



**British  
Geological Survey**  
NATURAL ENVIRONMENT RESEARCH COUNCIL

# PropBase Scoping Study

Urban Geoscience and Geological Hazards Programme  
Internal Report IR/06/088



BRITISH GEOLOGICAL SURVEY

URBAN GEOSCIENCE AND GEOLOGICAL HAZARDS PROGRAMME  
INTERNAL REPORT IR/06/088

# PropBase Scoping Study

RP Shaw

## *Main Contributors*

SDG Campbell, D Entwisle, I Gale, P Jackson, A Kingdon, DJ  
Lowe, AE Milodowski

The National Grid and other  
Ordnance Survey data are used  
with the permission of the  
Controller of Her Majesty's  
Stationery Office.  
Licence No: 100017897/2005.

## *Keywords*

Rock properties; PropBase;  
databases; 3D modelling.

## *Bibliographical reference*

RP SHAW. 2006. PropBase  
Scoping Study. *British  
Geological Survey Internal  
Report*, IR/06/088. 35pp.

Copyright in materials derived  
from the British Geological  
Survey's work is owned by the  
Natural Environment Research  
Council (NERC) and/or the  
authority that commissioned the  
work. You may not copy or adapt  
this publication without first  
obtaining permission. Contact the  
BGS Intellectual Property Rights  
Section, British Geological  
Survey, Keyworth,  
e-mail [ipr@bgs.ac.uk](mailto:ipr@bgs.ac.uk). You may  
quote extracts of a reasonable  
length without prior permission,  
provided a full acknowledgement  
is given of the source of the  
extract.

Maps and diagrams in this book  
use topography based on  
Ordnance Survey mapping.

## BRITISH GEOLOGICAL SURVEY

The full range of Survey publications is available from the BGS Sales Desks at Nottingham, Edinburgh and London; see contact details below or shop online at [www.geologyshop.com](http://www.geologyshop.com)

The London Information Office also maintains a reference collection of BGS publications including maps for consultation.

The Survey publishes an annual catalogue of its maps and other publications; this catalogue is available from any of the BGS Sales Desks.

*The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as its basic research projects. It also undertakes programmes of British technical aid in geology in developing countries as arranged by the Department for International Development and other agencies.*

*The British Geological Survey is a component body of the Natural Environment Research Council.*

*British Geological Survey offices*

### **Keyworth, Nottingham NG12 5GG**

☎ 0115-936 3241 Fax 0115-936 3488  
e-mail: [sales@bgs.ac.uk](mailto:sales@bgs.ac.uk)  
[www.bgs.ac.uk](http://www.bgs.ac.uk)  
Shop online at: [www.geologyshop.com](http://www.geologyshop.com)

### **Murchison House, West Mains Road, Edinburgh EH9 3LA**

☎ 0131-667 1000 Fax 0131-668 2683  
e-mail: [scotsales@bgs.ac.uk](mailto:scotsales@bgs.ac.uk)

### **London Information Office at the Natural History Museum (Earth Galleries), Exhibition Road, South Kensington, London SW7 2DE**

☎ 020-7589 4090 Fax 020-7584 8270  
☎ 020-7942 5344/45 email: [bgs london@bgs.ac.uk](mailto:bgs london@bgs.ac.uk)

### **Forde House, Park Five Business Centre, Harrier Way, Sowton, Exeter, Devon EX2 7HU**

☎ 01392-445271 Fax 01392-445371

### **Geological Survey of Northern Ireland, Colby House, Stranmillis Court, Belfast BT9 5BF**

☎ 028-9038 8462 Fax 028-9038 8461

### **Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB**

☎ 01491-838800 Fax 01491-692345

### **Columbus House, Greenmeadow Springs, Tongwynlais, Cardiff, CF15 7NE**

☎ 029-2052 1962 Fax 029-2052 1963

*Parent Body*

### **Natural Environment Research Council, Polaris House, North Star Avenue, Swindon, Wiltshire SN2 1EU**

☎ 01793-411500 Fax 01793-411501  
[www.nerc.ac.uk](http://www.nerc.ac.uk)

## Foreword

This report results from a scoping study undertaken between September 2005 and July 2006 to define the scope of, and assess the priorities for, the PropBase project. The PropBase project is intended to provide information on the physical, mechanical, chemical and mineralogical properties of UK rocks and soils and their interrelationships, to enable attribution of the 3D geological model and modelling of the properties themselves, and to obtain a better understanding of how these properties change as a result of geological processes. While one of the key drivers for PropBase is to allow 3D geological models to be attributed with property information, there are other geoscience activities for which the availability of systematic rock property information is important including enquires and data sales. Rock and soil property information can be used, for example: to improve the understanding of the distribution of engineering properties; and to better attribute groundwater flow and transport models with physical properties and so to be able to ascribe confidence limits to their 3D distribution. Information in PropBase will be a key resource for a number of sectors including radioactive waste disposal studies and ongoing enhancement of GeoSure. PropBase will also play an important role in promoting increased awareness of, and greatly improved accessibility to, corporate rock and soil property data for external clients, and will be especially beneficial to those involved in ground investigation in general.

PropBase will not be just another corporate database. Primarily, it will build on existing databases to enhance them where appropriate to meet its needs. Additional databases will be proposed where there are gaps in corporate data coverage. The most important aspect of PropBase will be the development of a 'portal' that allows seamless access to, and extraction from, corporate rock and soil property information databases and the provision of tools to summarise these data for use in a range of project types, including attribution of 3D volumes in geological models. Eventually it is envisaged that the 'PropBase Data Portal' will be one of several, linked, corporate portals that provide easy access to BGS information without the need for specialist IT skills.

## Acknowledgements

Key contributors are acknowledged as contributors to this report but many BGS staff, too numerous to name here, have contributed to discussions with this core team and we gratefully acknowledge their enthusiasm and support for the project, as well as thank them for their valuable contributions to this report and to the supporting reports.

# Contents

<b>Foreword .....</b>	<b>i</b>
<b>Acknowledgements .....</b>	<b>i</b>
<b>Contents.....</b>	<b>ii</b>
<b>Summary .....</b>	<b>1</b>
<b>1 Introduction .....</b>	<b>1</b>
<b>2 Properties .....</b>	<b>4</b>
<b>3 Vision for PropBase.....</b>	<b>5</b>
<b>4 Existing BGS data.....</b>	<b>6</b>
4.1 Corporate Data .....	6
4.2 'Private' data.....	6
4.3 Propbase Core Data.....	7
<b>5 External Datasets.....</b>	<b>8</b>
5.1 Key data holders .....	8
<b>6 Disciplines.....</b>	<b>10</b>
6.1 Geochemistry .....	10
6.2 Geophysics.....	10
6.3 Geotechnics.....	12
6.4 Hydrogeology .....	13
6.5 Mineralogy/Petrology .....	14
6.5.1 Petrographical Notes and Petrographical Reports .....	14
6.5.2 Engineering Properties of UK Rocks and Soils Project .....	15
6.6 The Nirex Digital Geoscience Database.....	15
6.6.1 PADAMOT Project Palaeohydrogeological database.....	15
6.6.2 Other datasets .....	16
6.6.3 Types of data produced by mineralogy and petrology .....	16
6.7 Palaeontology/Biostratigraphy.....	16
<b>7 Dealing with Legacy Data.....</b>	<b>17</b>
<b>8 Attribution of 3D rock volumes.....</b>	<b>17</b>
<b>9 Uncertainty.....</b>	<b>18</b>
<b>10 Potential Clients.....</b>	<b>18</b>

10.1 Internal Clients.....	18
10.2 Potential external clients.....	19
<b>11 Problems to Address.....</b>	<b>20</b>
11.1 General.....	20
11.2 Legacy data.....	20
11.2.1 Standards.....	21
11.2.2 Definitions.....	21
11.2.3 Internal consistency.....	21
11.2.4 Validation.....	22
11.2.5 Accessibility.....	22
11.3 Current data.....	23
<b>12 The way forward.....</b>	<b>23</b>
12.1 Immediately (Remainder of this FY).....	23
12.2 Short Term (years 1 and 2).....	24
12.3 Medium Term (years 3 and 4).....	24
12.4 Longer term (year 5 and beyond).....	24
12.5 Resources.....	25
<b>13 Key Recommendations.....</b>	<b>26</b>
13.1 Additional Evaluation.....	26
13.2 System/database development.....	26
13.3 Data needs and priorities.....	26
13.4 Project Management.....	27
<b>References.....</b>	<b>27</b>

## Summary

This scoping study outlines PropBase. This is not another corporate database but more an integrated means of accessing, synthesising and extracting data related to rock and soil properties, including anthropogenic soils that are held in existing, corporately managed, databases. It will also provide high-level metadata and contact information for relevant data that are held in BGS in non-corporate databases or analogue forms. PropBase will provide seamless access to rock and soil property data in a user friendly, intuitive, way as well as the means of providing summaries of the data. It is likely to be part of or, at least, linked to, the Internet Data Access (IDA), Geoscience Data Index (GDI), and Digital Geoscience Spatial Model (DGSM), data portal systems and will underpin, and enable extension of, existing enquiries services.

