DESIGN AND IMPLEMENTATION OF “FACTUAL” DATABASES TO SUPPORT GIS/RS ANALYSES OF EARTH SYSTEMS FOR CIVIL SOCIETY

Tsehaie Woldai [*] , Ernst Schetselaar

Department of Earth Systems Analysis (ESA), International Institute for Geo-information Science and Earth
Observation (ITC), P.O. Box 6, 7500 AA, Hengelosestraat 99, ENSCHEDE, The Netherlands, Woldai@itc.nl ,
Schetselaar@itc.nl

KEY WORDS: ‘factual’ databases, field data capture, geographic information systems, geoscience, multidisciplinary surface categorizations, remote sensing

ABSTRACT:

With the increasing access to various types of remotely sensed and ancillary spatial data, there is a growing demand for an independent source of reliable ground data to systematically support information extraction for gaining an understanding of earth systems. Conventionally maps, ad-hoc sample campaigns or exclusively remotely sensed data are used but these approaches are not well suited for effective data integration. This paper discusses avenues towards alternative methods for field data acquisition geared to the extraction of information from remotely sensed and ancillary spatial datasets from a geoscience perspective. We foresee that the merging of expertise in recently developed mobile technology, database design and thematic geoscientific knowledge, provides the potential to deliver innovative strategies for the acquisition and analysis of ground data that could profoundly alter the methods how we model earth systems today. With further development and maturation of such research methodologies we ultimately envisage a situation where planners and researchers can with the same ease as they nowadays download remotely sensed data over the web, download standardized and reliable ‘factual’ datasets from the surface of the earth to process downloaded remotely sensed data for effective extraction of information. This could result in a worldwide reduction of duplicative ground surveys and at the same time provide a wealth of information that is far more relevant to the societal needs, as the deliveries of today.

1. INTRODUCTION

The notion that Geological Survey organization are the sole keeper of information only is slowly changing to accommodate the needs of the civil society with up to date and real time information on dynamic geoscience processes. Clastic sedimentary systems, barrier island systems, delta’s etc. can be studied in the present as well as in the past. Neotectonic activity and seismicity demonstrates that suture zones and fault systems are dynamic geologic systems in terms of geometry and stress state. Natural hazards (landslides, flooding), desertification processes, changes in the coastal zone etc. again illustrate the dynamics of system Earth. Acknowledging this implies a change from a static resource inventory to a dynamic, process-based approach with the ultimate goal of not only gathering information on natural geologic resources, but aiming at describing and modeling the dynamics of the environment and the processes driving the Earth dynamic system (Bohlen et al., 2001).

Many of the developed countries are well aware of these pressing demands and are aiming at delivering nation-wide geologic map databases that will better serve society. Geoscience data are used in the search for mineral resources, for scientific research, for land-management decision making, engineering design, for assessing geological disasters, for groundwater exploration, etc. The surveying methods, however, remain strongly focussed on monodisciplinary map production. Despite the complex task of acquiring data from the Earth surface by human observation, only a fraction of the knowledge acquired in the field is represented. Many of the observed spatial relationships between objects are not systematically captured. As a result only a fraction of the acquired knowledge is represented in spatial databases, and often limited to the generalized and abstracted form of geological maps. This synthesized cartographic form of representation evolved because objects associated to a particular theme (e.g. geological units, soil units) are to variable extent covered by objects from another theme (e.g. vegetation). The delineation of partly covered objects as coherent mapping units, therefore, relies on substantial interpretation of their continuation in the near subsurface. Moreover, there is no method to objectively classify geospatial objects and their relationships. The accumulated knowledge and experience of the surveyor bred within a particular theoretical school, inevitably results in subtle differences in weighing criteria to identify and classify objects. This often results in a particular mapping style, which may further be affected by advancing geoscience theories and the evolving working hypothesis used as a model for the study area.

Despite these aspects of subjectivity in natural resource mapping and years of effort that went into it, only one interpretation is published. Many of the fundamental datasets recorded in the field are not available and as such this data reside in inaccessible formats, such as notebooks or reports disaggregated from their positional attributes and fragmentarily distributed within the syntax of human language. The real data is usually only structured in the notebook or worse the mind of the surveyor and therefore lost forever when he or she retires from the institute.

---

[*] Corresponding author.

In developing countries geological map coverage is far from complete. Geological map data are widely distributed, poorly linked and scattered around too many institutions. The geodetic network is often not established and as a result most of the geological maps are products of uncontrolled photo mosaics or reconnaissance sketches (Woldai, 1995; Woldai and Limptlaw, 2000). Basic geological maps cover only parts of countries and many are based on outdated earth science theory that drastically changed in the late sixties and seventies by breakthroughs in the concepts of plate tectonics and uniformitarian settings. Most of all, survey organizations cannot afford to send their staff to fieldwork for a long time due to financial constraints. This problem is rather acute in light of the efforts that still need to be implemented for nationwide geological resource inventories.

