Progress Towards a Consistent Exchange Mechanism for Geodetic Data in Australia and New Zealand

Roger FRASER, Australia and Nic DONNELLY, New Zealand

Key words: eGeodesy, Geodetic Data Exchange, Datum Maintenance, Standards, ICSM

SUMMARY

The core function of the Australian and New Zealand Intergovernmental Committee on Surveying and Mapping (ICSM) is to coordinate and promote the development and maintenance of key national spatial information. Integral to the maintenance of national geodetic infrastructure is the sharing and collation of vast amounts of geodetic data and metadata amongst a number of custodians, proprietary databases and geodetic related software. To this end, the Geodesy Technical Sub-Committee (GTSC) of ICSM has established the eGeodesy project to develop a standard for the exchange, delivery and archiving of geodetic information among Australia and New Zealand's State and Federal agencies. This paper reviews the objectives of the eGeodesy project, and the work undertaken to produce a comprehensive data model for managing geodetic data and metadata.

The outcomes of eGeodesy are critical to maximising the value gained from the extensive amounts of geodetic data held by the various ICSM jurisdictions and the wider geospatial industry. Managing geodetic data in a standardised way will further assist the maintenance of next generation geodetic datums for Australia and New Zealand. This is particularly the case where seamless access, efficient use and platform-independent exchange of epochal measurements captured from various sources will be key requirements.

Progress Towards a Consistent Exchange Mechanism for Geodetic Data in Australia and New Zealand

Roger FRASER, Australia and Nic DONNELLY, New Zealand

1. INTRODUCTION

One of the tasks involved in the development and maintenance of geodetic infrastructure is the management of extremely large amounts of data. Information that is typically recorded can include, for example, several decades of raw and reduced measurements, measurement processing parameters and results, station and physical mark information, published and historical coordinates, and – with the rise in the number of Global Navigation Satellite Systems (GNSS) continuously operating reference stations (CORS) – detailed information about station equipment and site-specific observation parameters. In addition, various forms of metadata are associated with this information to assist in making decisions about its correct use.

Amongst the geodetic and GNSS community, the management and sharing of this data has been no less than problematic due to the fact that there is no single, standardised way for recording geodetic information. By and large, data management has been software-driven. That is, the way in which data has been stored has generally relied upon the proprietary data file format of the hardware and software adopted to perform routine geodetic operations. Given the availability of a number of different software packages (including a host of custom-developed applications), the exchange of this data amongst different jurisdictions and organisations has relied heavily upon file format translation and labour intensive data checking. What is more, the task of translation can be rather complex and is not always free from error. Following the view of many in the field of information management, a viable solution is to manage geodetic information according to a single, standards based object oriented data model.

It is widely recognised that a variety of data models exist for the structure and format of geographic and survey related information. Some well known models amongst the spatial information community include LandXML, Geography Markup Language (GML) and Google's KML. Whilst these models in themselves are quite comprehensive, none can alone cater for the complete range of information geodesists routinely deal with. The adoption of Receiver INdependent EXchange (RINEX), Solution (or Software) INdependent EXchange (SINEX), International GNSS Service (IGS) products, the Radio Technical Commission for Maritime Services (RTCM) Differential GNSS standard and Networked Transport of RTCM via Internet Protocol (NTRIP) standards by numerous GNSS hardware and software vendors has led to significant efficiencies in the way in which raw GNSS measurements and station solutions can be recorded, shared and processed. However, as with the survey and geographic data models, the GNSS measurements managed by these standards deal with only a subset of the total range of geodetic information.

The adoption of a single data model for managing geodetic information is clearly a fundamental prerequisite for efficient and effective geodetic data management. This is particularly the case where seamless access, efficient use and exchange of epochal geodetic information captured from various sources are key requirements for collaborative geodetic infrastructure management. With the growing acceptance of distributed computing and data sharing via Web service technology, the development of such a model is not only desirable but essential. In this context, the purpose of this paper is to review the work being undertaken by the Geodesy Technical Sub-Committee (GTSC) of the Australian and New Zealand Intergovernmental Committee on Surveying and Mapping (ICSM) to develop a comprehensive data model for managing geodetic data and metadata.