There are two main aspects to PropBase. Firstly, there are the database-related activities. For data in digital databases to be usable by PropBase they must have attributes that enable them to be both located spatially and related directly to rock types. To achieve this 'PropBase compliance', data will need to be attributed with XYZ co-ordinates and LexRock codes.

Secondly, there is the development of a 'PropBase Data Portal', which will provide easy access to 'PropBase compliant' datasets wherever they may reside. The DGSM Portal, which provides access to the results of 3D geological modelling activities and their related data, and the Groundwater Portal, developed at BGS Wallingford, are good examples of such data portals already developed by the BGS. Given the clear potential benefits to PropBase, it is proposed that the Geochemical Properties Interface (GPI) be developed and extended, initially for a limited range of rock properties, starting with porosity. The advantage of this portal is that, if required by the user, it already has the capability to provide statistical summaries of the data that it extracts.

PropBase will be a relatively modestly funded project and may, therefore, not have the resources to undertake significant enhancement of existing databases or to digitise analogue datasets. It must, however, be able to ensure that all relevant new data acquisitions are 'PropBase compliant' and to support development of the 'PropBase Data Portal'. A suggested outline programme for PropBase is provided.

## 1 Introduction

The PropBase project is intended to provide access to information on the physical, mechanical, chemical and mineralogical properties of UK rocks and soils, including anthropogenic deposits such as made ground, fill, etc, and their interrelationships, to enable attribution of the 3D geological model and modelling of the properties themselves. This will promote a better understanding of how these properties change as a result of geological processes. PropBase will become a key component in underpinning the strategic geological and geo-environmental knowledge and the essential 3D geoscience baseline for the BGS Core Strategic Programme.

The project links to all projects that generate or use information on the physical, mechanical, chemical and mineralogical properties of soils and rocks. This includes a very wide range of Science Budget (SB), Co-funded (CF) and Commercially Funded (CR), projects in most BGS Programmes. PropBase will act as both a source of



information for users and a means of managing properties information generated by projects.

It is a major part of the current BGS strategy to develop a 3D spatial geoscience model of Britain. PropBase complements this strategic aim by providing the means to attribute the model with data on physical, mechanical, chemical and mineralogical properties. As such, it relates to all of the eight Science Themes in the BGS 2005-10 Core Strategic Programme and addresses aspects of three of the Cross-cutting Themes (Climate change, Sustainable management of energy and natural resources, and International development).

Most research into, and investigation of, the shallow subsurface involves the collection and interpretation of physical, mechanical, chemical or mineralogical property data. However, until now no concerted effort has been made to bring such information together in a rational and linked way. PropBase will provide the framework for storage and recovery of these data. In addition, it will develop much-needed research into inter-property relationships and the control of geological and other processes on the properties. The project will also provide a significant enhancement to the 3D spatial modelling in LithoFrame by enabling attribution of models with property information.

It should be noted that library and world wide web-based searches indicate that there is no equivalent project to PropBase being undertaken elsewhere, other than for small, localised areas.

This scoping study seeks to:

- Identify internal and external user needs for physical, mechanical, chemical and mineralogical property information and research.
- Identify the data needed, and where there are gaps: (a) where samples/data cannot be cross-correlated between different disciplinary datasets and, (b) where the gaps are across the board in all areas (i.e. no data available).
- Determine the databasing requirements of PropBase, identify relevant existing databases, and the links needed between them, and develop and implement a strategy for corporate databases of physical, mechanical, chemical and mineralogical property information.
- Identify and prioritise the collection and management of new information on properties, involving field and laboratory activities as necessary.
- Develop systems to quantify and capture data quality information so that retrieved data are fit-for-purpose for the scale or function for which they are required.
- Develop a strategy for the attribution, with respect to properties data, of LithoFrame/DigMap/Lexicon, including generic issues (e.g. uncertainty).
- Provide attributes for 3 and 4D modelling.
- Support development of methodologies for the 3 and 4D modelling of property information
- Identify research needs with regard to: a) inter-property relationships and, b) the relationship between properties and geological processes.
- Produce a medium- and long-term strategy for PropBase covering: a) the 2005-10 BGS Programme and, b) the period to 2055 (fifty years) (or as long as is deemed necessary).
- Implement the data management and research strategies.

The project is envisaged as having a life of at least fifty years. This may seem both ambitious and excessive. While the development of a PropBase data framework and methodology for continued population will only take a few years and, similarly, methodologies to attribute rock volumes with properties will initially be developed by 2 or 3 projects, ongoing research on interrelationships, etc., will continue to be needed but should be demand-led, and a significant effort is required to update and digitise legacy data to 'PropBase standards'. If a comprehensive database of physical, mechanical, chemical or mineralogical property information is to be created, and the relationships and process controls on them investigated, then the project length indicated is both realistic and visionary.

This scoping study has focused on answering the following questions and preparing a detailed short-term strategy, and outlining medium- and long-term project strategies:

- What, within the four elements (physical, mechanical, chemical and mineralogical properties) is needed by users, which will require a testing of internal and external markets?
- What information is already available (internally and externally to BGS) within each of the four elements, and where are the gaps?
- What additional information is needed to meet user needs and how should it be collected (existing data, field work, laboratory testing, etc)?
- How will the new and existing information be stored and managed?
- What research is needed on: the relationships between properties; the effects of processes on properties (4D effect); property modelling; and the determination of uncertainty?
- Preparation of short-term (5 years) and long-term (50 years) strategies for PropBase, to include Task descriptions, staff requirements and costings.

Activities within PropBase will fall into five discrete areas. These are:

- Provision of an intuitive web-based interface to allow data to be stored, accessed, extracted, derived and summarised by users in such a way that data/summary data can readily be made available for other uses including 3D model attribution.
- Future management of data relevant to PropBase, including recommendations with respect to database structure and systematic entry of key information.
- Support research into property relationships, etc, particularly focused on attribution of 3-D volumes from point measurements, and assessment of confidence limits.
- Confirm which relevant datasets are already 'PropBase compliant' and prioritise those datasets that are not compliant for enhancement.
- Bringing existing digital and analogue data records up to 'PropBase standards'.

During this scoping study, reviews of data holdings have been undertaken and selected opinion canvassed at the main BGS sites. These are more fully reported elsewhere (Gale, 2006; Campbell, 2006; Self, 2006 and Shaw, 2006) and, where appropriate, summarised here.

## 2 Properties

The range of properties potentially encompassed by PropBase is broad and should include, amongst others, the following.

a) Physical properties, e.g.:

- porosity;
- permeability;
- density;
- magnetic properties – including magnetic susceptibility;
- natural radioactivity;
- electrical properties such as conductivity, dielectric polarization, resistivity;
- seismic velocities;
- fluorescence.

b.) Mechanical Properties, e.g.

- strength (compressive, shear);
- elasticity;
- plasticity;
- hardness;
- particle size.

c.) Chemical Properties, e.g.:

- bulk composition, such as in terms of whole rock analyses, silicates, carbonates, phosphates, sulphates, etc.;
- composition of groundwaters;
- reactivity, e.g. acids.

d.) Mineralogical Properties, e.g.:

- mineralogy in hand specimen, thin section, from X-ray diffraction, etc.;
- petrology and texture;
- hardness;
- streak and lustre;
- cleavage and fracture;
- refractive index;
- colour;
- cation exchange capacity and specific surface area;
- etc.

It should be noted that a number of these properties are relevant to more than one discipline.

Structural data should also be considered more widely, given their relative importance, especially, but not exclusively, in relation to the characteristics of rock masses, as should data related to weathering. Palaeontological information is also considered for completeness.

It is important that BGS continues to be pro-active in gathering appropriate third party information. This is particularly important for those types of data for which BGS is not fully self-reliant, and for example in geotechnical data. Data gathering is already extensively done in some disciplines, and most notably, in relation to PropBase, for physical and mechanical property data, but, clearly, the BGS cannot accept and retain all data that are offered because of the storage space and management cost implications. Equally, however, it is likely that other organizations dispose of digital (or analogue) datasets that would be of scientific value to the BGS. PropBase should be responsible for at least raising awareness among data owners about potential data donation of key PropBase information and PropBase could be used to provide a push for more co-ordinated data collection rather than the more ad-hoc local area project interest which is often the case at present.

### 3 Vision for PropBase

PropBase is not another corporate database but more a means of accessing, synthesising and extracting data relating to rock and soil properties that are held in existing, corporately managed databases and providing high-level metadata and contact information for data that are held in BGS but not (yet) in corporate databases. Nor is it proposed to join existing databases together to build a 'super database'. PropBase will provide seamless access to property data, and summaries of these data, in a user-friendly, intuitive way through a portal. It is likely to be part of or, at least, linked from, and to, the IDA, GDI and DGSM data portal systems. Such a portal will need to access 'PropBase compliant' datasets and extract and/or summarise the data according to user specification. This will need to include the ability to use geological and user specified areas as query constraints. It is envisaged that, eventually, PropBase will be part of a much broader suite of portals that provide access to the full range of BGS information. As such it will need to be designed in such a way as to allow this integration in future. In effect, the 'PropBase Data Portal' will be a query layer and it is probable that for this to function, all the digital databases that are accessible through PropBase will need to be held and managed in the corporate Oracle Relational Database Management System (RDBMS).

