Combining numerical and cognitive 3D modelling approaches in order to determine the structure of the Chalk in the London Basin

Katherine R. Royse, British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

k.royse@bgs.ac.uk

Abstract

In order to determine the structure of the Chalk in the London Basin, a combined cognitive and numerical approach to model construction was developed. A major difficultly in elucidating the structure of the Chalk in the London Basin is that the Chalk is largely unexposed. The project had to rely on subsurface data such as boreholes and site investigation reports. Although a high density of data was available problems with the distribution of data and its quality meant that, an approach based on a numerical interpolation between data points could not be used in this case. Therefore a methodology was developed that enabled the modeller to pick out areas of possible faulting and to achieve a geologically reasonable solution even in areas where the data was sparse or uncertain.

By using this combined approach, the resultant 3D model for the London Basin was more consistent with current geological observations and understanding. In essence, the methodology proposed here decreased the disparity between the digital geological model and current geological knowledge. Furthermore, the analysis and interpretation of this model resulted in an improved understanding of how the London Basin evolved during the Cretaceous period.

Keywords: London, faulting, Cretaceous, Chalk, Numerical and knowledge driven modelling

1. Introduction

Since 3D geological modelling became an economic and technical reality in the late 1980s (Rosenbaum 2003), there has been a remarkable growth in computer modelling applications able to proffer 3D modelling solutions (Gibbs 1993, Perrin et al. 2005, Sobisch 2000, Turner 2006). It is now possible not only to view and manipulate 3D models on a standard desk top computer but also to integrate disparate digital datasets (De Donatis et al. 2009). This has enabled 3D geological models to move from the sole used of the petroleum and mining industry to becoming a standard geological tool used by all (Kessler et al. 2009, Rosenbaum , Turner 2003, Royse et al. 2009, Xue et al. 2004).

One of the key developments within the UK has been the increased availability of digital geological data. The first major step was achieved through the digitisation of the Geological map (Jackson , Green 2003). In subsequent years, data, for example borehole logs, tunnel maps and site investigation reports, became increasingly available in digital formats (Bowie 2005, Jackson 2004). This necessitated changes in data management practice (Culshaw 2005, Turner 2006), such as the requirement for data to be spatially registered in nationally recognised coordinate and elevation systems and a move towards corporate databases which have nationally agreed data standards and validation procedures (Baker , Giles 2000, Kessler et al. 2009). This increased accessibility of digital data has resulted in 3D models moving from the conceptual model of (Fookes 1997) towards the ‘real’ geological model of (Culshaw 2005, Royse et al. 2008). In order to fully complete this process, improvements will be needed in the current algorithms and concepts used in current computer modelling packages (Wycisk et al. 2009).

Geological 3D modelling software currently works in one of two ways, either using numerical algorithms to interpolate between data points such as borehole data (Krige 1966, Mallett 1992) or by using a more cognitive interpretative approach, which allows for the incorporation of expert geological knowledge between observational data points (Hinze et al. 1999, Sobisch 2000). In this paper a numerical 3D modelling method is defined as one where numerical algorithms are used to interpolate between data points (Wycisk et al. 2009) and a Cognitive 3D modelling methodology is one where the modeller incorporates his own geological knowledge to connect between data points (Kessler et al 2009). Both systems have their advantages; however, for many ‘real life’ situations, the best answer is one where a combination of both approaches should be used. This was the case with the London Chalk Model (LCM) which comprises of a series of seven faulted layers, representing six Chalk formations and the overlying undivided Palaeogene strata (Royse 2008). Producing as realistic a geological model as possible becomes more significant when the model is to be used to generate further numerical datasets, for example, a groundwater model (Wycisk et al. 2009). The work presented in this paper was funded by the Environment Agency, Thames Region, to support work on the production of a new hydrogeological model for the River Thames catchment.

2. Geographical and geological context

The model encompasses an area within the catchment of the River Thames; it extends from Hornchurch Marshes in the east to Hounslow in the west, up to Enfield in the north and down to Croydon in the south (Fig. 1). Geologically, the London Basin is a broad, gentle synclinal fold, whose axis can be traced from Chertsey through to Southend-on-Sea (Fig. 1). The basement rocks (Palaeozoic strata) of the region belong to 2 distinct structural provinces. To the north is the London Platform which is part of the Midlands Microcraton and in the south is the Variscan Fold Belt (Ellison et al 2004, Fig 3.). The geological structure of the Cretaceous and Palaeogene strata has in the past been considered to be ‘relatively simple’ (Ellison et al 2004) for example, on the current geological maps for the region only two faults are shown, the Wimbledon and Stratham fault and the Greenwich fault (Fig 2). There is however a growing body of data, particularly from recent deeper engineering projects such as the Channel Tunnel Rail link (CTRL, (Harris et al. 1996, Newman 2009), CROSSRAIL and the Docklands light railway, suggesting that the structure of London is far more complex.