2. GEODETIC DATA MANAGEMENT IN AUSTRALIA AND NEW ZEALAND

2.1 Geodetic Data Management in Australia

Generally speaking, there is no standardised approach for managing geodetic data across Australia. Whilst the procedures for measuring and processing geodetic information in Australia are reasonably consistent and are based upon well established standards and practices, there is no formalised standard for geodetic information management. Matters concerning data format, interoperability between systems, persistent storage and database management are usually left to each jurisdiction to manage. Where information has been stored in files rather than in databases, the file format adopted for recording station information, measurements, metadata about stations and marks and the like is typically dependent upon the proprietary file format of the hardware and software adopted to perform each geodetic task.

For the most part, the geodetic data managed by Australian jurisdictions has been collected during projects and campaigns, or is being collected continuously in real time. For the majority of the campaign-based data, this information has been captured to establish, maintain and enhance Australia's geodetic horizontal and vertical datums and supporting infrastructure. Keeping in mind that Australia's geodetic datum development has involved over fifty decades of measurement, processing and adjustment, Australia's core geodetic data is characterised by a high level of variability in quality, consistency and completeness – due in large part to the many instruments, working environments, standards, organisations and people involved in the capture, manipulation and computation process.

With the advent of real-time CORS streams, continuous GNSS data is being archived at an hourly rate. Due to the variation in equipment, network size, epoch rate and use of precise satellite clock and orbit products, raw GNSS data files vary dramatically in format, file size and quantity. Typically, most raw GNSS data is stored in RINEX or GNSS receiver manufacturer proprietary formats. Where a CORS network solution is being used to provide precise, real time corrections, additional information will be created and stored according to the needs of the software and its users. This information may exist in proprietary or open standard format. Notwithstanding the availability of standards such as RINEX, SINEX, IGS product specifications, RTCM and NTRIP, for example, there is still no standardised model in

Australia for exchanging geodetic information among the different packages employed to manage CORS network services and products.

As can be appreciated, there is an almost overwhelming amount of digital data being managed within Australia for the purpose of maintaining the numerous ground marks and CORS sites. Whilst users of digital geodetic data in Australia are becoming accustomed to working with large files and translation between various software packages, there is a growing need for geodetic data to be presented in a standardised format for the purpose of consistent, instantaneous exchange between internal software applications and databases and external data custodians and users.

A final point worth noting is that there is some discrepancy in the use and interpretation of various terms relating to geodetic infrastructure in Australia. Whilst there is a high level of correlation between the types of information each jurisdiction manages, the vocabulary for naming and describing geodetic entities is not completely consistent. For instance, information about geodetic stations maintained by each jurisdiction will usually include a reference to a physical mark. The vocabulary for describing a physical mark's type, material, condition, recovery, location, access and so on is known to vary from one State to the next. Even a superficial reading of condition indicators such as good, appropriate and OK brings a level of doubt as to how consistency in interpretation may be achieved from one mark to the next.

2.2 Geodetic Data Management in New Zealand

Land Information New Zealand (LINZ) is the government department mandated to develop and manage New Zealand's geodetic system and the associated data. Most geodetic data is stored within the Landonline database, which also manages the nation's cadastral, electoral and land titles data. A copy of the key Landonline geodetic data is also contained within a Microsoft SQLServer database called the Geodetic Database (for example, coordinates and mark location information). The Geodetic Database also contains supplementary data, which has been collected since the design of the geodetic component of Landonline in the late 1990s (such as cell-phone coverage details). Information in the Geodetic Database is made available to the public via the LINZ website (www.linz.govt.nz/gdb).

Geodetic survey and physical maintenance work in New Zealand is carried out almost exclusively by private survey firms under contract to LINZ. Geodetic data collected by contractors is submitted to LINZ as RINEX files (for GNSS data), SNAP files (for adjustment results using the LINZ-developed SNAP software) and csv files (for mark and observation data). Once the quality of the work has been checked, it is converted to an in-house format (Core Records System format – CRS) and loaded into Landonline and the Geodetic Database. With the exception of the RINEX data, the data exchange formats used do not follow any international standard.