An important issue is that properties held in multiple datasets, for example porosity or density, might have been determined in different ways and to different standards. It is essential that all users are fully informed of the constraints this may place on use of the data, so that they are fully aware of all limitations.

Although the majority of rock and soil property data that the BGS holds relates to on-shore data, PropBase will also include coastal, near-shore and off-shore information. The spatial reference systems for the offshore databases are typically latitude and longitude, as opposed to the National Grid Reference (NGR) system used for onshore data. This means that there is an issue with co-ordinate systems that will need addressing. The DGSM project considered the issues relating to the conversion of national grid co-ordinates to latitude/longitude. However, the issue is further complicated by the metadata for the offshore databases commonly indicating 'system not known' for the system of latitude/longitude used (e.g. the commonly used ED50 or

WGS84 systems), and the differences between systems can result in substantial differences in location. Work is in progress to try to resolve this issue for existing databases, and there are conversion packages built into both ArcGIS and ORACLE. Therefore, this should be further reviewed by PropBase and recommendations for how this issue is to be managed within the BGS made, and then appropriate conversion tools built.

## 4 Existing BGS data

### 4.1 CORPORATE DATA

Metadata are available for all corporate datasets (i.e. those managed corporately under the guidelines established by the BGS-GeoIDS project) and these data are secure (they are backed up and they are appropriately maintained). There are currently over 450 separate datasets managed corporately (these are listed in Appendix 1 of Shaw, 2006), many of which contain information of direct relevance to PropBase. Further assessments of those datasets maintained in Murchison House are also contained in Campbell (2006, Appendix 1).

Viewed as a whole this is an extensive data collection and a significant proportion of it is of direct relevance to the aims of PropBase, although a significantly smaller proportion is currently fully 'PropBase Compliant'. The resources available to PropBase in the next few years are limited and are unlikely to permit significant routine enhancement of existing digital data or conversion of analogue to digital data. It is, therefore, likely that individual projects will have to undertake any enhancement of these data to 'PropBase standards' that may be necessary as and when required and that bulk enhancement will not be possible at present.

All digital data added to corporate datasets should in future include XYZ co-ordinates (or XY co-ordinates and depth from a specified datum and the datum elevation) as well as LexRock code attributions. Where necessary, corporate data tables will need to be modified to allow these attributes to be added. Without these attributes, datasets will be of minimal use to PropBase.

### 4.2 'PRIVATE' DATA

An unknown number of datasets held by individuals in the BGS are relevant to the aims of PropBase. These are data not currently held in a 'corporate' database (nor necessarily backed up) and are, therefore, vulnerable to loss or accidental corruption. Data known, or believed, to exist in the BGS and held by individuals include in particular most data on Mineralogy/Petrology (see 6.5 below) and an amount of borehole geophysical log data (though the latter are currently in the process of being added to the corporate collection). Other data are likely to exist and a brief trawl through parts of the W drives suggests that there may be a significant amount of data held here, though not all relevant to PropBase. A review of various non-corporate datasets held, for example, in Murchison House is presented in Campbell (2006), but a thorough assessment of what may be available in all locations within BGS is required. While it is important that PropBase is involved in such an assessment, and it would be important to establish whether such data are 'PropBase Compliant', this should be undertaken at a corporate data management level. Guidelines on data management need to be re-enforced to ensure that all relevant data are managed in corporate databases.

The benefits of corporate management of data need to be re-emphasised to encourage migration of 'private' datasets into corporate management. This may need to be supported by the application of some form of sanction where corporately valuable data are not migrated.

### 4.3 PROPBASE CORE DATA

Table 1 lists the disciplines that hold datasets of significance to the aims of PropBase. This overview is only to the discipline level and is split into measured data and derived data. Apart from mineralogy/petrology, the availability of measured data (or, at least, databases to store the data!) is fairly good. It must be noted, however, that many of these datasets may need enhancement before they become a resource for PropBase. The main area that requires development by PropBase is that relating to derived data (creating or adapting databases to store the data and metadata on derivation, etc, as well as the generation of derived datasets, such as porosity or rock strengths from wireline geophysical logs) or the generation of such data 'on the fly' to meet specific project requirements. It is important to note that specific property information is commonly relevant to two or more disciplines (e.g. porosity, density, etc.) and, therefore, that the discipline boundaries are 'artificial'. However, the parameters are likely to have been measured by different methods and/or to different accuracies and an important aspect of PropBase will be to ensure that users are aware of the limitations of the data to which they have gained access.

Table 1: PropBase Core Datasets by Discipline

Discipline	Existent	Current Extent	Notes
<b>Measured Data</b>			
Geochemistry	Yes	Very Good	For G-BASE but not complete coverage as yet - other geochemistry fairly poor. Limited depth data.
Geophysics (field, borehole and laboratory derived data)	Yes	Good	Variable density. High density in areas of mineral and groundwater resources
Geotechnics	Yes	Good	Variable density
Hydrogeology	Yes	Good	Variable density
Mineralogy/petrology	a few	Poor	
Palaeontology/Biostratigraphy	Yes	Good	
<b>Derived Data</b>			
Geochemistry	No	N/A	
Geophysics	Yes	Fair	
Geotechnics	Yes	Good	From categorisation of LexRock codes
Hydrogeology	Yes	Good	From categorisation of LexRock codes
Mineralogy/petrology	No	N/A	

Palaeontology/Biostratigraphy?	?	From stratigraphical information
--------------------------------	---	----------------------------------

Table 1 is arranged alphabetically and is not prioritised. All these disciplines provide important rock and soil property data and are, therefore, difficult to prioritise in order of relevance.

For rock property information to be useful, say for providing information for the population of 3D models, etc., the data need to be attributable to lithostratigraphical/lithodemic units (and other less formal stratigraphies such as those established for artificial deposits) and to be located in 3-dimensional space. This can be achieved if all of the data are attributed with 3D coordinates and LexRock codes. It is perhaps beyond the scope of this report, but one of the issues that will need to be addressed in future by PropBase is how (or even if) data from the same location (or sample) in different tables/databases are linked by the use of a primary (unique) key. It is likely that to revisit existing digital data will be impractical and that a pragmatic solution utilising attributes already included in many database tables, for example 3D co-ordinates, is the most appropriate way forward. The requirement for, and the implications of, such an approach, and the most suitable attributes, need to be considered in detail at an early stage in the next phase of PropBase.

## 5 External Datasets

### 5.1 KEY DATA HOLDERS

Many of the potential external clients of PropBase are also holders of geoscientific property data, which may be relevant to PropBase. The data are likely to comprise especially Site Investigation reports, including borehole records and laboratory test results for a wide range of geotechnical data, monitoring data and geochemical data (soils and groundwater) and in some cases geophysical data. The key external data holders include:

- statutory bodies such as the EA and SEPA and the local authorities;
- commercial developers, consultants and contractors, and;
- academic institutions.

BGS already holds large volumes of site investigation data including related laboratory results acquired from statutory bodies, local authorities, consultants and contractors. However, it is known that many reports are not acquired for a variety of reasons. Chiefly amongst these appears to be the limitations of resources in local authorities for what they would regard as the low priority task of extracting and forwarding reports. Many authorities do not know that we collect site investigation data/reports and will only provide site investigations when reminded on a project basis. The burden could be substantially alleviated if the data were already available in digital form, e.g. AGS format data (see below) and could therefore be readily transmitted. At present only a limited, but increasing, number of AGS format site investigations exist. There are problems of getting hold of the data including those that we should get as a matter of course, for example, the Highways Agency. This would also relieve the burden on the data holder of storage of the data.

Data reception from stakeholders in AGS Format would be likely to improve data capture overall due to reduced time and costs in handling, and would facilitate downstream usage and third party transfer.

Geotechnical, engineering geological, environmental and hydrological/hydrogeological consultants are all likely to retain substantial databases of geoscientific property data. Such data, if acquired on behalf of local authorities or statutory bodies, may be accessible, with the compliance of the client body (and for, example, in Scotland under the standing instruction to do so from the Scottish Executive). However, those data deriving from work for private clients may be less likely to become accessible, unless it can be demonstrated that collective ownership of such data is advantageous in terms of greater efficiency, or, as in the case of DEAL, data are made available collectively subject to appropriate license conditions and fees.

Using the DEAL model, BGS may be able to develop a system for PropBase whereby consultants/contractors and local authorities can discharge their legal obligation to retain site investigation, etc. data (typically for seven years) by transferring them at a suitable time (end of works, settlement of any claims, etc.) to a discrete framework within the NGDC.

It will be particularly important to test the extent to which consultants and contractors currently acquire/maintain geotechnical data in AGS format and to assess their willingness to submit data to BGS in this format. In this regard, the support of major clients who own the data, and especially the local authorities, statutory bodies and Government Departments/devolved Administrations, would be essential. A list of consultants and contractors who have registered their AGS format document with the Association of Geotechnical and Geoenvironmental Specialists in the UK is contained in Appendix 2 of Campbell (2006). This list contains many of the leading geotechnical and geoenvironmental companies operating in the UK, and who currently provide ground investigation and laboratory data to BGS.