The London Basin is thought to have formed in the Oligocene to mid-Miocene times during the main Alpine compressional event (Ellison et al. 2004). Formations in this region range from Cretaceous (144 to 65 Ma) to Quaternary (2 Ma to present day) in age. The Cretaceous Chalk is present at subcrop throughout the London basin and comes to the surface along the southern margin (the North Downs) and along the northwest margin (Chiltern Hills) and is locally at or close to the surface e.g. along the Greenwich and Purfleet anticlines in East London.

The Cretaceous Chalk is typically a fine grained white limestone. (Bristow et al. 1997) provides a detailed description of the Chalk lithostratigraphy (Fig 4). The Chalk in the London area can be divided into 6 Formations; West Melbury Marly Chalk, Zig Zag Chalk, Holywell Chalk, New Pit Chalk, Lewes Nodular Chalk, and the Seaford and Newhaven Chalk undivided (Fig 3). These are distinguished by changes in their, hardness, colour and lithology and by the presence or absence of marl and flint bands. In the London area the total thickness of the Chalk is between 170 and 210 m and generally thins from the west to the east. The London Basin succession is a relatively thin succession compared to that of the Hampshire – Dieppe Basin where the Chalk is over 400 m thick (Ellison et al 2004). Overlying the Chalk is the oldest Palaeogene deposit, the Thanet Sand Formation. This formation consists of a coarsening upwards succession of fine grained, grey sand. The formation reaches a maximum thickness of around 30 m in the area. Above the Thanet Sand Formation lies the Lambeth Group. This group consists of three formations: the Upnor, the

Woolwich and the Reading Formations. The Lambeth Group is between 20 and 30 m thick in the area and lithologically, the group is highly variable, consisting of variable proportions of sands, silts, clays and gravels. Overlying the Lambeth Group are the Eocene sediments of the Thames Group which consist of the Harwich and London Clay Formations. The Harwich Formation (formally known as the Blackheath or Oldhaven Beds) consists predominantly of sand and pebble beds up to 4 m thick. Above this is approximately 90 to 130 m of London Clay. The London Clay Formation consists of grey to blue grey, bioturbated, silty clay. 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. Deposits include alluvium, peat, brickearth and river terrace deposits (for example the Kempton Park, Taplow and Shepperton Gravels).

3. Data sources and acquisition

This section describes the data collected for the LCM. The LCM project area is entirely within the city of London and as a consequence there is a huge quantity and variety (both in age and type) of geological data which can be incorporated into the model. This data has been collected by the British Geological Survey over a period stretching from the 1830s to the present day. Therefore the quality as well as the quantity of data available to define the position of each geological surface in the model is highly variable. In general, uncertainty in the thickness and geometry of any modelled geological unit is greatest in areas where the data is sparse and or of poor quality. Conversely, confidence is highest where there is a high concentration of good quality data (Kaufmann , Martin 2008b). Therefore the first stage in the modelling process was to collect, sort, interpret and validate this data (Kaufmann , Martin 2008b). The data used in this project, described below, can be divided into two main types: interpretative (geological maps, cross sections, research reports and memoirs) and observational (boreholes, site investigation reports, and outcrop descriptions)

3.1 INTERPRETATIVE DATA

Four digital 1:50 000 scale geological maps published by the BGS cover the LCM project area [sheets 256 (North London), 257 (Romford), 270 (South London) and 271 (Dartford)]. These maps were all re-surveyed during 1970–1995. The London Memoir (Ellison et al. 2004) covers all four map sheets within the study area and has been used as the definitive text in this study (additional information sources are listed below). The map sheets 256, 257 and 270 all use the traditional three-fold subdivision of the Chalk. However, map sheet 271 uses the new lithostratigraphic scheme developed for the Chalk over the last eleven years (Bristow et al. 1997, Rawson et al. 2001). For a full list of interpretive information sources used in this project, see table 1.

3.2 OBSERVATIONAL

In this study, 12,400 lithostratigraphic and 200 geophysical (natural gamma and resistivity) borehole records were looked at; these records are held in the National Geological Records Centre and by the Environment Agency. The records are of variable age and quality and many lacked useful lithological (or lithostratigraphical) information, the descriptions being too vague, imprecise or inaccurate. In the end, some 4,300 borehole logs were found to provide useful information about at least one stratigraphic boundary.