3. DESIGN GOALS

It is clear from the preceding overview that there is a need to manage geodetic information in a consistent, structured, homogeneous, and platform independent way. The questions that naturally follow are – what attributes are necessary to describe the characteristics of geodetic data and processes? How should geodetic information be structured? What standards are available? Are there any existing standards based data models that can be built upon? How can information stored according to such a model be seamlessly incorporated within existing business processes? What is the best way to make this information accessible to end users and collaborative maintenance partners?

The answer to the first question is presented in detail in section 4. The answer to the remaining questions requires an understanding of the available technologies and recommended standards for modelling data and metadata. Leaning upon the justification presented by Picco et al. (2006) to adopt the Unified Modelling Language (UML) and XML for documenting the model and encoding data, this section will review the project aims of the eGeodesy project, the design goals of the current model and some of the benefits and practical implications of its implementation.

3.1 Project Objectives

In an attempt to develop a comprehensive data model for managing geodetic data and metadata, the eGeodesy project commenced with the following objectives:

1. Standardise the exchange of geodetic information
2. Standardise the publishing of results
3. Standardise the recording of observations
4. Enable the archival of observational data
5. Streamline the process for generating adjustment files
6. Seamlessly interface with and between vendor packages

To date, all except objective 6 have been achieved through the process of modelling business practices, roles and fundamental data types. The development of the model with respect to objectives 1 – 5 is discussed in section 4.

3.2 An Object–Oriented and Open Standards–Based Approach

Critical to the seamless, platform-independent exchange of information between various users and applications is the adoption of open standards. But the task of geodetic data management and exchange is not a simple one due to the sheer amount of data being managed and the various file formats and standards that exist. An open standard which can adequately deal with this problem is the eXtensible Markup Language (XML). XML is a W3C (World Wide Web Consortium) metamarkup language for describing data. Data is described by surrounding it with context sensitive, text-based tags that give information about the data itself as well as its hierarchical structure. The benefits of using XML are varied and well known, the most notable of which is that it enables users to seamlessly exchange information between organisations, users and a wide variety of applications without the need for complicated translation. A second benefit of no lesser significance is that XML is directly compatible with Web service technology.

Web services provide a standardised way of exchanging information and exposing application logic. Web services are designed to enable different software applications running on various platforms, distributed on a network or over the Internet, to interact in a completely seamless way. A Web service is a piece of application logic residing at a specific network location (or network address) that is accessible to programs via standards-based Web technologies including Hypertext Transfer Protocol (HTTP), Simple Object Access Protocol (SOAP), Web Service Definition Language (WSDL) and XML. During the process, a Web service may call other Web services and/or run other software applications. The response may be in the form of an entire dataset such as a map, a database entry, or simply the result of a computation.

The benefits gained by adopting open standards such as XML and the Web services architecture for the dissemination of geodetic information and geodetic services cannot be underestimated. Some examples of how this approach can be used are given in section 5.

4. MODELLING GEODETIC DATA

At the time of writing, a draft geodetic data model has been developed. For this purpose, UML was used to document fundamental processes and geodetic data types used in Australia and New Zealand. Class diagrams and enumerators have been developed to manage information pertaining to datums and coordinate systems, physical station and GNSS CORS information, mark monuments, measurements, raw measurement files, adjustments, quality and uncertainty, projects and campaigns, and roles. In order to provide consistency in the use of geodetic vocabulary and meaning, effort has been given to the standardisation of terms for fundamental geodetic data types and behavioural processes. Where possible, terminology has been based upon previously established standards or the most widely accepted value. To assist users in the use and interpretation of each entity, detailed descriptions have been provided.

The task to archive geodetic information at the database level using this model will inevitably require some form of custom mapping between the model and the user's database schema (or physical model). Given the variation in database models in Australia and New Zealand, the development of such a map is assumed to be the responsibility of the custodian and is therefore not considered here.

4.1 Datum

Information about coordinate systems and the transformations between them provides the basis for all other parts of the geodetic model. A datum is the fundamental reference frame for geodetic work. It has an ellipsoid associated with it, as well as metadata such as the reference epoch.