Some potentially relevant data to PropBase are acquired and held by universities, etc. However, if they are the owners of the data, the data are likely to be limited in number and localised in coverage, given the limitations of the universities in data gathering. If the data are not owned by the university, they should instead be obtained from the primary sources. Therefore, universities, etc. are unlikely to be a major source of data relevant to PropBase, although some specialised data, for example some specific types of geochemical data not routinely acquired by BGS, may be worth accessing for PropBase. NERC has relatively recently re-enforced its policy of acquiring copies of all data collected or generated by work funded on their grants in universities. It is the responsibility of the NERC Data Centres, including NGRC, to obtain and manage these data. A new project has been set up within the IM Programme to target the acquisition of such "academic data" as part of the BGS Earth Science Academic Archive.

Miscellaneous groups might hold some potentially relevant material that may be useful to fill gaps in datasets. For example, information and samples collected from quarries and other localities are held externally by local RIGS (Regionally Important Geological and Geomorphological Sites) groups.

It is important that BGS is pro-active in gathering appropriate third party information to enhance our data holdings. This is already done extensively in some disciplines. Clearly the organisation cannot accept and keep all data that are offered. However, it is likely that digital (and analogue) datasets of which we are unaware, which would be of scientific value to the BGS, are disposed of by others. PropBase should be responsible for at least raising awareness among data owners about potential data donation of key PropBase datasets. It should also be responsible for reviewing the suitability of any



data offered for donation, including the acceptability of any terms and conditions that may be applied to the use of the data, and recommending whether or not they are fit for purpose from a PropBase point of view.

## 6 Disciplines

These brief discipline summaries are ordered alphabetically as in Table 1 and no priority is implied by this order.

### 6.1 GEOCHEMISTRY

In general, the management of geochemical data in the BGS is maintained at a high standard and the Geochemistry Database, comprised mainly of G-BASE and MRP datasets, as well as being one of the earliest digital datasets in the BGS and the first to be managed corporately, is probably the most comprehensive (Coats; 2004, Johnson et al; 2004 and Johnson; 2005), but the data are of limited depth extent. All data entered into the database have XY co-ordinates and depth attributes and the sample site geology is recorded. The latter is currently to BGS dictionary standards and earlier 'pre-Lexicon' geological descriptions can be matched to current standards via look up tables. As such, G-Base is 'PropBase compliant'. All data in G-Base have unique keys composed of site number, sample type, 'project code' and a sample number system code to enforce uniqueness of sample numbering of samples derived from different projects within BGS. Samples are collected according to prescribed methods and standards (Johnson et al; 2003) and analytical methods are well documented. Quality statements can be attached to all data by attributing a qualifier field.

G-Base holds data from the UK land geochemical survey (stream sediments, stream waters and soils). The data held are essentially inorganic elemental determinations and a number of other parameters, such as pH, CEC, Eh, TDS, etc. Routine determinations of organics are not undertaken.

Outside the G-Base project geochemical data are less well managed. Radiochemical determinations are not systematically databased and nor are data from site-specific contaminated land, etc investigations. Whereas NIGL isotopic determinations are published routinely, they are not databased systematically. Bulk rock geochemical analyses have been added to the corporate Geochemistry Database but not comprehensively.

Geochemical data are now common in urban site investigation, so-called environmental investigations. However, these data are not currently added to geochemical databases but may be added to the National Geotechnical Property Database.

Some information on databases containing geochemical data collected off-shore is provided in Campbell (2006).

### 6.2 GEOPHYSICS

Geophysical data potentially relevant to PropBase include wireline, remote sensed and surface geophysics data and derived data and laboratory measurements.

Geophysical ("wireline") logs of boreholes are usually understood in terms of the interpretations that can be derived from these data, such as, sedimentological or stratigraphical interpretations. However, fundamentally the data recorded are

effectively physical properties (or close proxies thereof) measured at close spacings along the length of a borehole. As such these logs provide a 1D continuous sample of rock properties through a rock mass and provide a highly valuable source of information from which to make predictions of 3D rock property distributions. More details are provided in Shaw (2006).

As physical property measurements they are a fundamental dataset that must be included in consideration of the design of PropBase. In particular, geophysical borehole logs from deep boreholes are one of the very few datasets with genuine national coverage of high density, high quality physical property data across the UK. Any attempt to understand the physical property variation in the UK in 3D will require a data framework that is based either wholly or largely on values derived from geophysical borehole logs.

Of the various log data holdings available to BGS, those collected for coalfield, oil exploration and radwaste disposal have left a legacy high-quality dataset.

It is important to be aware that geophysical logs record bulk, in-situ measurements so, for example, density logs are recording bulk rock density of the sediments and fluids of the borehole environment at the point of measurement and not the grain density, which excludes the volume of pore water measured.

The ability to understand, for example, rock strength and density variations at depth is potentially very important to civil engineers. However, in practice for the immediate future, one of the core 'markets' for PropBase output is likely to be in supporting the understanding of aquifer properties.

Understanding of porosity and permeability (and hence transmissivity) distribution in three dimensions, including any linear features, e.g. sealing faults or clay stringers, that might affect lateral flow, is important in the construction of groundwater models. Data supporting this understanding will be key outputs for PropBase. Improved understanding of the distribution of these parameters is needed, together with knowledge of the confidence limits that can be placed on the derived data. Other key information includes the reliability of data used to decide on boundary conditions in groundwater models and the distribution and significance of fracture networks in flow and transport. Understanding the evolution, including diagenesis, of aquifers helps greatly in defining models and assessing the confidence that can be placed in resultant predictions.

BGS holds a variety of seismic data, especially the vast amount of offshore data, but also on-shore data from oil and coal exploration and the Nirex investigations, which, together with other geophysical data will be important in conditioning the attribution of 3D volumes. Campbell (2006) provides some further details of the off-shore datasets.

As part of the Nirex Sellafield Investigations, BGS was involved in a project that attempted to attribute 3D rock volumes for which high-quality 3D seismic data had been acquired with rock quality and bulk rock permeability indices. These were derived from wireline log data that was calibrated with the results of core testing (Brereton; 1997). The key link between the seismic survey data and the wireline log data was acoustic impedance, which can be determined from both datasets. Whereas uncertainties about the methodology remain (Evans et al; 2003 and Kingdon et al; 2003), the method offers potential as a means of populating 3D models with rock property attributes that are related directly to specific rock volumes rather than geostatistical distributions or average values. This approach merits further evaluation by PropBase when resources permit.

The basic quality and reliability of geophysical logs and also the understanding of hydrogeophysics have improved very significantly in recent times. This means that

aquifer properties are more clearly understood and can be more easily modelled than has previously been possible. However, it is essential that such data are of the highest quality if meaningful quantitative interpretations are to be derived.

### 6.3 GEOTECHNICS

One of the key databases for PropBase is the National Geotechnical Properties Database (NGPD), population of which was started in the mid 1980s. This database (fully described by Self; 2006) contains data extracted from site investigation reports produced by commercial geotechnical contractors for various clients.

Originally, a coherent series of project-specific flat-file geotechnical datasets were created during the applied geological mapping of Exeter, Deeside, Coventry, Nottingham, Bath, Castleford/Pontefract and the Black Country, and mapping of the Thame 1:50k sheet. For each of these projects data were abstracted from the available site investigation reports within the project areas.

A further series of project datasets, for Wrexham, Leeds, SW Essex and Stoke, were created in the late 1980s. All of the data for each project were stored on paper datasheets that were later input into a computerised spreadsheet, replicating the design of the analogue datasheets. These spreadsheets were later combined and restructured, and stored as a project database.

The Association of Geotechnical and Geoenvironmental Specialists (AGS) established a common Data Interchange Protocol (First Edition, 1992; Second Edition, 1994; Third Edition, 1999) which has become widely accepted by the geotechnical community. The NGPD was, therefore, redesigned in order to maximise compatibility with data provided in AGS format and, from 1992, this has meant that data presented in site investigation reports could be transferred electronically. This also provided a means whereby large amounts of data could be received digitally in AGS format and loaded directly into the database, although the quantity of AGS format data that has been made available to date is variable, and, for example, no such data have been submitted to Murchison House, other than on a project-specific basis.

Data from all 'Engineering Geology of UK Rocks and Soils' projects have since been entered in the NGPD and most of the data from the earlier projects have been updated to the newer format, but data migration from the Leeds and Stoke projects remains incomplete.

Further upgrades to the database in 2003 allowed data to be loaded directly into the corporate Oracle tables, and corporate audit triggers were added.

The present database has been designed as a stand-alone, AGS compatible, database for use within 'Engineering Geology of UK Rocks and Soils' projects. It was hoped that data from the NGPD would be made accessible via the DGSM Portal, but this work is currently 'on hold'. It was intended that this database would be compatible with BGS corporate databases, such as the Single Onshore Borehole Index (SOBI) and BGS.BOREHOLE\_GEOLOGY (BoGe). This has not proved possible, partly because AGS datasets are added to the NGPD when they arrive, but the update of SOBI requires more processing.

