial geology of London and parts of the wider Thames valley has been developed using GSi3D software at approximately 1:50,000 scale. The model covers 4800Km^2 and has been developed using approximately 7000 borehole records through which 922 geological cross-sections have been constructed. 3D geological volumes exist for over 20 units from Cretaceous through to the Paleogene. The superficial geology model is comprised of over 70 Quaternary (including anthropogenic) units. Refinements to the chalk 3D geological model to incorporate major faults and the chalk subdivisions across the Thames Basin are being made to enhance numerical modelling of the groundwater systems.

New databases to serve physical properties data to industry would bring significant benefits to the UK economy. Engineering data from large-scale public funded major projects such as HS2 (High Speed Two http://www.hs2.org.uk/) and CrossRail(http://www.crossrail.co.uk/) needs to be made easily available to the communities that these projects affect, as well as to the geotechnical consultants, developers, and public sector development organisations. BGS has for many years stored all digital scientific analysis and records in relational databases (currently over 200) to ensure the long-term continuity of this information. However the structure of these databases is, by necessity, complex; each database, as well as containing positional reference data and model information, also contains metadata such as sample identification information and attributes that define the source and sample processing. This metadata is critical to detailed assessment of the value of these analyses but it can over complicate a dataset when only a simple understanding of the variability of a single physical property is required. Another problem is the issue that property information is held within several databases, for example, in order to understand the variability of porosity within the Lambeth Group in London, there will be a need to interrogate a number of databases, as it is likely that porosity will have been measured for a variety of uses in a variety of ways. This means that the extraction of physical properties from these databases for a first look at e.g. porosity in the Lambeth Group is not a straightforward task; therefore the 'PropBase Query Layer' has been created to allow a simplified aggregation of and extraction of all related data held within the BGS databases.

The concept of the Query Layer is the presentation of complex data in simple tables. The PropBase Query Layer brings together property information from various databases (each with its own database structure that reflects the nature of the data) into a single system. This means that data from all of the BGS's subsurface data holdings can be viewed together within a simple interface. The query layer structure comprises of a set of tables, procedures, functions, triggers, and views. The structure contains a main table PRB_DATA which contains all of the data with the following attribution:

- A unique identifier for each record
- Source of the data
- Corresponding unique identifier of the record from its parent database for traceability
- Geographic co-ordinates
- Depth values
- Type of property
- Value of the property

- Units of measure
- Appropriate qualifiers
- Precision values and a full audit trail for the record

The data source, property type and units of measure are constrained by a series of dictionaries collated from the values used in the different databases from which data is extracted to populate the query layer. The property dictionary is a key component of the structure as this defines what properties and inherent hierarchies are to be coded and also guides the process as to what and how these are extracted from the structure. The structure incorporates some flexibility allowing extra tables to be added as required linked and for these to be linked off the main PRB_DATA table to capture extra attribution with a 1-to-many relationship or even a 1-to-1 relationship.

Given the size of the query layer's structure and the many property types and their values from various data sources, it's important that there is a co-ordinated technical approach to keep the layer synchronised. The query layer therefore makes use of oracle procedures written in PL/SQL containing the logic to carry out the data manipulation (inserts, updates, deletes) to keep the layer synchronised with the underlying databases. These procedures and/or packages are run as scheduled jobs at regular intervals (weekly, monthly etc.) or can be invoked on demand.

3. Application of Environmental modelling systems to London

As pressures increase on water resources from greater demand and environmental change, there is a need to understand, model and manage catchments and societal demand and forecasts on a holistic basis. The Thames basin, offers a large number of complex inter-relations in which population growth, increases in per capita consumption of resources, changes in land use in combination with the impacts of climate change interact to provide a highly challenging and complex system in which to manage water resources. In order to understand the basin hydrology and the impacts of pressures on the groundwater systems within the basin an integrated approach is necessary. The groundwater regime in the Thames Basin is made up of about 20 distinct aquifer units from Jurassic age limestones in the west of the catchment through to Palaeogene sands and silts in the east and including the major Chalk aquifer system (Figure 2). These are then overlain by variable thicknesses of Quaternary superficial deposits and alluvium. The majority of these aquifer units are not in hydraulic connection with each other and the only way water is exchanged between these aquifers is via the River Thames and its tributaries. In addition, the Thames catchment is extensively exploited for public water supply. The Thames Basin exhibits highly variable spatial rainfall patterns and hence recharge varies across the basin and requires significant management and intervention by water companies and regulators. Understanding (or capturing) the variability in recharge, in combination with the heterogeneity in both geology and hydrogeology across the basin is key to the management of water resources at the catchment scale. This is further complicated by the variability in demand both spatially and temporally, from the public, the environment, agriculture, and industry. It is also important to have an understanding of the vertical and lateral variability of aquifer properties in the basin at a variety of scales, from regional down to borehole scale. Much of the hydrogeological heterogeneity can be gleaned from aquifer property data through interrogation of PropBase Query layer.