A given datum may have many different types of associated coordinates. For example, New Zealand Geodetic Datum 2000 (NZGD2000) coordinates could be represented as geodetic coordinates (lat, long, height) or in terms of a projection (east, north).

To resolve the many-to-many relationships among datums and coordinate types, a coordinate system is used. For example, coordinates may be expressed as latitude and longitude (a coordinate type) on NZGD2000 (a datum). The coordinate system which represents this is also called NZGD2000, but we know that it specifically expresses latitude and longitude coordinates.

There are numerous transformation methods that are used in geodetic computations. Associated with a transformation method will be a number of transformation parameters. For example, the Transverse Mercator transformation method includes parameters such as the central meridian and false origin.

A specific instance of a transformation method is represented through transformation functions and transformation parameter values. Continuing the above example, the Transverse Mercator function might output coordinates in terms of the New Zealand Transverse Mercator (NZTM) projection, for which the function needs the specific NZTM parameter values.

Ongoing work is being carried out to ensure that, as far as possible, datum-related modelling is consistent with that which is outlined in the standard AS/NZS ISO 19111: 2008 Geographic information – Spatial referencing by coordinates.

4.2 Measurements

Geodetic measurements can exist in a variety of forms. A single measurement can be associated with one, two or three nodes and can be expressed in one of a number of units and optionally with respect to a datum. Each measurement has an associated set of quality indicators relating to method, measurement characteristics and precision.

The model provides the capability to record metadata about each measurement, including date, ID, reduction parameters, associated project/campaign, and observer. Parameters affecting the adjustment of each measurement, such as include/exclude and apriori scalars are also provided. Clusters of measurements are handled by an unbounded list of single measurements. Correlations between single measurements are maintained within the single measurement class.

The model caters for the following measurement types – ellipsoid chord distance, horizontal angle, geodetic azimuth, ellipsoid arc distance, GNSS baseline, GNSS point, orthometric height, orthometric height difference, astronomic latitude, astronomic longitude, astronomic azimuth, mean sea level (MSL) arc, geodetic latitude, geodetic longitude, ellipsoidal height, slope distance, zenith angle, vertical angle, absolute gravity and gravity difference. In order to enable the reproduction of processed measurements, the ability to record information pertaining to raw measurement files, instruments, instrument settings, measurement conditions and processing parameters has been provided.

4.3 Geodetic Station and CORS Information

Geodetic station information relates primarily to nodes and coordinates. A node is an entity which represents a three-dimensional position in space. To the lay user, a node often appears identical to a mark, but a mark is what physically represents the node. For example, a node may be physically represented by a stainless steel pin (the mark). The distinction between marks and nodes is useful. For example, an iron tube might rust and be replaced by a stainless steel pin in an identical three-dimensional position. In this case the mark has changed, but the node remains the same. Conversely, the physical location of a mark may change (perhaps due to nearby construction activities). In this situation the mark is the same, but is located at a new node. Since geodetic coordinates are typically three-dimensional, a change in the height of a mark is enough to require that it be associated with a new node.

Each node may have many coordinates associated with it. These coordinates may be in terms of one or more coordinate systems. For example, a node may have several coordinates in terms of the official datum as recalculations are carried out to improve coordinate quality. The best quality (usually most recent) coordinate would be the authoritative coordinate. Others would have a status of decommissioned. There may also be coordinates associated with historical datums, or vertical datums. A key piece of logic is that there should only be one authoritative coordinate for each coordinate system with which the node is associated. Often coordinates also have a velocity associated with them. Since velocity may change with time, start and end dates for each velocity value associated with a coordinate can be recorded.

For CORS, Scripps Orbit and Permanent Array Center (SOPAC) has developed an XML schema for managing GNSS site logs structured according to the International GNSS Service (IGS) site log format. The schema includes provision for details of equipment, monuments, sites and meteorological data. The geodetic model will take advantage of the comprehensive nature of the SOPAC model, supplementing it where required to meet any additional requirements of the ICSM jurisdictions. One of the advantages of using XML is that it is easy to reference other XML schemas.