Where possible, the level of each stratigraphic boundary recorded in these logs was determined and stored centrally in an oracle database called Borehole Geology (Kessler et al 2009). The database contains information on each borehole’s unique identification code, its national grid reference, its height relative to UK Ordnance Datum and information on the depth to base of each stratigraphic boundary encountered in the borehole along with a free text description of that boundary. The digital borehole data was then downloaded form a data portal (Kessler et al 2009, (Howard et al. 2009) into a tab separated table which was compatible with the data formats required for GSI3D and GoCad.   As errors can occur in any portion of the borehole data for example, in the original record, in its subsequent interpretation and in the recorded location of the borehole, (Aldiss et al. 2004) these were checked for in each individual borehole. The National Grid coordinates for boreholes were taken from the BGS Single Onshore Borehole Index (SOBI). The ground surface level (relative to Ordnance Datum) for each borehole was taken from the borehole record, where documented. Recorded levels were checked against the NEXTMAP DTM. Where ground levels were not recorded, or were obviously incorrect, the level was interpolated from the NEXTMAP DTM elevation data.

The lithological boreholes were interpreted using the new Chalk lithostratigraphy (Bristow et al. 1997). Borehole logs intersecting the top of the Chalk beneath the Palaeogene were extrapolated downwards to the base of each of the new Chalk formations, using an estimated thickness for each (Aldiss et al 2004). It should be noted that the thickness of each unit is known to vary slightly across the area, and so these ‘phantom data points’ are correspondingly uncertain. The ‘phantom data points’ were incorporated into the production of the digital geological cross-sections, which were drawn up as part of GSI3D modelling procedure (see section 4.1). The cross-sections provided a means of checking each phantom point’s position relative to other boreholes in the near vicinity. In this way the ‘phantom data points’ made a valuable contribution to elucidating the position of each Chalk formation within the model.

Geophysical logs (natural gamma and resistivity) stratigraphic interpretation was based on work by Mortimore and Pomerol (1987b) and Murray (1986) and is described more fully by Woods (2001, 2002). Geophysical boreholes were scrutinised in a similar way to those of the lithological boreholes; each record was first interpreted individually, and then each interpretation was compared with that of its nearest neighbours, as a further check on the consistency of the interpretation.

Interpreted borehole data was then used to generate the 3D model, enabling the borehole records to be considered relative to each other, in their local context. Borehole records which gave rise to obvious anomalies in the modelled surfaces and which seemed to be in some way unreliable (e.g. over-simplified drillers' logs) were noted within the modelling metadata files and then discarded. It should be noted that borehole records which are somehow incorrect but which are nevertheless consistent with the model will generally remain unsuspected (Aldiss et al 2004).

4. Geological modelling

Modelling was carried out to ascertain not only the distribution of the six Chalk formations found within the London Basin but also the Chalk's structure. One of the major difficulties in elucidating the structure of the Chalk within the London Basin is that the Chalk is largely unexposed and where it is exposed, it is either covered by superficial deposits (drift) or obscured from view due to urban development. Therefore the project had to rely to a large extent on the Geologist's interpretation of the subsurface data and geological observations made in the mid to late 1800s. Although few faults are indicated on the current published geological maps, there is a growing body of data, particularly from recent deeper engineering projects such as the Channel Tunnel Rail link (CTRL), (Harris et al. 1996, Mortimore et al. *In prep*), that suggests that faults are far more numerous. These data are further supported by the mounting evidence that tectonic and sea-level movement occurred in phases throughout the upper Cretaceous (Evans , Hopson 2000, Evans et al. 2003, Mortimore , Pomerol 1987a, 1991, Mortimore et al. 1998).

A methodology was needed that enabled the Geologist to apply his geological knowledge intuitively into the 3D model, as would be the case when producing a traditional geological map Therefore a workflow was needed to mirror as much as possible the methods used when drafting traditional cross-sections across areas with sparsely distributed control data (Fig. 5). This allowed the modeller to pick out areas of possible faulting and to achieve a geologically reasonable solution even in areas where the data was sparse or uncertain (kaufmann , Martin 2008a, b, Lemon , Jones 2003). Therefore a methodology was developed that combined a cognitive and numerical approach using the combined functionality of GSI3D (version 2.5) and GoCad (version 2.1.3). This approach allowed the modeller to capture his/her own interpretation of the geometry and thickness of each geological unit (Kessler et al. 2009), to pick out areas of faulting and generalise the faults into a coherent fault network, and finally, using numerical techniques in GoCad, to smooth and cut the model by the fault network generated.