The NGPD was designed as far as possible to mirror the structure of data received in AGS format, with each table in the database representing an AGS data group, although, not all AGS data groups are represented by a database table.

Importantly from the PropBase point of view, the attributes stored include lithostratigraphy and lithology codes and text descriptions for depth intervals down the borehole. The codes used to identify the lithostratigraphy are taken from the BGS

Lexicon. The databases also included 3D co-ordinates in the form of XY co-ordinates, and depths and borehole OD where available, and can relatively easily be converted to XYZ co-ordinates and so are effectively 'PropBase compliant'.

New data can be added manually or by digital download as noted above. The database can currently be queried using SQL or an MS Access front-end query.

A number of future developments are under consideration. These include the completion of the migration of all geotechnical data into the NGPD, extending the database to include BGS Laboratory data from non-borehole sources and the creation of a user-friendly interface to the data. The current MS Access data query system allows data to be selected by area, type and geology. A future development could be to rewrite the system in Cold Fusion and add it to the Intranet as part of the IDA system. As the data are spatially referenced they could be integrated with a GIS data retrieval system, and the positions of boreholes and pits in this database are held in the GDI.

Campbell (2006) provides some further details of the off-shore datasets that contain some geotechnical data.

## 6.4 HYDROGEOLOGY

BGS corporate datasets for the UK include aquifer properties, water quality and borehole, well and spring information for c.107,000 sites, about 97% of which are in England and Wales. These data are widely used for commercial enquiries and to provide background and raw data for CR and SB projects. Some data are organised so that they can be accessed from the WellMaster relational database via the GDI platform. Digitisation of this vital hydrogeologically-oriented system is now advanced, with more than 80% of 10 km x 10 km sheets in England and Wales now processed. However, there are a number of valuable datasets, mainly but not exclusively located at Wallingford, that are not yet subsumed into this comprehensive and flexible system, being instead located in separate databases that have accumulated *ad hoc* over many years. These databases have attributes and features not currently available within WellMaster that are useful to (and employed by) users. They need to be brought into the GDI, RECALL and in some cases integrated into the WellMaster system where they can be accessed by users.

The principal areas of information involved are aquifer properties, hydro-geophysical borehole logs and hydrogeochemical data. Details of groundwater databases managed from Wallingford are provided by Gale (2006).

As noted elsewhere, one of the delivery mechanisms proposed for PropBase is to be through the GDI. The GDI will continue to evolve towards a web-based system accessible through the internet/extranet as well as the intranet. Data will be stored on the IDA and databases will form a live basis for interrogation and presentation, including value-added interpretation to a range of levels. One route to access relevant information will be through the Groundwater Portal.

Tasks to achieve the objectives of PropBase therefore fall into three categories, the first two needing to be integrated into the BGS Information Structure to ensure compatibility and to avoid duplication, and the third requiring iteration with the other two:

- Entering current and backlog data into databases;
- Continued development of systems to facilitate access to data, either in its raw form or as derived datasets, maps, etc;

- Develop the “user end” of the system to produce queries to service both internal and external enquirers, e.g. enquiries service, water companies, EA, consultants.

These categories are summarised in Table 2 of Gale (2006), to act as a basis for the identification of tasks that can be tackled in the short- and medium-term, as well as ongoing.

## 6.5 MINERALOGY/PETROLOGY

Mineralogy and petrology can potentially provide a wide variety of data that are relevant to PropBase objectives. Mineralogy and petrology properties have a fundamental influence on the geochemical, geotechnical, geophysical and hydrogeological properties of rocks, soils and sediments, and as such these data can be extremely useful for predicting or explaining the physical and chemical behaviour of geological materials. For example:

- The presence or absence of swelling clay minerals (e.g. smectite) has a major control on the plasticity, shrink-swell characteristics, ion-exchange and sorption characteristics of mudrocks;
- The fabric microstructure is fundamental to understanding the collapse behaviour of metastable soils such as brickearths (loessic soils);
- The degree and distribution of cementing minerals such as calcite or quartz exert major controls on the porosity and permeability of sandstone aquifer rocks;
- Groundwater and soil porewater chemistry is strongly influenced by the nature of, and interactions with, the mineral surfaces that it contacts.

As such, mineralogical and petrological data are potentially very valuable attributes that could be applied to 3D geological models in a large number of applications.

Unfortunately, the management of mineralogical and petrological information within the BGS is very poor to virtually non-existent, and no formal corporate mineralogical and petrological databases currently exist. In fact, little systematic corporate archiving of mineralogical data has been carried out (or supported) within BGS, following the merger of the Petrology Unit and Applied Mineralogy Unit, and their subsequent transfer from London to Keyworth in the mid-1980s. However, a considerable amount of data exists in paper records and as BGS reports, but much of this is non-digital. Digital mineralogical and petrological data are stored largely within project-specific databases and Excel spreadsheet files that are not maintained under the corporate umbrella, or individual BGS scientists hold the data informally. Another major problem is that there are no metadata to describe the type of information available adequately, and to identify its location. Knowledge of the data rests largely with the specialist scientists who sometimes regard the information as ‘theirs’ and, consequently, the data are not readily accessible to corporate users.

However, a very large amount of mineralogical and petrological information is held by the BGS, and that could potentially be used to attribute 3D geological models. The principal sources of mineralogical and petrological data and datasets are summarised here, but are discussed in more detail in Shaw (2006).

### 6.5.1 Petrographical Notes and Petrographical Reports

‘Petrographical Notes’ comprise a large series of formalised proforma descriptions of individual thin sections in paper record format.

'Petrographical Reports' is a series of short reports summarising more comprehensive petrographical studies. These are based primarily on work carried out in support of the primary geological survey activities.

### **6.5.2 Engineering Properties of UK Rocks and Soils Project**

Mineralogical and some petrographical/petrological data have been acquired in support of 'The Engineering Properties of UK Rocks and Soils Project'. The data are focussed on UK mudrock lithologies and are traceable to parent samples for which location data are archived within the Engineering Properties of UK Rocks and Soils project databases.

Bulk and clay mineral, and surface area data are available for the "Lias", Mercia Mudstone Group, Gault Formation, London Clay Formation and "Brickearth", and data are also currently being acquired for the Lambeth Group. It covers the stratigraphical range of each specified rock unit, and also a broad geographical area. The data vary from semi-quantitative to quantitative mineralogy, held in Excel spreadsheet format, within project-specific workspace on corporate drives. These mineralogical data are not included within the UK Geotechnical Database.

## **6.6 THE NIREX DIGITAL GEOSCIENCE DATABASE**

The Nirex Digital Geoscience Database (NDGD) was constructed and operated by the BGS on behalf of United Kingdom Nirex Limited (Nirex) to manage and store the geoscientific data being produced from their site investigations at Sellafield and Dounreay between 1989 and 1997.

Although restricted in its geographical coverage specifically to Sellafield (west Cumbria) and Dounreay (Caithness), the NDGD is a relational database that contains one of the largest single datasets of mineralogical and petrological data available within the BGS. The data are also well-constrained and fully integrated with other geoscientific information obtained from the same (or related) samples, including: geology, stratigraphy, major and trace element whole rock geochemistry, geophysical, geotechnical and engineering properties, hydrogeological properties (poroperm), and palaeontological.

### **6.6.1 PADAMOT Project Palaeohydrogeological database**

The PADAMOT Palaeohydrogeological database is a fully-relational database designed by the BGS to handle palaeohydrogeological information, and was set as part of the EC Framework V PADAMOT Palaeohydrogeology Project (McCormick et al., 2004). Through arrangement with Nirex, the PADAMOT database continues to be maintained under BGS corporate management. This database contains detailed information for specific sites in the UK and elsewhere.

These data are linked to full 3D geospatial data, geology; groundwater geochemistry; chemistry; palaeohydrogeology/palaeoclimate evolution, fracture distribution and geometry.

The PADAMOT database was designed in close liaison with BGS petrologists, and is capable of handling a wide range of geological, mineralogical, petrological, isotopic and geochemical data, and includes petrographical images from optical microscopy, CL microscopy, SEM, BSEM and electron microprobe X-ray elemental mapping. However, due to resource limitations, not all data obtained from the PADAMOT project have been included in the database.

The format and design of the PADAMOT database could potentially be expanded to form the basis of a more general database for mineralogical and petrological information for wider use within the BGS.

### 6.6.2 Other datasets

Significant amounts of mineralogical and petrological data exist in BGS Technical Reports, and in non-corporate project-specific databases, or are held by individual staff either in hand-written formats in project/laboratory notebooks or digitally on individuals' PCs. Some of these data might be available digitally but they are not readily accessible as little or no metadata are available, and knowledge of the whereabouts and type of data is largely limited to individual mineralogists. Datasets exist from single sample descriptions up to several hundreds of samples per project. Many of these legacy data are potentially very valuable but would require significant effort to compile and enter into databases, and would necessarily require the involvement of individual specialists. They include:

- Electron microprobe microanalytical data.
- Electron microprobe X-ray maps.
- Petrographical images.
- Heavy Mineral Collection.
- BRITROCKS. (This is the main rock collection maintained corporately by the BGS in the NGDC.)