A crucial managerial question for geological surveys in developing countries is if the nation-wide geological map coverage and production should be completed in the traditional manner in the form of (digital) maps. We argue that in the light of the limitations sketched above and huge investment requirements in terms of cost and training of personnel, it may proof worthwhile to restructure the surveying methodologies for natural resource inventories. This is not only relevant to better meet the demands of the user for reliable and relevant spatial geoinformation, but also to establish geodata infrastructures that exploit rapidly evolving earth observation technologies. In this paper we provide some arguments for developing more 'factual' representations of spatial data as a potential solution to reach these objectives.

2. WHAT DO WE MEAN WITH FACTUAL SPATIAL DATABASES?

The source data underlying natural resource maps are data collected by direct human observation, possibly augmented with chemical analysis of sampled materials and physical measurements by portable instruments. These data are georeferenced by traditional positioning increasingly replaced by methods based on global positioning systems (GPS). The word factual here is not meant in an absolute sense. Even direct observations in natural resource mapping are interpretations, as they are contextual in nature and subject to taxonomic and classification ambiguities. The data is considered 'factual' because it is based on direct observation at or in the near vicinity of a visited location (e.g. field station). According to this definition spatial data outside the field of view of the observer based on indirect inferences from observations between known occurrences are excluded.

Dependent on the scale of mapping, the observed geospatial objects could be too small to be represented as a two-dimensional object on a map. Nevertheless, the bounding surfaces between material volumes observed at a particular station may, because of it limited intersection with the Earth surface be the only piece of evidence on which to base the location and interpret the nature of boundaries between the units of a thematic map layer stored in a GIS. Although this type of spatial data provides the fundamental source to develop the spatio-temporal reasoning mechanisms underlying map interpretation, this data is usually not represented in the spatial data structure of a GIS.

The ambiguity in classifying an object has not only aspects related to its spatial extent, but also on the expertise required to classify a geospatial object correctly. In the context of land use mapping, there will be a general consensus to classify an object as an agricultural field, residential area, road or park. Such objects may even be considered 'factual' if they can be identified by interpretation of remotely sense data of high spatial resolution. However, the spatial extents as well as the nature and identity of objects become more subtle and open to dispute if one has to delineate and identify soil, rock or vegetation units. Clearly, the question if data collected by observation can be considered factual or not depends on the knowledge and experience that is required to make a statement on its spatial extent and identity at a particular level of detail in the taxonomic hierarchy pertaining to a particular geoscience discipline. Hence, surveying expertise should be a crucial component of any metadata scheme proposed to accompany natural resource inventories.

In summary factual spatial data can be considered known occurrences defining geospatial objects or parts of them by means of direct observation or measurement that can be attached to positions in geographic space. The identity of the represented objects are either based on formal taxonomies used by experts or belong to class types on which a general consensus can be reached among non-experts.

3. WHY DO WE NEED 'FACTUAL' SPATIAL DATABASES?

In every geological survey, geological maps have by far remained their most important thematic assets. As new data are collected, geoscience theories change and interpretations are challenged; there is no method of backing out the factual data recorded in the past, let alone deriving alternative interpretations from the same primary data source. Knowledge acquisition and its conceptualisation by mapping in the geoscience disciplines are inherently based on scientific inference from observed occurrences by experts. This fundamental knowledge acquisition is often incremental, complex, uncertain and dynamic, and frequently results in multiple valid models for a geographic region (Brodaric and Gahegan, 2001). Hence the geological map, depicting a generalized spatio-temporal synthesis of the substrate according to a particular school of thought, provides invaluable background information but contributes very little to delivering specific information of the substrate that society demands for decision-making. Under current surveying techniques society is getting a small benefit from a major effort.

The very same lack of site-specific ground reference data for meeting society demands hampers progress in developing realistic and cost-effective applications of remote sensing analyses. Contrary, to the level of sophistication of today in image processing algorithms for the extraction of information from RS data, little has been done to develop data acquisition methods on the ground that support their application. This is arguable related to two kinds of impediments in the applications of GIS in an operational context:

a) The lack of an infrastructure that provides access to the source data and associated knowledge underpinning the thematic map layers in a GIS.

b) The lack of structured methods oriented towards establishing relationships with remote sensing measurements and ground observations.

The first deficiency is currently being resolved at geological surveys in a number of developed countries (Australia, Britain, Canada, and USA) where geologic concept modelling has been an active field of research since the late nineties. For example, a North American Map Data Model NADM has been developed that is under continuous refinement in different physical implementations (Soller and Berg, 2000). The NADM, which later evolved in the National Geologic Map Database NGMDB, provides standards and a consistent framework for geologic concept modelling from the hand sample to a regional tectonic province, providing information relevant to land management, engineering design, earth resources, natural hazards and scientific research. From the experienced gained in developing prototypes for the NGMDB, it was shown that such conceptual data models will ultimately include a coherent knowledge base to check the consistency of existing information against new interpretations and data (Richard, 1999).