4.1 Cognitive modelling methodology

GSI3D modelling methodology (Sobisch 2000) allows the modeller to model the distribution and geometry of geological units by using the modeller's geological knowledge (Wycisk et al. 2009). The modelling procedure within GSI3D is based on the creation, by the user, of a series of intersecting cross-sections. The Cross-sections are generated from borehole information and 2D geological map and surface data. A generalised vertical section (GVS) is then defined for all the rock units in the study area. The package then interpolates between nodes along the sections and produces a series of triangulated irregular networks (TINs), for each rock unit modelled (Kessler et al 2009). Because GSI3D uses a 'constructive method' (Wycisk et al 2009) the package provides the modeller with the ability to connect areas in the model, where there is either only partial data coverage or where the geometry of the geological units is poorly understood. The LCM was constructed by correlating outcrop data with boreholes that were linked together in a network of intersecting cross-sections. Data was included from a considerable distance beyond the project area in order to ensure that regional trends were correctly represented (Fig 6a)

The cross-sections were constructed in roughly orthogonal directions (north-south and west-east), which allowed for borehole correlations to be checked iteratively across the area (Fig 6c). Where possible, cross-sections were placed at right angles to known geological structures. Shorter, ancillary cross-sections on other alignments were constructed, in order to encompass local variations and anomalies. Errors caused by data deficiencies were checked against the supporting data and removed or smoothed. A total of 100 sections were constructed (Fig 6c).

During model construction, metadata was recorded describing the geologist's decision-making (cognitive) processes and any boreholes found to be erroneous. This is an essential part of the procedure. Firstly, it is important that the model is repeatable; therefore the modeller needs to record what assumptions or actions were made as part of the cognitive modelling method. Secondly, it allows the eventual model to be reused at a later date when the originator may not be reachable, thereby future-proofing the data. Once the model was assembled in GSI3D, the sections were revisited to check that fault determinations were valid.

4.1.1 **Determination of faulting**

As mentioned in section 2 only two faults have been mapped in the London Basin yet a growing body of evidence from recent site investigations suggests that in reality the structure of the Basin is more complex (Newman 2009, Skipper et al. 2008). However determining the exact nature of faulting within the London basin is difficult because the majority of the bedrock is either at subcrop and/or covered by the built environment of the city of London or by thick superficial deposits related to the development of the River Thames. To further add to the problem, elucidating faulting within the Chalk outcrop of Southern England is known to be problematic (Aldiss et al. 2004). This is due to the fact that when faulting is observed in the Chalk, the displacement has often been accommodated by movements on numerous small-scale faults within a zone tens or even hundreds of metres wide. For example, known (mapped) faults in the London Basin such as the Greenwich fault (Ellison et al. 2004), occur as a single plane on the geological map, but is in reality a zone of disruption which includes a number of closely spaced faults and fractures. Therefore, in unexposed Chalk terrain, it is rarely possible to distinguish the difference between a broad, gentle anticlinal fold and a broad fault zone (Aldiss et al. 2004). Therefore to elucidate the structure of the London Basin an approach was needed that would allow a geologist trained in traditional field surveying techniques and specialising in the geology of the London Basin the ability to capture his specialist knowledge and understanding in a 3D geological model. It was found that by using the GSI3D cognitive approach (see section 4.1) with its methodology based on the long-standing relationship between the geological map and cross-section generation (Kessler et al 2009), a structural model for the London Basin could be achieved. During this process a set of criteria, that suggested areas where faulting in the Chalk Strata was probable, was documented, see Table 2.

At this stage 90 individual fault traces were picked out. As discussed above, known faults in the London Basin are in reality zones of disruption which consist of a number of closely spaced en echelon faults. Therefore the individual fault traces were viewed in a more regional context and compared with the gravity anomaly and interpreted datasets in ArcGIS (Fig 3, Table 1). This was then used to produce a regional fault pattern for the London Basin. The resulting fault network consisted of 13 major fault zones cutting across the project area (Fig 6 d). It should be noted that the relatively sparse distribution of subsurface data did not allow for the delineation of any but the most obvious structures, particularly where the occurrence of small to medium scale faults in the Chalk is less than the general spacing of the boreholes.

4.2 Numerical modelling

Once these steps were completed, the data was exported into GoCad. GoCad operates on the premise that the geometry of any geological object can be defined by a set of points. An object is modelled by the links connecting these points. The Discrete Smooth Interpolation algorithm (DSI), which sits in the interior of the GoCad programme, was designed to model the geometry of complex geological objects and account for any constraints, such as boreholes data, placed upon it (Mallet 1997).

The data imported consisted of digital cross-sections generated in GSI3D, the original borehole data, which were all imported into GoCad as 3D geo-registered point data, the NEXTMAP DTM was brought in as a surface and the generalised fault network work (Fig 3) and digital geological line work was imported in as 3D line datasets. Data exchange between the two programmes (GSI3D and GoCad) was simply made through existing file exchanges. This data provided the constraints to the final modelled surface produced in GoCad.