### 6.6.3 Types of data produced by mineralogy and petrology

A very complex and wide variety of mineralogical data and data formats are currently or potentially produced by the BGS that need to be catered for, and captured in, any future databases. These are listed in Shaw (2006).

## 6.7 PALAEOLOGY/BIOSTRATIGRAPHY

Superficially, it is arguable that biostratigraphy and related disciplines are of no relevance to the aims of PropBase and it might well be true that many users do not require this type of information. However, this discipline provides fundamental information, used, for example, defining the geological boundaries, etc. needed for the construction of digital 3D geological models and, as such, is highly relevant to the aims of PropBase. The High-Resolution Borehole Stratigraphy Dataset is a fully 'PropBase compliant' stratigraphical database, with XY co-ordinates, depths and LexRock attributes. The database has been set up in prototype for the Kimmeridge Clay Formation, and a demonstration prototype of a Biostratigraphical Geographical Information System (for the Chalk formations of the Winchester district) has also been set up (see Woods, 2006) for more details). Information held in the former will be useful for defining the geological boundaries within which rock property attributes can be applied, and hence will be a valuable asset to PropBase. Additionally, the variability of properties within some formations can be related to the more specific properties of biostratigraphical units.

Campbell (2006) provides details of the off-shore datasets that contain some biostratigraphical data and of palaeontological datasets in general that are maintained in Murchison House.

## 7 Dealing with Legacy Data

The BGS holds a vast quantity of legacy data, much of which is of relevance to PropBase. Whereas some of these data are available digitally and are held in corporately managed databases, a significant number are available only in analogue (paper) form or are held in a variety of non-corporate ('private') forms and therefore at risk of loss or corruption in the future.

Whereas an ultimate aim is that all information relevant to rock and soil properties should be available digitally and attributed in such a way that it can readily be synthesised to provide user-specified summary information, realistically, achievement of such an aim is beyond the resources likely to be available to PropBase in the short-to medium-terms. PropBase can only provide resources to deal with high priority data requirements for which there is a good justification and a clear corporate need. Projects requiring data will have to fund, or largely fund, the data entry and any related activities needed to prepare such data for addition to 'PropBase compliant' databases. Focused use of unfunded time to undertake these activities for high priority or specialist data might provide an appropriate means of gradually bringing datasets up to 'PropBase standards'.

Given the quantity of legacy data, methods of bulk attribution should be investigated by PropBase to establish whether and how some of the existing analogue and digital data can be brought up to 'PropBase standards'. It is appropriate that this be undertaken in a pilot area where existing data coverage is good.

For data to be suitable for PropBase they must have information on location, lithology and lithostratigraphy (or compliance with another approved stratigraphical scheme). The former will be provided by accurate (or at least as accurate as attainable) 3D grid co-ordinates (X, Y and Z) and the latter by allocation of LexRock codes to all databased information. In the absence of the former, the latter are critical, though would only allow data to be used in compiling summary information, and without either, data are of minimal use to PropBase. All relevant existing (and new) databases need to include these attributes if they are to be used by PropBase.

If resources are available from PropBase, or elsewhere, to deal with any legacy data consideration should be given to focussing work on datasets that are managed or held by staff approaching retirement. This is especially important where the BGS has minimal expertise in the dataset elsewhere in the organization.

## 8 Attribution of 3D rock volumes

Currently, most 3D modelling undertaken in the BGS is the modelling of surfaces, such as unconformities, major lithostratigraphical boundaries and faults, and little modelling of 3D volumes has, as yet, been undertaken. Modelling of the latter is required to enable 3D distributions of rock mass and soil properties to be modelled. How property distributions are to be applied to the modelled volumes will be dependant on various issues including the availability of a number of datasets. This might vary, in the case of rock masses, from a simple attribution of model volumes using statistically derived average rock property values, through rock property distributions derived geostatistically where data volumes permit this type of approach, through to variations of the approach, trialled by Nirex and noted in section 6.2 above.

Modelling of spatial variability of rock mass properties by the use of geostatistics (there might be a university collaboration opportunity here) will be an important



approach. Other approaches, such as identification of structural domains, as was undertaken for Nirex (1996), and using these to constrain attribution, will help to constrain 3D distributions.

## 9 Uncertainty

Whereas there might be uncertainties associated with measured data and their spatial information, these are likely to be understood and be quantifiable (but they still need to be recorded with the data). However, the creation of derived information will have variable, but largely identifiable and quantifiable, uncertainties and it is essential that these also are documented with the derived data. Subsequent use of any of these data, for example to provide attribution of defined volumes in 3D digital geological models, already having inherent uncertainties, must also take account of these uncertainties. Uncertainty modelling is currently being carried out for 3D models and these methods developed under the DGSM Programme could be applied to PropBase.

Where appropriate, it is important that quality/methodology information is recorded with all data, so that the data can readily be assessed for their fitness for purpose.

One of main uncertainties is commonly likely to be related to XYZ coordinates. In an ideal situation accurate 3D location is highly desirable but if the rock property data can be attributed with Lexicon and Rock Classification Scheme (RCS) codes they can contribute effectively to summary information on units without accurate spatial location. Lexicon and RCS coding are critical for the use of property data for PropBase related activities, and, whereas 3D co-ordinates that are as accurate as possible are highly desirable, they are not necessarily essential.

Whereas uncertainty modelling is outside its scope, where possible PropBase needs to provide access to the information that allows such modelling to be undertaken by projects as and when needed. This will include information relating to methodologies, best practices, etc, followed during sample testing, quality statements and calibration records.

## 10 Potential Clients

### 10.1 INTERNAL CLIENTS

PropBase provides an important means by which to expand the culture of corporate ownership of data within BGS and there is general support internally for its development.

The main internal clients of PropBase are considered likely to be in relation to:

- revision mapping;
- urban mapping;
- engineering geology;
- hydrogeology and groundwater modelling;
- enquiries services and data sales;
- geohazard evaluation and GeoSure;

- building stones;
- modelling seafloor changes and coastal erosion;
- attribution of hydrogeological, geotechnical, geophysical and mineralogical property information to LithoFrame, DigMap, and the BGS Lexicon;
- research into property information relationships and the effect of geo-processes, and;
- 3D and 4D modelling of property information, including research into statistical methods to summarise the data and to define uncertainty.

Internal usage in relation to marine geology, seismology and geomagnetism appears likely to be limited as access to relevant data is already provided to a substantial degree. However, in future, further consideration should be given to this issue.

## 10.2 POTENTIAL EXTERNAL CLIENTS

A comprehensive review of potential external clients from the Scottish perspective has been undertaken and reported in Campbell (2006). This is to a large extent applicable to the whole country, and is the basis of the following summary.

The main potential external clients are likely to be:

- Central Government and the devolved administrations, and their consultants;
- Local Authorities, and their consultants, especially engaged in:
  - Planning;
  - Environmental Assessments;
- The EA and SEPA, and their consultants;
- The natural heritage bodies and their consultants;
- NGOs, including those involved in conservation;
- Nirex and similar organisations;
- Commercial clients, consultants and contractors, including:
  - Construction industry and planning;
  - Extractive industry.
- Academics.

It is likely that the demands of a new programme in the UK for the selection of a site for radioactive waste disposal will mean that such a programme will be a significant potential external user of PropBase information. Resources available to the other potential users mean that their use of PropBase data may well be variable, and possibly limited.

There is reasonable current demand in some areas for reliable and up-to-date geoscientific analysis and information for the urban environment, for example in the areas covered by the Clyde and Manchester projects. The Clyde Basin Environmental Project has demonstrated a specific Local Authority demand for linkage of onshore soil geochemical data with sediment geochemistry in near-shore and estuarine environments. There is also a similar need apparent for linkage of geotechnical data in these environments, for example in relation to the Clyde Waterfront Regeneration Plan and the related Clyde Gateway project.

At present, enquiries for offshore property data come mainly from oil companies or consultants to the oil industry, and are typically related to the planning of offshore site investigations. They are generally borehole or area based and will request any geotechnical data that may be available. Typically the data are used either to help them plan their own survey, or to satisfy insurance or legislative requirements.

Potential growth areas in relation to demand for data are seen in relation to offshore wind farms and tidal power schemes, where foundation conditions would be the issue, and in relation to seabed trenching for power lines, etc.

## 11 Problems to Address

### 11.1 GENERAL

Following completion of the baseline geological mapping programme around 2010, revision mode will predominate. This will link logically to PropBase, because the borehole and other property data that will be accessible through PropBase will be important data on which map revisions are likely to depend. The integration of project areas (onshore and offshore) in terms of scales and consistent handling of data will also be essential.

It is vital in this context that PropBase is:

- seen as a long-term infrastructural project,
- able to capture new data efficiently.

The data must be:

- internally consistent (e.g. to BS 5930 (Anon., 1999), BS 1377 1-9 (Anon 1990) :1990; UKAS accredited laboratories, etc.), and
- in a form that can readily be accessed digitally rather than as hard copy, and with mechanisms to resolve IPR issues efficiently.