To overcome the second hurdle is, as apparent from the lack of structured methodologies and the nature of the problem more daunting. Although field campaigns based on ground-based remote sensing devices or sampled surface materials have been widely used in establishing causal relationships between surface characteristics and remote sensing measurements, the difference in represented surface area is at least an order of scale in magnitude. The immediate problem that arises is to what extent the sampled surface segments can be considered representative for the heterogeneous surface area covered by the instantaneous field of view of the sensor. Moreover, the semantic issues of translating the object modelling approach used by humans to the field approach used as a model for the spatially continuous distributions of remote sensing measurements over geographic space, appears to be an overwhelmingly complex problem (Camara et al., 2001).

Hence, from a user s perspective there is a demand for reliable databases of ‘factual’ data that model the aggregated nature of multiple geoscientific themes at the earth surface. The challenge is to build bridges between the intrinsic information content of datasets build from monodisciplinary geoscientific perspective and remotely sensed data to arrive at a more complete understanding of the multidisciplinary problems addressed by society. In the words of T.S. Eliot:

Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?

Working towards a solution will ultimately deliver better methods for integrated analyses and information extraction and lead to an increased understanding of the potential and limitations in the use of remotely sensed data. Some of the advantages envisaged are:

Methods to improve on the archiving, modelling and representation of factual geoscientific datasets that are less dependent on advances in geoscience theories and allow developing alternative map interpretations from a common knowledgebase.

An increased understanding of the relationships between remote sensing measurements and human landscape characterization, enabling better integration of knowledge and data-driven methods for the extraction of information from remotely sensed datasets.

An increased potential for separating target responses relevant to a theme within a monodisciplinary context from unwanted responses from themes from other disciplines.

Holistic and integrated approaches of information extraction that better meet societal demands related to environmental monitoring and interdisciplinary approaches of resource management.

4. DEVELOPMENT IN EARTH OBSERVATION THAT FOSTER THIS CHANGES IN FOCUS

Research in geological remote sensing has since the launch of the first earth orbiting platforms, been focussed on spatial characterization of geological patterns at the earth surface. Geological information has traditionally been extracted by human interpretation (photo-geology) and is gradually augmented (but only partly replaced by) numerical extraction methods.

In principle all geological remote sensing applications aim to use 2-D image representations of the surface and subsurface to ultimately understand 4-D geological systems. The physical properties measured by the sensors provide indirect “signatures” off the finite states of objects in geographic space. It is from these finite states that earth scientists try to unravel the internal and surface processes that shaped the Earth through geologic history (deformation, sedimentation, erosion, metamorphism, mineralisation, metasomatism, hydrothermal alteration, weathering etc.). Seen in this light, remote sensing should not only be used to tell us where distinguishable objects are and how they coincide with geological features, as they have been mapped and classified on the ground. Ultimately they should tell us *what* things are and *how* they came into being.

An understanding of the genetic processes by which geological resources are formed and how they give rise to remotely sensed “signatures” is a research field on its own that strives maximising the contribution of RS in assessing the spatial distribution and geological significance of imaged features. A process-based perspective in developing this multidisciplinary field of earth science and earth observation is considered of fundamental importance in understanding the distribution of geological resources.

4.1 Status in Earth Observations Systems

Almost 30 years after the first Earth Observation Satellite (EOS) went to orbit, optical and microwave remote sensing such as, NASA *Landsat MSS/Thematic Mapper, SPOT, IRS, ERS and JERS* still remain the most popular and the most used remotely sensed data in many scientific researches and productions. The number of scientific publications and maps covering many applications outweighs the purpose of this paper.

Unfortunately, the applications derived by such remotely sensed methods, particularly in the field of geological mapping, have been widely oversold, ignoring limitations related to their applicability in the real world (Drury, 1993):

- No single system provides all the information to fully characterize the features of interest.

- Many of the sensors provide information of the earth's surface only (no depth penetration).

- Mineralogical compositions cannot be uniquely identified nor their proportions be precisely estimated (a multitude of minerals and their mixtures may yield similar spectral features or do not exhibit diagnostic features at all).

- Direct estimation and identification of mineralogical composition is hampered by weathering materials (desert varnish), soil and vegetation covers.

- Key information such as relative age-relationships, sedimentary and tectonic structures, petrographic textures, visible in outcrop, can not be identified with any conceivable remote-sensing instrument.

To apply remote sensing for geological resource management, it is gradually becoming clear to the user-community that besides the on-going development of automated spatial and spectral feature extraction methods, field surveys directed to the understanding of the underlying geological processes are important to make progress in an operational context.