Using scripts ‘wizards’ within GoCad, triangulated surfaces were generated for each geological formation and fault plane. The surfaces were constructed using the DSI algorithm to compute the location of the nodes (Mallett 1997). This algorithm produces a geometry which is smooth, but can also takes account of a set of constraints, in this case the borehole and cross-section data (Galera et al. 2003). Once this is done, a series of steps are followed which removes cross-over errors between the surfaces. This is done through either applying thickness constraints or moving surfaces above or below a reference surface i.e. the surface with the highest quantity of good quality well distributed data. Once these stages were completed the resultant model could be visualised and assessed (Fig 7).

4.3 Comparison of the proposed 2 step methodology with a single step numerical modelling method

After the modelling work was carried out, a comparison was undertaken between the combined cognitive and numerical workflow with a more numerical workflow using script ‘wizards’ within GoCad to interpolate between borehole points. In Figure 8 part of the base Palaeogene surface has been remodelled using a numerical workflow. The same borehole dataset was used as in the combined approach discussed in sections 4.1 and 4.2. The base Palaeogene surface was specifically chosen for this comparison because it has the highest number of borehole data points defining its surface. The location was picked as it is an area where faulting is not recorded on the current geological maps but where observations from deeper engineering works would suggest that faulting may be present.

The comparison of the two surfaces in Figure 8 shows clearly the effects of the combined approach on surface construction and fault determination on the base Palaeogene surface. For example the northern boundary fault, NW and ENE trending faults described in section 5 (Fig 9) are clearly observed in the combined method however the more numerical workflow does not provide a clear indication of all of these structures. In this case even though a large number of boreholes are available for the base Palaeogene surface, where the geology was faulted the numerical workflow was not able to achieve a model that was as consistent with current geological knowledge and observations as the combined methodology attained (see section 6; Newman 2009). Subsequent layers beneath the base Palaeogene surface have significantly less borehole data defining their surfaces, for example, the Seaford Chalk Formation contains only 54% of the total number of boreholes used in the project. With depreciating amounts of borehole data intersecting each succeeding lower layer the results achieved with a single stepped numerical workflow become increasingly inadequate. In essence the single stepped numerical modelling methodology requires a high concentration of boreholes which are evenly distributed for each surface to be modelled.

5 The Structure of the Chalk under London as derived from the London Chalk Model

By using a combined cognitive and numerical method, the resultant 3D model for the London Basin was consistent with current geological observations and understanding. The analysis and interpretation of this model, discussed below, has resulted in an improved understanding of how the London Basin evolved during the Cretaceous period.

The geological structure of the London Basin was generally thought to be a relatively simple north-east trending syncline (Ellison et al. 2004). However, the LCM suggests that, in detail, the London Basin is a much more complex structure, being a collection of at least 5 fault-bounded basins (Fig 9 and 10). The model also suggests that the project area can be split into two sections or regions, which have behaved differently during the evolution of the basin. This split can be related to the two structural provinces observed within the basement strata in the region (Ellison et al. 2004): the northern portion being underlain by the London Platform (part of the Midlands Microcraton) and the southern portion by a zone of transition between the London Platform and the Variscan fold-thrust belt (Fig 3). This change in basement material across the Basin has determined, to a large extent, the type and intensity of the geological features found in each region.

For example, folding within the project area (Fig. 11) can be divided into two groups: the first group found south of the London Basin Axis (Fig1) and coincidently South of the River Thames consists of east-north-east trending periclinal folds, including the Greenwich and Streatham anticlines. These features are generally high amplitude and short wavelength folds, many of which are asymmetric, usually with steeper north-facing limbs. The second group are confined to the northern part of the project area and are in the main low amplitude, long wavelength folds.

Faulting is predominantly confined to the south-eastern portion of the project area; its distribution within the London Basin again appears to have been controlled by the properties of the basement which underlie it. The faults, broadly speaking, can be divided into 3 groups (Fig 9): ENE trending faults, which downthrow to the north (the majority of faulting within the south-eastern sector); ENE trending faults, which downthrow to the south (northern boundary faults); and northwest trending faults, which downthrow to the west. Displacements range between 10 to 50 m. The LCM modelled Chalk surfaces also suggest the presence of a central structural high. The central structural high is bound to the west by the NW trending faults and to the north by an ENE trending fault.

6. Summary and Conclusions

This paper has described a combined cognitive and numerical modelling methodology.

In order for this approach to work, two key developments were necessary; the availability of digital geological data within the UK and the inter-operability between modelling packages, which provided the tools necessary to integrate different types of digital geoscientific data and modelling approaches. This methodology was developed in order to overcome the problem of having an uneven distribution of borehole/subsurface data which was clustered around linear routes e.g. infrastructure developments and a limited amount of surface exposure of the Chalk in central London, (either because the stratum was at sub-crop or because it was covered by superficial deposits and/or the built environment). It was found, that to produce the most realistic 3D model possible, large quantities of data was not enough; it was also essential to use the correct processing method. The method had to produce surfaces (faults and stratigraphic horizons) that not only honoured the data but were also geologically reasonable and finally, the resultant model had to be repeatable, in other words the hypotheses or concepts used to generate the model had to be captured.