The data should be in a form capable of being used to:

- update existing maps (DiGMap);
- upgrade existing 3D models, and;
- develop other corporate datasets, including the BGS Lexicon.

### 11.2 LEGACY DATA

Large volumes of legacy data identified in this report appear potentially useful to a wide range of end-users. However, several overriding problems are apparent in integrating these datasets, namely a:

- lack of common standards,
- lack of consistent definitions,
- lack of internal consistency,
- lack of previous validation, and
- lack of ready accessibility.

If left unresolved these problems would significantly undermine the defensibility, and hence the commercial viability, of the data, and so would limit the extent to which they could be used.

### 11.2.1 Standards

The metadata for most of the BGS corporate datasets are generally insufficient to resolve many of the uncertainties in terms of methodologies and equipment used to acquire the data, the levels of precision attained, and the definitions of the parameters quoted. However, in some instances, these uncertainties can be resolved by referring to supporting data in published documents (such as in-house Technical Reports, Site Investigation Reports, etc.) and files. In other cases the data quoted might have been acquired subject to a British or international standard, or acquired by an organisation subject to UKAS accreditation.

### 11.2.2 Definitions

The issue of definitions is key to the linking, integration and expansion of datasets. Where different definitions are used for the same term in different datasets, their linkage potentially degrades the value of both, and will be masked by subsequent quantitative analysis. The problem of lack of definitions is also exacerbated by the lack of generally accepted definitions for some important concepts that are highly relevant to PropBase. The scope and definitions of relevant terms defined in the Corporate Thesaurus need to be reviewed by PropBase to confirm suitability. New terms will need to be recommended for incorporation in the thesaurus if necessary.

The classification of, and related definitions of, superficial deposits, including the related classification of artificial deposits/made ground/derelict ground, have been addressed internally by McMillan and Powell (1999). McMillan et al. (2001) present useful definitions for engineering soils and engineering rock from BS5930 (Anon., 1999).

BGS's urban geoscience projects follow the British Standard for lithological description as far as possible, so that data provided by the projects are interchangeable between most end-users. However, this would not necessarily be the case for other BGS lithological data, which (since 1999) would generally conform to the BGS Rock Classification Scheme. The engineering properties of soils should be acquired by soil mechanics testing techniques (BS 1377:1990 and its earlier versions).

Colour in BGS's urban geoscience projects is recorded using the widely-adopted Munsell soil or rock colour charts, but these are not used in all other areas of BGS's work, although they are used, for example, for projects relating to building stones.

Texture, discontinuities (fractures and joints) and weathering are again described in urban geoscience projects using British Standard 5930:1999 descriptors.

With regard to the in-house mineralogical and petrological data, many are largely descriptive (thin section, hand specimen and exposure descriptions) and there have been no established corporate standards for such descriptions, though in 1993 some guidelines were included in the procedures of the (then) Thematic Mapping and Onshore Surveys Division of the BGS. However, by contrast, more recent GeoReports provide petrographical descriptions that conform to British and European standards.

### 11.2.3 Internal consistency

Inevitably some dataset contents are subject to considerable operator bias, therefore affecting the internal consistency. For example, the sets of structural data include data

collected by many individuals, generally working independently. As a result, although definitions of features measured may be unambiguous, the identification of the measured features might not be.

The in-house mineralogical datasets are subject to substantial operator bias in terms of their accuracy and content. These largely descriptive data commonly reflect the particular skills and interests of the individual, and the aims of the project in relation to which the descriptions were made, and so can generally be regarded as partial descriptions only.

The geochemical datasets are amalgamations of data obtained from a variety of laboratories (in-house and external), using a wide range of equipment and analytical techniques with varying levels of analytical precision. Although individual sets of data would typically be calibrated to reference samples material, in many cases the calibration data are not readily available.

Physical and mechanical data are obtained from a very large number of external sources and, even if they apparently conform to specified standards, there might be inconsistency between different data providers. The problem is exacerbated where the data pre-date any effective data gathering standards.

#### **11.2.4 Validation**

There is little information available on validation of the datasets and it appears that in many instances, the datasets have been subjected to little or no validation. With most of the older legacy data this shortfall would be impossible to rectify, and provision of retrospective validation in other instances would require considerable resources.

Ultimately, with respect to the legacy data in particular, there will be many gaps in the datasets that cannot be filled. Therefore, it will be important in such cases to acknowledge that these gaps exist, by use of one of the following:

- ‘Not applicable’, meaning information definitely ‘Not applicable’ in this context,
- ‘Not available’, meaning information would be ‘Applicable, but a value cannot be found despite a search’, and
- ‘Not entered’, meaning ‘a value has not yet been assigned (and it might not be applicable)’.

#### **11.2.5 Accessibility**

The data are held in various hard copy, analogue (e.g. magnetic tape), and digital (spreadsheets, databases, relational databases) formats and there may be issues of compatibility.

Many of the data are scattered within hard-copy files and reports, which may or may not be indexed, and the useful data are commonly hard to identify and retrieval and validation are labour intensive.

Some magnetic tapes (containing seismic reflection data) are in uncertain condition and retrieval of digital data from this storage medium might not be possible in all cases.

In some cases data are covered by confidentiality agreements and PropBase access must preserve these (assuming such agreements remain extant).

### 11.3 CURRENT DATA

Most of the problems associated with integrating legacy data also affect current data. With current data many of these problems are more readily overcome, such that overall data quality is higher, and their defensibility is more readily established, and therefore their corporate usefulness and their resulting commercial viability are both greater.

## 12 The way forward

Tasks to achieve the objectives of PropBase fall into three categories:

- Develop the “user interface” (PropBase Data Portal) system to query databases to service both internal and/or external enquiries;
- Entering current and/or backlog data into databases;
- Continued development to facilitate access to data in their raw form as well as providing access to derived datasets (map, databases, etc).

A number of specific tasks, discussed below, have been identified to meet these needs.

### 12.1 IMMEDIATELY (REMAINDER OF THIS FY)

1. Undertake a feasibility study to examine the merits of expanding the scope of the recently developed Geochemical Properties Interface to accommodate a wider range of physical properties to be accessed through the ‘PropBase Data Portal’. Initially this should be by extending it to include porosity data. Porosity is one of the key PropBase datasets and data are currently held in a number of corporate databases and can also be derived from wireline logs. As a first step, the ability to access, extract and summarise porosity data will be a valuable tool and the developments needed to achieve this will address most of the issues likely to be encountered for other PropBase datasets.

2. Undertake a feasibility study to examine potential methods of bulk attribution of existing digital data that are considered important to PropBase so that the issues of upgrading existing datasets to be ‘PropBase compliant’ can be understood and costed. It is proposed that this study be focused on the Midland Valley of Scotland, especially the Glasgow Integrated Project Area, because this is considered to be a representative and well-advanced ‘urban’ project that already utilises various property data encompassed by PropBase. The project also has significant stakeholder involvement, and stakeholder views could be sought during this process.

3. The modification of existing corporate database tables to include Z co-ordinates calculated from depth data and datum elevations should be investigated and costed for those datasets relevant to PropBase.

4. All internal stakeholders should be canvassed for their views on the key PropBase datasets identified in this report, to further assess priorities and to ensure that all appropriate datasets have been identified.

5. A few external stakeholders, selected to be representative of potential external users of PropBase related information, should be identified and their views on, and priorities for, PropBase sought. These could include, for example, the Glasgow Integrated Project Area participants, Nirex and ‘Government’ users (e.g. EA).

6. Various commercial software packages are available, such as Core Lab Reservoir Information Browser, for the management of rock property information. These are tailored largely to the needs of the oil industry but might provide an ‘off the shelf’

solution to some of the needs of PropBase. They should be evaluated early in the project to assess their potential suitability.

7. Start designing a publicity strategy on PropBase to potential stakeholders and data sources.

## **12.2 SHORT TERM (YEARS 1 AND 2)**

8. 'PropBase data standards' need to be defined and corporately agreed, and 'PropBase relevant' (or even all) corporate databases should be modified as far as possible to reflect the PropBase data standards. As noted in this report it is considered that the inclusion of XYZ co-ordinates and LexRock codes for a sample will make data sufficiently 'PropBase compliant'. Once these modifications have been completed all new data entered into databases will conform to the 'PropBase data standards'. The target is that all future rock and soil property data acquisitions are PropBase compliant, (i.e. incorporated in databases whose structure is PropBase compliant, even if not all existing entries are). Other attributes may be important for some rock properties.

9. Key 'missing' PropBase datasets will be identified and corporate databases established to hold them. The databases will subsequently be populated with new data to PropBase standards. The main discipline that currently lacks extensive corporately managed datasets is that of mineralogy and petrology, and creation of corporate datasets in this area must be prioritised. Legacy data for these datasets will have to be managed with the other legacy data.

10. A key focus early in the PropBase project will be the development of the 'PropBase Data Portal' and its integration with existing BGS data access tools, such as the DGSM Portal, so that data can be accessed, extracted and added to the PropBase datasets. It is expected that this will build on the porosity feasibility study noted above. (Estimated staff effort, based on the DGSM Portal project, for portal development is £75-100k).