4.2 New Developments

Earth observation in the past was highly mono-sensor, present and future. Earth observation space missions have a multi-sensor character. The major benefit is that they provide synoptic overview and a cost-effective and timely monitoring potential. In the field of natural disasters, they allow one to monitor the events during the time of occurrence while the force is in full swing. A glance at the present activities of the major players in the Earth observation field from the technological perspective, NASA, ESA, etc. learns that long term missions comprise of complementary but far better sensor quality, higher spectral/spatial resolution and better calibration.

The increase in spatial resolution of data acquired by newly developed earth orbiting sensors will provide cost-effective access to framework data and mapping bases. An important development in this direction is *IKONOS*. This sensor system acquires panchromatic imagery with one-meter spatial resolution and multispectral imagery at four meters. With ground control, the imagery boasts a two-meter horizontal and three-meter vertical accuracy, equivalent to 1:2,400-scale map standards. The satellite's ability to swivel in orbit enables it to collect imagery anywhere on earth with a revisit frequency of just one-and-a-half days.

Potential applications for one-meter satellite imagery in a GIS environment are limitless. The imagery can serve as detailed base map upon which thematic map layers can be overlaid, or it can be used as an up-to-date data source from which various geological features, land cover, soil degradation, hydrology and other activities related to elevation features are extracted to populate multiple GIS layers.

With the operational use of sensors with higher spectral resolutions, such as ASTER, the quantification of the composition of earth surface materials becomes feasible. *ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)* uses 14 spectral bands as opposed to Landsat TM's 7 bands. These bands were carefully chosen to differentiate a number of rock types. It has 5 bands compared to the Landsat TM single band 7. This is in an area where OH and H_2O bending and stretching can distinguish many important rock types such as limestone from dolomite, various clays and sulfates. In addition it has 5 bands in Thermal infrared and these are capable of mapping silicate mineralogy.

ASTER data also serve to obtain maps of land surface temperature, emissivity, reflectance and elevation. *One of the most important aspects is that ASTER data is currently accessible (free of charge) to the public; a trend unheard in space data acquisition.*

NASA's present flagship is the TERRA platform that complements the general purpose Landsat program. On board of TERRA are various dedicated instruments that, like ASTER no longer provide raw data that require extensive calibration (system, atmospheric and geocoding), but geophysical products (radiance, reflectance, emissivity, temperature, height, deformation, etc). Typical instruments of such missions include interferometers, lidar (laser altimeters), imaging spectrometers and the like.

ESA follows a similar policy. The planned *Envisat mission* once in orbit, will involve a laser altimeter, a SAR interferometry system (ASAR), an imaging spectrometer (MERIS) etc. An important issue within the context of ESA is the Global Monitoring for Environment and Security (GMES; part of a larger framework of three such monitoring systems, the G3OS) initiative.

Data collected by *airborne spectrometers* have already demonstrated that it is possible to identify certain types of exposed mineralogy, to label the minerals present and to determine the fractions of the minerals occurring in small, sub-pixel units. Thus, provided that the non-uniqueness in the solutions obtained from spectral unmixing methods can be constrained, a new type of maps, indicating "mineral abundance at surface", can now be made. This will help explorationists to home-in on zones of mineral alteration around mineral deposits, detect previously unrecognised mineral patterns across whole mineralised districts, document mineralogical components of the weathered regolith and to locate waste products, such as sulphate minerals, causing acid mine run-off from mine tailings.

Pioneering studies so far carried out using *SAR interferometry* (InSAR) have already being earmarked as a new development in understanding our earth and its dynamics. InSAR can provide with unprecedented precision, high-resolution images of earthquake-prone areas, topographic data (DTM's using stereopairs of radar images with differing viewing angles) and a map of coseismic deformation generated by an earthquake. The precise monitoring of surface deformation allows accurate zoning, mapping and prediction of volcanic eruptions, landslides and ground subsidence. Differential interferrometry allows one to measure surface movements with sensitivity of the order of a few centimeters over large surfaces (Massonnet et al., 1993).

In most developing and some developed countries accurate topographic base map and digital terrain model (DTM) of the area under investigation is missing. In any GIS work therefore, this problem remains a handicap. An exciting development towards solving this acute problem is envisaged from the new *Shuttle Radar Topographic Mapping (SRTM*) acquired by Space Shuttle Endeavour in February 2001. The SRTM instrument captured allows one to create very detailed topographic maps of the Earth's surface using interferometry. This radar system gathered data that will result in the most accurate and complete topographic map of the Earth's surface that has ever been assembled. Once processed, the SRTM radar data will allow to obtain accurate knowledge of the shape and height of the land, and to assess: flood, soil degradation, deforestation and reforestation, landscape changes. SRTM was launched into an orbit with an inclination of 57 degrees. This allowed most the Earth's land surface that lies between 60 degrees north and 56 degrees south latitude to be covered by the SRTM radars. This is about 80 percent of the Earth's land mass.