The project therefore had to incorporate specialist geological knowledge from a geologist more at home with traditional field surveying techniques than ‘state of the art’ computer modelling packages. Consequently it was essential that a methodology was developed that enabled the

Geologist to not only capture his knowledge and understanding of the geology of Chalk in London but to also provide a means of selecting areas of possible faulting and finally to achieve a geologically reasonable solution even in areas where the data was sparse or uncertain.

Therefore the accuracy of any 3D digital model will depend not only on the data, its density and quality, but also on the theoretical understanding of the underlying geology by the modeller. It follows therefore that, when assessing the confidence or uncertainty of a model, a key component should be the modeller's theoretical knowledge and experience (Royse et al. 2009). This becomes more critical when the model is to be used to generate further numerical datasets as is the case in the London Chalk Model. All users of 3D models must be able to understand the limitations of the data on which they base their assessments. Improvements in 3D modelling methods are allowing geoscientists to introduce a far greater level of realism into their 3D models. It is therefore essential, particularly where cognitive modelling techniques have been used, that users are able to understand how the model was produced as well as the density and quality of the data used. One way to achieve this is to compile metadata files during the modelling process. These files should contain information on exactly what modelling processes were undertaken, the modellers understanding of the geological setting, what data was discarded and why these actions were taken. As Users, ultimately, need to be able to assess the risk associated with using 3D models, so that sound decisions can be made (Royse et al 2009).

The methodology combined together the combined functionality of GSI3D and GoCad. This approached allowed the modeller to capture an interpretation of the geometry and thickness of each geological unit (Kessler et al. 2009), to pick out areas of faulting and generalise the faults into a coherent fault pattern, and finally, using numerical techniques in GoCad, to smooth and cut the model by the generated fault pattern. In essence it provided a conduit through which the capture of specialist geological knowledge could be achieved and used within a 3D modelling environment. It was essential that metadata was kept with the modelling project, so that a record of the concepts and processes performed on the model were recorded. This would mean that the modelling procedures could, at a later date, be reproduced.

The resultant model is more consistent with current geological observations and theories and as a consequence the model is a closer representation of geological reality. For example the model predicts that the Greenwich fault continues into north east London and that there is faulting to the south of the River Lea (Fig 6d). Ground investigations, including rotary cored boreholes, carried out as part of the Thames Tideway tunnelling project (Newman 2009) has shown that these predictions can be substantiated. Further evidence for validation of the modelling methodology has come from chalk-cored boreholes from the Thames Waters Lee Tunnel and Thames Waters Ring Main extension, where site investigations recently reported by Mortimore et al (In prep) suggest the presence of a major north-south offset which has again been predicted by this model. Current work underway on production of a new hydrogeological model for London has found that in using the new fault model the resulting groundwater level pattern fits better with groundwater level observations (Steve Buss pers. comm.)

In conclusion, the increasing accessibility of digital data along with a combined cognitive and numerical approach to model development will result in 3D models moving from the conceptual model of Fookes (1997) towards the ‘real’ geological model of Culshaw (2005). To fully complete this process, modelling software that combines both cognitive and numerical approaches is required. If this can be achieved, then the future proposed by Culshaw (2005), where ground investigations and the development of groundwater models will start by testing the validity of the ‘real’ geological model, will become a reality.

Acknowledgement

The author would like to thank her many colleagues at BGS for their help and support with the production of this paper. The two referees are thanked for there constructive comments on the original submission, which have helped her, make significant improvements to the paper. This article is published with the permission of the Executive Director of the British Geological Survey (NERC).

References

Aldiss, D. T., Bloomfield, J. R., Buckley, D. K., Doran, S. K., Evans, P. M., Hopson, P. M., Royse, K. R., Woods, M. A. 2004. A geological model of the chalk of East Kent. Volume 1 , 2. British Geological Survey Commissioned Report. CR/04/92, 97pp.

Baker, G. R., Giles, J. 2000. BGS Geoscience integrated database system: a repository for corporate data. Earthwise 16, 12-13.

Bowie, R. 2005. Creating a digital landslide. Geoscientist 15(3), 4-7.

Bristow, C. R., Mortimore, R. N., Wood, C. J. 1997. Lithostratigraphy for mapping the chalk of southern England. Proceedings of the Geologists' Association 108, 293-315.

Culshaw, M. G. 2005. From concept towards reality: developing the attributed 3D model of the shallow subsurface. Quarterly Journal of Engineering Geology and Hydrogeology 39, 231-284.