11. Write publicity strategy and keep potential stakeholder and users informed of progress. Identify what data are needed and how they should be provided. (AGS and other organisations might be useful).

## **12.3 MEDIUM TERM (YEARS 3 AND 4)**

The 'PropBase Data Portal' will be enhanced to allow data to be summarised, etc and to provide property information in a suitable format for use in attributing 3D geological models.

With completion of some of the initial developments identified above, resources should become available during this period to allow some systematic enhancement of key corporate datasets to make them 'PropBase compliant'. This will need to be prioritised on user needs and will be limited by resource availability.

## **12.4 LONGER TERM (YEAR 5 AND BEYOND)**

Ongoing development and maintenance of the 'PropBase Data Portal' will be required to provide additional functionality and to utilise improved software/hardware, etc. that are likely to become available in future.

An ultimate aim will be to ensure that all 'PropBase relevant' datasets are suitably attributed and available through the 'PropBase Data Portal', though the resource implications for achieving this for all legacy data are vast and currently beyond the

PropBase project. With time, advances in methods of data capture, particularly from analogue datasets, will permit a more sophisticated and cost effective means of dealing with legacy data.

## 12.5 RESOURCES

It is likely that resources for the development of PropBase will be limited to about £80k to £100k per year for the next few years. This will place severe constraints on what PropBase can achieve, and prioritising activities will be important. Targeted use of unfunded time could be a valuable tool in dealing with upgrading legacy data to PropBase standards. This could make use of the expertise of available individuals to bring datasets in which they have an interest up to PropBase standards.

Estimated costs for the activities noted above are provided in Table 2 below. It is stressed that these are first-pass estimates and that these costings may be subject to significant revision in the light of experience gained during the remainder of this financial year (2006-7).

Activity	Cost (£k)
<b>Immediately</b>	
Feasibility study of the development of the Geochemical Properties Interface to include other properties and delivery through the 'PropBase Data Portal'. Trial area looking at porosity.	15-20
Feasibility study to examine potential methods of bulk attribution of legacy datasets.	10
Investigate the modification of existing corporate database tables to include Z co-ordinates by contributing to the Borehole Users Group review of this issue.	10
Canvass internal stakeholders re key PropBase rock properties.	5
Solicit views of a few key external users of PropBase related information.	5
Review any available commercial software packages for suitability/cost.	5
<b>Short Term</b>	
Define and agree 'PropBase data standards'.	10
'PropBase relevant' corporate databases modified to reflect the PropBase standards.	25-50
Agree key 'missing' PropBase datasets and establish new corporate databases as required.	15
Expansion of scope of 'PropBase Data Portal' to all relevant datasets.	75-100
Evaluate and agree latitude/longitude vs. National Grid co-ordinates issues.	10
<b>Medium Term</b>	



Enhancement of 'PropBase Data Portal'.	20/yr
Systematic enhancement of key digital datasets to be 'PropBase compliant'.	80/yr
<b>Longer term</b>	
Ongoing development and maintenance of the 'PropBase Data Portal'	5-10/yr
Continued systematic enhancement of digital datasets and data digitisation.	70-100/yr

Table 2: Cost estimates for PropBase.

## 13 Key Recommendations

### 13.1 ADDITIONAL EVALUATION

Initially, key and later all, potentially relevant corporate datasets will need to be examined in detail to determine how easily they can be made 'PropBase compliant' for future data entry and to assess the magnitude of the task to bring all existing data to the same standard. This should include an assessment of user needs for each dataset.

### 13.2 SYSTEM/DATABASE DEVELOPMENT

Development of part of the 'PropBase Data Portal' will be an early win. Expansion of the Geochemical Properties Interface developed at Wallingford (Chris Milne) to handle other data as the PropBase Data Portal will be straightforward. This can be done for porosity, data for which are held in several corporate databases and can also be derived from wireline logs, and which is a fundamental rock/soil property used by several disciplines. While this is being undertaken the issues relating to the integration of the 'PropBase Data Portal' into various BGS Intranet and data search applications can be fully evaluated.

In the longer term the Portal will need to be expanded to allow access to, and synthesis of, all corporate digital rock and soil property information. PropBase will be one of several similar 'portals' providing access to corporate data and this development needs to be undertaken in such a way as to ensure future compatibility.

### 13.3 DATA NEEDS AND PRIORITIES

The data handling aspects of PropBase are a massive undertaking. The main (and perhaps the only realistic) attributes are the combination of lithostratigraphy (etc) and lithology, and spatial information (X, Y, Z). These will allow data to be summarised in order to extract and synthesise data and therefore to populate, for example, rock properties of 3D volumes. All new data entered into all corporate databases should have these attributes and many current databases may need to be modified to include this. Obtaining and entering this information in respect of existing datasets will be an enormous undertaking and, if it is to be done for all data, cannot realistically be resourced in the short or medium terms by PropBase. However, opportunities to utilise unfunded time to undertake some of this task could be taken if they arise and could be focussed on the expertise of available staff.

### 13.4 PROJECT MANAGEMENT

PropBase should be overseen by a project Board, which should meet a minimum of once per year. The Board should be kept small and it is suggested that its initial membership is:

- Programme Manager Physical Hazards Programme;
- Programme Manager Groundwater Management Programme;
- Programme Manager National Geoscience Framework Programme;
- Programme Manager Information Management Programme.

If appropriate, this can be varied in due course and other relevant guidance can be acquired on an ad hoc basis, depending on project-specific-needs.

The PropBase Project Manager will also attend.

Ad hoc meetings should be held between various user groups and the PropBase Project Manager. These will allow the Project Manager to inform users about progress, etc within PropBase and to receive feedback from user groups on their specific requirements.

## References

- Anon. 1981. BRITISH STANDARD 5930: 1981. *Code of Practice for Site Investigations*. British Standards Institution, London.
- Anon. 1991. BRITISH STANDARD 1377:1991. *Methods of test for soils for civil engineering purposes*. Volume 1-9. British Standards Institution, London.
- Anon. 1999. BRITISH STANDARD 5930: 1999. *Code of Practice for Site Investigations*: British Standards Institution, London.
- Brereton, N.R. 1997. Porosity, permeability and the scale of measurement - an integrated approach to rock mass property predictions. British Geological Survey Report WK/97/14.
- Campbell, SDG 2006 PropBase - Murchison House Perspective British Geological Survey Internal Report, IR/06/089. 77 pp.
- Coats, J S. 2004. The BGS Geochemistry Database: history, design and current usage. *British Geological Survey Internal Report*, IR/04/033. 29pp.
- Evans, C J, Cuss, R J and Kingdon, A. 2003. Wireline Geophysical Logging of the Nirex Deep Borehole in the Sellafield Area: Review of the Status of Wireline Logs and Their Use to Measure Porosity. British Geological Survey Commissioned Report, CR/02/166N. 27pp.
- Gale, I N, Morris, B L and McKenzie, A A. 2006 PropBase Scoping Study - Hydrogeological data. British Geological Survey Internal Report, IR/06/090. 11 pp.
- Johnson C C, Brown S E and Lister T R. 2003. G-Base Field Procedures Manual version 1.1. *British Geological Survey Internal Report*, IR/03/096N. 102pp.
- Johnson, C C, Coats, J S, Breward, N, Ander, E L and Mackenzie, A C. 2004. Geochemical data as a standard reference data set for the SIGMA project. *British Geological Survey Internal Report*, IR/04/026. 77pp.
- Johnson, C C. 2005. 2005 GBASE Field Procedures Manual. *British Geological Survey Internal Report*, IR/05/097. 130pp.
- Kingdon, A, Evans, C J and Cuss, R J. 2003. Sellafield borehole geophysical data: Acquisition Parameters, Environmental Corrections and QA status. *British Geological Survey Commissioned Report*, CR/01/065N. 67 pp.
- McCormick, T, Nayembil, M and Giles, J R A. 2004. Design and Compilation of Database: Final Report. PADAMOT Project Technical Report WP3. 92pp.
- McMillan, A A, Ellison, R A, Fordyce, F, Golledge, N R, Holmes, K A, Jordan, C J, Kessler, H, Monaghan, A A, Northmore, K J, Robins, N S and Taylor, C. 2001. Superficial deposits characterisation in the urban environment, a best practice guide to mapping and research. *British Geological Survey Internal Report*, IR//01/68.

- McMillan, A A and Powell, J H. 2001. BGS Rock Classification Scheme Volume 4: Classification of artificial (man-made) ground and natural superficial deposits applications to geological maps and datasets in the UK. *British Geological Survey Research Report*, RR/99/04.
- Nirex. 1996. Sellafield Geological and Hydrogeological Investigations: Structural Atlas of the Sellafield Boreholes.
- Self, S. 2006. BGS UK Geotechnical Database. *British Geological Survey Internal Report*, IR/06/092. 78pp
- Shaw, R P. 2006. PropBase – Keyworth Perspective. *British Geological Survey Internal Report*, IR/06/091. 42 pp
- Woods, M A. 2006. A scoping study for the digital capture, storage, retrieval and use of high-resolution stratigraphical data: a preliminary report. *British Geological Survey Internal Report*, IR/06/025R. 13pp.