4.3 Future Trends

Turning to the future, spaceborne remote sensing offers a more challenging task to the geoscientific community. Three new data types and their technologies are about to become a commonplace. These include:

- Multi-attribute imaging radars having multiple frequencies and polarisations.
- Highly detailed digital elevation models (DEM) of anywhere in the world, allowing sophisticated modelling and visualisation of the landscape and processes acting on it. Such data will become extremely valuable for better modelling geophysical and geochemical responses including geological processes, determining geological structures and logistics and also for visualising regions prior to fieldwork.
- Imaging spectrometry or hyperspectral imaging systems, such as the planned Australia's ARIES-1, using sensors with 100's of spectral bands providing spectroscopy and compositional information about earth surface materials based on the principles of spectroscopy. This brings the exciting new possibilities of not just discriminating the materials but actually identifying it and putting a label to the major mineral components present in every pixel of an image.

5. DEVELOPMENT IN GIS THAT FOSTER THIS CHANGES IN FOCUS

Modern image processing facilities and methodologies of digitally formatted data has revolutionized the interpretation of large-scale planetary landscape scenes. Personal computers can nowadays handle large amounts of remote sensing data, providing access to universities, resource-responsible agencies, small environmental companies, and even individuals.

The data required to understand geological or geoscience problems is coming from different disciplines which need integration in order to arrive at sound multi-purpose geological and environmental maps. From the perspective of the user data integration is one of the strongest element of Geo-Information System (GIS). The zonation, integration and modelling of various geoscientific data should supply planners and decision-makers with adequate and understandable information within a relatively short period of time.

Desktop-GIS with full-fletched image processing capabilities is now being used for a wide variety of applications ranging from environmental assessment to marketing. For mineral exploration desktop-GIS provides the means to statistically analyse and classify geochemical data, enhance geophysical images, determine spatial relationships between features, and produce charts, tables and maps to report the analysis. Pattern recognition, feature extraction, texture analysis by mathematical morphology, and a variety of unsupervised and supervised classification techniques (including neural nets) represent techniques of information extraction increasingly used in the earth sciences. Integration modeling by favourability functions has provided the mathematical framework for predictive spatial data analysis, a developing but still relatively unexploited area of GIS analysis (Chung and Fabbri, 1993).

Computerized field data capture using GPS, palm-top and laptop computers, field spectrometers and other field sensors, are becoming critical inputs to systematic spatial data analysis of maps and images. They anchor to ground truth much of the pre-processing of images to obtain spatially and spectrally corrected images for further processing.

Digital geologic mapping applications have also rapidly evolved over the years. Data that are digitally input directly in the field are preferable because they reduce error and save time in the data processing and interpretation stage. The importance of digitising information directly in the field to reduce error and duplicity in the map production process cannot be underestimated. There have been several attempts to make digital field mapping software since the mid 1980's. Brodaric and Fyon (1989) produced the first widely used field mapping system called Fieldlog. By the end of the 1990's field mapping crews in Canada were going right into final geologic mapping production without pen touching paper. Currently, Fieldlog is shareware that links with several GIS systems. Geomapper is a commercial PenMap software (for pen-based computer application) developed at the Berkeley Earth Resources Center Digital Mapping Lab., University of California (Brimhall et al., 1998) GeoMapper provides a computerized mapping legend which contains both the geological features needed to map the earth as well as a visual interface to use all the digital electronic equipment a user selects. The mapping tools include a pen stylus, which serves the purpose of a full set of colored pencils. In combination with digital topographic maps or color ortho-photos on the screen of a portable computer for positioning. Additional digital tools include sub-meter accuracy GPS, laser range finders, digital cameras and visible/infrared (IR) spectrometers (Kramer, 2000).

6. HOW HAVE WE DEALT WITH THESE DEVELOPMENTS AT ITC AND WHICH RESEARCH PROBLEMS HAVE WE IDENTIFIED THAT NEED TO BE ADDRESSED?

6.1 Developments within ITC

The Geological Survey Division (GSD), (since 2002 amalgamated Earth Systems Analysis) in its more than 40 years standing within ITC have managed to bridge the gap between the "traditional geologist" and the "computerised spatial analyst" and to keep the balance between a traditional geological background and the development of geoscience information system management. The GSD amalgamated, since January 2002, with the divisions of Mineral Exploration and Exploration Geophysics in a new department called Earth Systems Analysis, reflecting the current trend towards integrated approaches in analysing earth systems.

Since the 70's, ITC has developed as a centre of excellence in geo-information knowledge transfer. Training and education was geared to the maximum usage of tools acquired from satellite earth observations. This, together with aerial photograph interpretation has proven indispensable in mapping geological features and extracting many geoscientific details including natural hazard and pollution monitoring, water resources evaluation, topographic and relief mapping, ecosystem monitoring, monitoring of crop and plant health, land use and land cover mapping, forestry, etc. There are many examples of new geological discoveries in "well mapped" areas due to mapping with these data. All these are in pursuit of sciences dealing with features, processes and phenomena operating in the Earth's surface both in space and time.