De Donatis, M., Borraccini, F. , Susini, S. 2009. Sheet 280 - Fosombrone 3D: A study project for a new geological map of Italy in three dimensions. Computers & Geosciences 35, 19-32

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 (South London) and 271 (Dartford) (England and Wales). British Geological Survey. 114 pp.

Evans, D. J., Hopson, P. M. 2000. The seismic expression of synsedimentary channel features within the chalk of southern England. Proceedings of the Geologists’ Association 111(3), 219-230.

Evans, D. J., Hopson, P. M., Kirby, G. A., Bristow, C. R. 2003. The development and seismic expression of synsedimentary features within the chalk of southern England. Journal of the Geological Society, London 160, 797-813.

Fookes, P. G. 1997. Geology for engineers: the geological model, prediction and performance Quarterly Journal of Engineering Geology 30, 293-424.

Galera, C., Bennis, C., Moretti, I., Mallett, J. L. 2003. Construction of coherent 3D geological blocks. Computers & Geosciences 29, 971-984.

Gibbs, B. 1993. Mineral Industry Software. In: Bird, D., Brewis, T., Chadwick, J., Clifford, D., Ellis, R., Kennedy, A., Morgan, R., Pearse, G., Rosin, N., Schooling, J., Spooner, J., West, M. (Eds.) Mining Annual Review, 1993. Mining Journal Publications, London, p. 57-63.

Harris, C. S., Hart, M. B., Varley, P. M., Warren, C. D. 1996. Engineering Geology Of The Channel Tunnel. Thomas Telford Publishing. 526 pp

Hinze, C., Sobisch, H.-G., H-H, V. 1999. Spatial modelling in geology and its practical use. Mathematische Geologie 4, 51-60.

Howard, A., Hatton, B., Reitsma, F., Lawrie, K. 2009. Developing a geoscience knowledge framework for a national geological survey organisation. Computers & Geosciences 35(4), 820-835.

Jackson, I. 2004. Britain Beneath Our Feet. British Geological Survey Occasional Publication No. 4, Hawthorns, Nottingham. 114pp.

Jackson, I., Green, C. A. 2003. The Digital geological map of Great Britain. Geoscientist 13(2), 4-7.

Kaufmann, O., Martin, T. 2008. 3D geological modelling from boreholes, cross-sections and geological maps, application over former natural gas storages in coal mines. Computers & Geosciences 34, 278-290.

Kessler, H., Mathers, S., Sobisch, H.-G. 2009. The Capture and dissemination of integrated 3D geospatial knowledge at the British Geological Survey using GSI3D software and methodology. Computers & Geosciences 35, 1311-1321.

Krige, D. G. 1966. Two-dimensional weighted moving average trend surfaces for ore valuation. Journal of South African Institution of Mining and Metallurgy (Proceedings from the Symposium on mathematical statistics and Computer Applications in Ore Valuation) 66, 13-38.

Lemon, A. M., Jones, N. L. 2003. Building solid models from boreholes and user-defined cross sections. Computers & Geosciences 29, 547-555.

Mallett, J. L. 1992. GOCAD: a computer aided design program for geological applications. In: Turner, A. K. (Ed) Three-dimensional Modelling with Geoscientific Information Systems. Kluwer, London. p. 123-141.

Mallett, J. L. 1997. Discrete modelling for natural objects. Mathematical Geology 29, 199 - 219.

Mortimore, R. N., Pomerol, B. 1987a. Correlation of the Upper Cretaceous White Chalk (Turonian to Campanian) in the Anglo-Paris Basin. Proceedings of the Geologists' Association 98, 97-143.

Mortimore, R. N., Pomerol, B. 1987b. Correlation of the Upper Cretaceous White Chalk (Turonian to Campanian) in the Anglo-Paris Basin. Proceedings of the Geologists’ Association 98, 97-143.

Mortimore, R. N., Pomerol, B. 1991. Stratigraphy and eustatic implications of trace fossil events in the Upper Cretaceous Chalk of Northern Europe. Palaios 6, 216-231.

Mortimore, R. N., Wood, C. J., Pomerol, B., Ernst, G. 1998. Dating the phases of the Subhercynian tectonic epoch: Late cretaceous tectonics and eustatics in the cretaceous basins of northern Germany compared with the Anglo-Paris Basin. Zentralblatt Für Geologie und Paläontologie 1(11/12), 1349-1401.

Murray, K. H. 1986. Correlation of electrical resistivity marker bands in the Cenomanian and Turonian Chalk from the London Basin to east Yorkshire. British Geological Survey Report 17/8. London: Her Majesty's Stationery Office. 18pp

Newman, T. 2009. The impact of adverse geological conditions on the design and construction of the thames water ring main in greater London, UK. Quarterly Journal of Engineering Geology and Hydrogeology 42(1), 5-21.