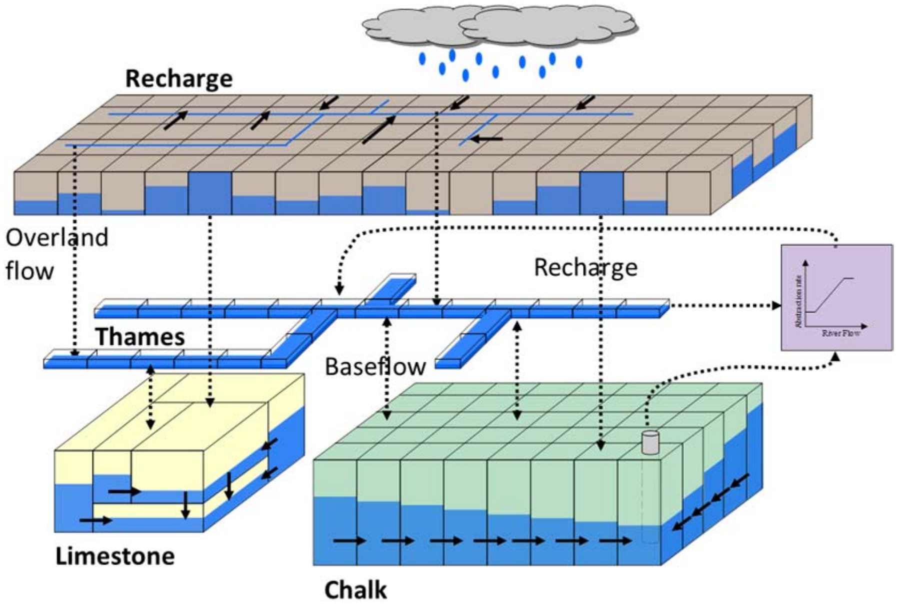

Figure 2: Conceptual diagram illustrating how the multiple aquifer units in the Thames Basin are connected together via the River Thames and are incorporated into different models.

To develop better environmental models of the Thames Basin a methodology has been developed to simulate multi-aquifer systems by linking independent groundwater models, a river model and an abstraction management module (Mansour et al 2013) using the linking modelling standard OpenMI (OpenMI Association, 2010; Gregersen et al 2005). The groundwater models have been developed at different times, in different model codes and with different levels of complexity which reflects both the level of understanding of the different aquifers and how they function. It has been possible to use OpenMI to link structurally quite different models together to improve the simulation of flow processes on a catchment scale. Operational water resource management has also been introduced to the environmental modelling composition with the inclusion of an abstraction management module. The project has demonstrated the flexibility of using OpenMI as a model linkage tool. The integrated model also provides a more realistic representation of the main aquifer processes and interactions in the Thames Basin and hence will provide a valuable tool for the integrated management of groundwater resources.

4. Conclusions

Sustainable development can only be achieved if an appreciation of the close links between the environment and society are understood (Hopwood *et al.* 2005). These links are many and varied. In this paper we have shown that with new developments, such as linked databases for environmental data and the integration of environmental modelling systems, we can start to take a whole system approach to sustainable urban development.

In London, difficult ground conditions cost developer’s time and money through project overruns and expensive engineering solutions. A more sustainable solution is in improving our understanding of the relationship between the geology and its properties. To achieve this, we need to improve how we access and share data between the commercial sector and the wider geological community (Barron et al 2013). The implementation of the PropBase QueryLayer within BGS, to find, display and interpret a greater number of datasets with greater ease, massively simplifies the process of populating 3D framework models volumes with physical properties for parameterisation and study of geological intra-unit heterogeneity. This has enabled more rapid data discovery and population of 3D models with data held in our databases, enabling different datasets to be easily compared improving the data verification process.

The need to develop better more conceptually realistic environmental models can be achieved through the use of new model linkage software such as OpenMI. These allow individual models to interact (in real time if necessary) to provide a more realistic model of the processes taking place within the basin. We have shown how legacy models, which represent a significant investment in both time and money, can be re-used and linked together. Finally, even though we can deliver information faster and more efficiently producing more realistic environmental models of the urban and peri-urban environment, communication and collaboration between stakeholders within the urban environment is equally important, (if not more so). It is only when true engagement between all partners is gained that sustainable developments can be delivered.

5. References

Barron, H., Bonsor, H.C., Bricker, S., Entwistle, D. C., Lawrence, D., Lawrie, K., Ó Dochartaigh, B. É. Ó., Campbell, S., and Watson, S. 2013. Maximising past investment in subsurface data in urban areas for sustainable resource management: a pilot in Glasgow, UK. *Ground Engineering*, Technical note February 2013

Ellison, R.A., Woods, M.A., Allen, D.J., Forster, A., Pharaoh, T.C., King, C., (2004) Geology of London. Special Memoir for 1:50 000 Geological sheets 256 (North London), 257 (Romford), 270 (Dartford) (England and Wales) British Geological Survey.

Erratt, D., Thomas, G., Wall, G., (1999) The evolution of the Central North Sea Rift. In: Fleet, A., Boldy, S. (Eds.), Petroleum Geology of Northwest Europe: Proceedings of the 5th Conference. Geological Society of London, pp. 63–82

Galipeau, G and Besner, J. (2003): The underground city of Montreal: a win-win approach in the development of a city. Presentation at the 1st international conference 'sustainable development and management of the subsurface. Utrecht, the Netherlands.

Gregersen, J.B., Gijsbers, P.J.A., Westen, S.J.P, and Blind, M., (2005). OpenMI: The Essential Concepts and their Implications for Legacy Software. Advances in Geosciences, 4, 37- 44, 2005.

Hopwood, B, Mellor, M and O'Brien, G. (2005) Sustainable Development: Mapping Different Approaches. Sustainable Development 13: 38-52

Mansour, M, Mackay, J, Abesser, C, Williams, A, Wang, L, Bricker, S, Jackson, C. 2013 Integrated Environmental modelling applied at the basin scale: Linking different types of models using the OpenMI standard to improve simulation of groundwater processes in the Thames Basin, UK. Modflow and More 2013: translating science into Practice, Colorado, USA, 2-5 June 2013

OpenMI Association, 2010. OpenMI Standard 2 Specification for the OpenMI (Version 2.0). Part of the Open MI Document Series. www.http://www.openmi.org

The Mayor of London. (2008) The London Plan: Spatial development strategy for Greater London. Consolidated with alterations since 2004. Greater London authority. 391pp

UNEP. (2006) Global Environment Outlook Year Book 2006: an overview of out changing environment. 91pp