The Division, realizing at least part of the problem discussed in sections 1-3, has since 1994 invested in applying *field data capture methodologies* in its educational programs. Field data capture techniques were for the first time implemented by the Division in a structural mapping exercise in the Rheinische Schiefer, Eifel Germany using Fieldlog (one of the first field data capture tools developed at the Ontario Geological survey, Canada (Brodaric and Fyon, 1989). Students of the GSD were equipped with laptop computers, GPS and small digitising tablets to digitally capture, display and output daily field observations. At the outcrop, site-specific data collected and recorded into standard field notebooks by day were later transferred to laptop computers (equipped with remotely sensed and other ancillary data) for analysis of this data in conjunction with framework and remote sensing data in a field camp GIS environment.

For the first time at ITC, Databases of source data were build and maintained over several years of progressing fieldwork activities. Students were able to analyse data collected by their colleagues of previous years and to efficiently upgrade the spatial database while mapping in the field. The power of querying site-specific field data in this project became immediately obvious, as it facilitated the compilation of various thematic map products, instead of one, and allowed to analyse and model the data in alternative and innovative ways (e.g. Schetselaar, 1995).

Since 1995 field data capture approaches have been used to support several inter-ITC Divisional M.Sc research projects with a fieldwork component, such as joint groundwater resource assessments, field spectroscopy, interpretation of RS datasets and geologic mapping. The approach proved also to be invaluable in research projects facilitating thematic map analysis and the compilation of real world training sets for image classification tasks (Schetselaar et al., 2000).

In 2001 personal digital assistants (PDA's) were introduced in the Professional Master fieldwork for data acquisition. Instead of the laptop used in past exercises, palmtops were equipped with RDBMS software for direct data acquisition in the field. A relational database was maintained by synchornizing daily observation to laptop GIS environment at field camps, tailored for geological data and spatial operations. The database structure and contents were designed in plenary sessions by intensive discussions with the students. This resulted in a factual database of the fieldwork areas, allowing students to process their data digitally from photo interpretation, field verification observation to the final map product. This methodology allowed for project driven database and map construction to proceed in the field itself, resulting in a digital representation of primary observations. This data can then be easily migrated to desktop and mainframe GIS environments. As the database is populated at the source, and intimately utilized during field mapping, the digital field data represents the most accurate data repository of the field survey and thus is ideal for integration into larger corporate GIS databases, without requiring duplicate data entry or error-prone data transcriptions sessions in the office. The same factual database served a number of applications, including the construction of WEB-based virtual field excursions and potentially serves a multitude of innovative approaches for analysis and visualizations. The database that was build in two subsequent educational fieldwork campaigns, allowed, for example to streamline flow of primary data to 3-D GIS environments to support 3-D modelling and visualization of structural elements (e.g. Schetselaar and de Kemp, 2001).

6.2 Problems Identified

The approach that we took at ITC, so far has been a pragmatic one serving the needs of educational groups and individual researchers on an ad-hoc basis. The data acquisition methods were tailored to the needs of specific applications, including structural geology, ground water resource assessment, geological mapping and the compilation of training datasets for supervised image classification methods. From these experiences, however, we identified several problems that need to be addressed to arrive at more generic and integrated approaches to arrive at more effective solutions:

1. How to represent knowledge in spatial databases that is important for the scientific synthesis of the study area, but is not suitable (in terms of its open structure, 3-D nature and complexity) to be integrated with the 'traditional' 2-D interpreted map layers? Geological examples include crosscutting and inclusion relationships between rock types observed in three dimensions at outcrops.

2. How to systematically encode field-observed features and ground-based physical measurements to better quantify relationships between objects inferred by human observations and remotely sensed data (at least in an empirical manner)?

3. How to designing data models that provides a structure for representing the multi-scale and complex phenomena of the natural environment without scarifying the currently developed user-friendly approaches.

4. How to build schemas for metadata that in some way allow us to quantify the uncertainty and semantic ambiguity in the cognition of geospatial objects. The schemas must also account for the different levels of geoscientific expertise and consider usability of the data by users not involved in the field survey.

6.3 How do we Want to Develop this Further from a User's Perspective?

We plan to design and implement alternative data models for archiving 'factual' geoscience datasets, underlying thematic map compilation that is also tailored to the extraction of information from remotely sensed data. Although each application will have specific user requirements, it should be possible to identify the common requirements for optimising GIS/RS spatial analyses. In the realist prospective, the process of representation of geographic reality involves the assignment of concepts to elements of the physical world, by virtue of collective agreement of a multi-disciplinary community that shares common perceptions (Searle, 1995). The recent developments towards ontology-driven GIS, where semantic integration can be achieved at a generic level at costs of loosing the details at the discipline-specific level, could provide part of the solution that we are seeking (Fonseca et al., 2002). We specifically need, however, generic categorizations that are congruent with remote sensing measurements. Considering that EM spectra from wavelength ranges (overlapping with the spectral range of human vision) used by optical sensor systems are mainly controlled by the molecular structure of chemical substances, would argue for a characterization of the earth surface that is congruent with classes that predominate the spectral variations of most multispectral scenes.