Perrin, M., Zhu, Z., Rainaud, J.-F., Schnieder, S. 2005. Knowledge-driven applications for geological modelling Journal of Petroleum Science Engineering 47, 89-104.

Rawson, P. F., Allen, P. W., Gale, A. S. 2001. The Chalk Group - a revised lithostratigraphy. Geoscientist 11(1), 21.

Rosenbaum, M. S. 2003. Characterisation of the shallow subsurface: implications for urban infrastructure and environmental assessment. In: Rosenbaum, M. S., Turner, A. K. (Eds) New paradigms in subsurface prediction. Lecture notes in earth sciences 99. Springer-Verlag, Berlin. p. 3-6.

Rosenbaum, M. S., Turner, K. 2003. New paradigms in subsurface prediction : characterization of the shallow subsurface implications for urban infrastructure and environmental assessment Springer, Berlin. 397pp

Royse, K. R. 2008. The London Chalk Model. British Geological Survey Commissioned Report, CR/08/125. 21pp.

Royse, K. R., Reeves, H. J., Gibson, A. R. 2008a. The modelling and visualisation of digital geoscientific data as an aid to land-use planning in the urban environment, an example from the Thames Gateway. In: Liverman, D. G. E., Pereira, C., Marker, B. (Eds) Communicating Environmental Geoscience. Geological Society, London, Special Publication 305, 89-106.

Royse, K. R., Rutter, H. K., Entwisle, D. C. 2009. Property attribution of 3D geological models in the thames gateway: new ways of visualising geoscientific information. Bulletin of Engineering Geology and the Environment. 68, 1-16

Skipper, J. A., Newman, T., Mortimore, R. N. 2008. The Engineering geology of the lee tunnel, east London, UK. Euroengeo 2008. International Association of Engineering Geology Madrid. p 1-3

Sobisch, H.-G. 2000. Ein difitles raeumliches Modell des Quartaers der GK25 Blatt 3508 Nordhorn auf der Basis vernetzer Profilschnitte. . Shaker Verlag, Aachen.

Sumbler, M. G. 1996. British regional geology: London and the Thames Valley. Her Majesty's Stationary Office for the British Geological Survey, London.173pp

Turner, A. K. 2006. Challenges and trends for geological modelling and visualisation. Bulletin of Engineering Geology and the Environment 65, 109 - 127.

Woods, M. A. 2001. Lithostratigraphical interpretation of borehole geophysical logs in the chalk group of the berkshire downs. British Geological Survey Internal Report, IR/01/150 10pp.

Woods, M. A. 2002. Correlation and lithostratigraphical interpretation of borehole geophysical logs in the chalk group of the north downs. British Geological Survey, Commissioned Report CR/02/247, 5pp.

Wycisk, P., Hubert, T., Gossel, W., Neumann, C. 2009. High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites. Computers & Geosciences 35, 165-182.

Xue, Y., Sun, M., Ma, A. 2004. On the reconstruction of three-dimensional complex geological objects using Delaunay triangulation. Future Generation Computer Systems 20, 89-104.

Figure Captions

Figure 1: Geological sketch map of project area. Adapted from Sumbler (1996)

Figure 2: Geological cross-section across region showing ‘relatively simple’ geological structure of region as previously proposed by Sumbler (1996). Section adapted from Sumbler (1996).

Figure 3: Colour-shaded bouguer gravity relief map showing location of two structural provinces dissecting project area (outlined in purple). OS data ©Crown Copyright. All rights reserved. BGS 100017897 / 2009

Figure 4: Detailed lithostratigraphy of Chalk in London. Adapted from Ellison et al. (2004)

Figure 5: Diagram of workflow developed to model Structure of Chalk under London

Figure 6: Data and fault distribution in study area. a) Distribution of boreholes in study area b) Distribution of fault traces as determined from cross-section analysis c) Fence diagram showing distribution of cross-sections within study area d) Regional Fault Network

Figure 7: 3D model of Chalk Group under London. OS data ©Crown Copyright. All rights reserved. BGS 100017897 / 2009

Figure 8: Structure contour plots of part of base Palaeogene to compare combined methodology proposed in this paper with a numerical modelling method based solely on interpolation between boreholes. OS data ©Crown Copyright. All rights reserved. BGS 100017897 / 2009

Figure 9: Structure contour plot of base of Palaeogene, showing major fault groups and location of structural high

Figure 10: Updated Geological cross-section across region showing more complex geological structure of London Basin as proposed in Figure 10.

Figure 11: Base of Seaford Chalk showing fold axial traces (lines: black with diamonds anticlines; magenta with crosses synclines and brown faults

Table Captions

Table 1: Interpretive information sources used in 3D modelling of Chalk in London Basin

Table 2: Set of criteria indicating a high probability of faulting within the sub-crop Chalk Strata in the London basin