Textbooks on remote sensing, invariably consider (green) vegetation, water and soil-rocks as the most important categories controlling spectral variations (e.g. Lillesand and Kiefer, 1994). The recent implementation of core-modules on GIS and remote sensing where these spectral concepts are taught to all students enrolling in ITC courses supports our view that from an educational perspective, such a common spectral categorization must be feasible. Adding other categories, such as unconsolidated substrate materials that are not soil (weathering materials and sediments), organic litter (leaves), anthropogenic objects and pavements, and other consolidated materials (stones, logs etc.) would provide one framework for a first-order categorization to characterize most rural (and natural) landscapes.

Experiments on cognitive categorization, such as those conducted by Mark et al. (1999) on geographic objects, will need to be conducted to test if such categorizations provide workable solutions for the various expert and non-expert user communities. In addition research is needed how the aerial percentages of such compositional categories can be quantified in various physiographic settings. Modern surveying methods based on outlining classes with the help of differential global positioning systems have shown promising results for coastal areas (Donoghue and Mironnet, 2002). Other useful techniques for surface characterization may employ geostatistical methods for modelling landscape mereology using digital photography.

We envisage that much of the concepts for designing data model for the approach can be directly borrowed from established geoinformation theory. Yet emphasis in data modelling should be given to typical and unique characteristics of ‘factual’ spatial databases that are of fundamental importance for the effective extraction of information.

1. It accommodates those field-observed and field-measured properties and attributes that allow associating targets of interest with remotely sensed measurements.

2. It allows deriving site-specific spatial models of semantic and spatial granularity (e.g. models of mixed pixels) were surface categories can be generalized at different hierarchically arranged scaling levels.

3. It allows encoding the variability and nature of boundaries, the modelling of which relate to the semantics of geoscience disciplines. Data structures are sought to locally include 3-D spatial relationships at the intersection between 3-D objects and the Earth surface without the necessity to represent objects as closed polygons. This will facilitate interpretations on the basis of spatio-temporal reasoning mechanisms in geology (Ady, 1993). Such topologic concepts can be extended to the bounding interfaces between rock, soil and vegetation volumes, having a potential for ecologic characterizations of natural environments, difficult to foresee yet. Similar hybrid 2-D – 3(4)-D approaches are being developed in cadastral applications (Billen and Zlatanova, 2001).

4. It provides interfaces to other discipline specific non-spatial and spatial data bases including international standards as currently under development at the Open GIS Consortium.

5. It provides appropriate metadata schemas with quality measures for spatial accuracy and the uncertainty and ambiguity in classifying geospatial objects.

7. CONCLUSIONS

The design and integration of factual spatial databases including information originating from multiple disciplines and user applications at different levels of detail is a challenge. The research towards this aim is at its preparatory stage and a clear solution is not at hand. In our opinion, further developments should rely on a close cooperation between geoinformation and geoscientific expertise, both available at ITC. A multidisciplinary integrated approach is envisaged to be essential for optimising feature extraction from remotely sensed data where targets of interest to a specific discipline are ‘hidden’ in mixed responses from different ground cover themes. On the other hand such an approach would foster building integrated conceptual models of the Earth surface founded on the over centuries matured concepts in geologic, soil and vegetation sciences supporting multidisciplinary issues related to environmental monitoring and natural resource management.

REFERENCES

Ady, B.E., 1993. Towards a theory of spatio-chronological relations for geoscience. Unpublished M.Sc. thesis, Department of Geology, University of Toronto, 75 p.

Billen, R. and Zlatanova, S., 2001. 3D spatial relationships model: A useful concept for 3D cadastres? In: Registration of Properties in Strata, Oosterom van P.J.M, J.E. Stoter and E.M.Fendel (eds.), Proc. International Workshop on "3D Cadastres", Delft, The Netherlands, 28-30 November, pp. 232-244.

November 28-30, 2001 Department of Geodesy, Delft University of Technology, Delft, the Netherlands.

Bohlen, S.R., Halley, R.B., Hickman, S.H, Johnson, S.Y., Lowenstern, J.B., Muhs, D.R., Plumlee, G.S., Thompson, G.A., Trauger, D.L. and Zoback, M.L., 2001. Geology for a Changing World: A Science Strategy for the Geologic Division of the U.S. Geological Survey, 2000-2010. U.S. Geological Survey Circular 1172. Available at http://pubs.usgs.gov/circular/c1172/.

Brimhall, G., Stanton, E., Nunnelley, G., Hillman, B., Kramer, J., 1998. Direct Digital Field Mapping Using Pen-Based PC

Computers Supported by Differential Global Positioning Systems and Laser Range Finders. Geological Society of America Annual Meeting, Abstracts with Program, 30 (7), p. A-256. Available at http://www.geosociety.org/cgi-bin/annual-meeting-abstracts-query.pl.

Brodaric, B. and Fyon, J.A., 1989. OGS Fieldlog: a microcomputer-based methodology to store, process and display map-related data. Ontario Geological Survey, Open File report 5709, 73 p.

Brodaric, B and Gahegan, M., 2001. Learning geoscience categories in-situ: implications for geographic knowledge representation. Proc. ACM-GIS 2001, 9th ACM international symposium on GIS, Atlanta, GA, Nov. 9-10. http://www.geovista.psu.edu/publications/NIMA/ACMGIS200.pdf.

Câmara, G., Egenhofer, M., Fonseca, F. and Viera Monteiro, A.M., 2001. What's in an image? http://www.spatial.maine.edu/~max/imageOntology.pdf

Chung, C.F. and Fabbri, A.G, 1993. The representation of geoscience information for data integration. *Nonrenewable Resources,* 2, 159-169.

Donoghue, D.N.M. and Mironnet, N., 2002. Development of an integrated geographical information system prototype for coastal habitat monitoring. *Computer & Geosciences,* 28, 129-141.

Drury, S.A., 1993. Image interpretation in geology, 2nd ed. Chapmann & Hall, London, 283 p.

Fonseca, F., Egenhofer, M., Davis, C. and Câmara, G., 2002. Semantic granularity in ontology-driven geographic information systems. *Annals of Mathematics and Artificial Intelligence* (in press). Available at http://www.dpi.inpe.br/geopro/papers/semantic_granularity.pdf.

Kramer, J.H., 2000. Digital Mapping Systems for Field Data Collection. http://pubs.usgs.gov/openfile/of00-325/kramer.html.

Lillesand, T.M. and Kiefer, R.W., 1994. Remote sensing and image interpretation, 3rd edition. John Wiley & Sons, 750 p.

Mark, D.M., Smith, B. and Tversky, B., 1999. Ontology and geographic objects: an empirical study of cognitive categorization. In: Spatial Information Theory: A theoretical Basis for GIS, Freska, C., Mark, D.M. (Eds.), Springer-Verlag, Berlin, Lecture Notes in Computer Sciences, pp. 283-298.

Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K. and Rabaute, T., 1993. The displacement field of the Landers earthquake mapped by radar interferometry. *Nature,* 364, 138-142.

Richards, S.M., 1999. Geologic concept modeling, with examples for lithology and some other basic geoscience features. Proc. Workshop Digital Mapping Techniques '99, U.S. Geological Survey Open-File Report 99-386, http://pubs.usgs.gov/openfile/of99-386/richard.html (accessed 7 March 2002).

Schetselaar, E.M., 1995. Computerized field-data capture and GIS analysis for the generation of cross-sections in 3-D perspective views. *Computer and Geosciences*, 21(5), 687-701.

Schetselaar, E.M., Chung, C.F, Kim, K.E., 2000. Integration of Landsat TM, gamma-ray, magnetic, and field data to discriminate lithological units in vegetated granite-gneiss terrain. *Remote Sensing of the Environment*, 71(1), 89-105.

Schetselaar, E.M. and E. A. de Kemp, 2001. 3-D symbolization of Planar and Linear Fabric Elements using Surfboards as an Aid to the Analysis of Geological Structures. Proc. Congress of the International Association of Mathematical Geology, Cancun Mexico, Session F (on CD-ROM). Available at http://www.kgs.ukans.edu/Conferences/IAMG/Sessions/F/Papers/dekemp2.pdf

Searle, J., 1995. The Construction of Social Reality. The Free Press, New York.

Soller, D. R. and Berg, T. M., 2000. The National Geologic Map Database: A Progress Report. Proc. Workshop Digital Mapping Techniques '00, U.S. Geological Survey Open-File Report 00-325. http://pubs.usgs.gov/openfile/of00-325/soller2.html

Woldai, T., 1995. Survey and mapping in Sub-Saharan Africa – Which way further? Keynote paper, UN-Outer Space Affairs Division (OSAD), United Nations Regional Conference, Senegal, West Africa.

Woldai, T. and Limptlaw, D., 2000. Land use change detections as an initial stage in environmental impact assessment on the Zambian Copperbelt. Proc. 28th International Symposium on Remote Sensing of Environment and 3rd AARSE Conference "Information For Sustainable Development", 27-31 March, Cape Town, South Africa, Category 6, pp. 23-26.