4.4 Physical Mark and Monument Information

Every measurement observed over a geodetic network is typically based upon an instrument located with respect to a physical mark or monument, whether it be a CORS GNSS antenna mount, a brass plaque or otherwise. All information relating to physical monuments and their association with nodes is managed within a mark metadata package.

For all types of marks, a permanent mark class has been defined for handling mark ID, multiple mark names, mark description, installation date, installation details, mark status and type, physical state, condition, site location and access information. Supplementary documents, such as reports, photographs, site access diagrams, mark diagrams (or sketch plans), network diagrams, GNSS site logs, field notes, station observation summaries, and monument construction plans are managed within a supplementary document class. Provision is also made for managing mark metadata recorded by multiple visits to the site over time and any maintenance activities associated with the mark. The condition or status of a mark, mark type, locality and mark name are all based upon enumeration lists.

Within any geodetic network, nodes of particular significance may have a protection structure or fence, a beacon and/or recovery (or eccentric or witness) marks in the vicinity. Information pertaining to such objects, including condition, associated maintenance activities and any supplementary documents, is represented by a recovery mark class. With all marks, a reference to a mark quality standard is provided so as to standardise on the terms used to describe the quality of a physical mark installation.

4.5 Adjustment and Acceptance of Results

The data associated with a geodetic adjustment includes input data (coordinates and measurements), constraints (on coordinates and measurements), details of the adjustment (such as methodology) and output data (coordinates, residuals and statistics).

Input coordinates and measurements are brought through from the Station and Measurement packages respectively. There is provision for the inclusion of both stochastic and non-stochastic constraints. For example, adjustments will have stochastic constraints associated with the observations (through a priori error estimates on measurements) and may have non-stochastic constraints associated with the coordinates (such as holding some coordinates fixed).

Metadata relating to each adjustment is provided for in the model. This includes the adjustment runtime, coordinate system and user, as well as key summary statistics such as the number of iterations required for convergence. There is also information associated with the adjustment method, such as the number of dimensions (1, 2 or 3D) and confidence limits used in generating statistics.

Adjustment outputs include adjusted coordinates and measurements, as well as accuracy information. Both coordinate changes and observation changes (residuals) are recorded.

4.6 Projects and Roles

Geodetic activities are typically carried out on a project basis. A project may include a number of activities, from campaign planning to computing final coordinates. Information about workflow can be stored, for example, when each stage of the project is completed.

The project class is essential for ensuring that data stored can be traced back to its source if problems arise. For example, if there are concerns about the quality of coordinates, these can be traced back to the project which created them, which would then identify field notes or site logs which could be examined for errors.

Within a project there will be a number of roles. A role may be carried out by an individual or corporate entity (for example, a survey firm contracted to do fieldwork). A person or corporation may have one or more roles within the project. There are numerous potential roles within a project, including project manager, observer, adjuster and quality assessor.

4.7 Quality and Standards

Information pertaining to quality and uncertainty in all aspects of geodetic datum maintenance is essential due to the fact that error is an inescapable component of measurement. It follows that quality is a key indicator for determining the extent to which a dataset may be used. To this end, elements for managing qualitative and quantitative information have been integrated within the model using published standards or widely accepted conventions. The terms adopted for quantitative and qualitative information are based upon ICSM's Standards and Practices for Control Surveys, the ISO 9000 family of standards, and the core quality metadata elements developed under the Australia and New Zealand Land Information Council (ANZLIC) metadata initiative.

Provision has been made to hold precision for coordinates and measurements in the form of an upper triangular variance co-variance (VCV) matrix. The number of elements within a single node or measurement (such as 1D, 2D, 3D points and vectors) will determine the dimensions of the matrix. Depending on the number of nodes or measurements within a cluster, off-diagonal elements from a fully populated cluster VCV matrix may be associated with each node and measurement. The unit of measure for each diagonal term is also managed so as to allow rigorous propagation of uncertainty between coordinate systems.

5. USE CASES

The following sections demonstrate the viability of adopting a standards-based data model for managing geodetic information, and open standards Web technologies for the exchange of this information between organisations, users and a wide variety of applications without the need for complicated file format translation.

5.1 Collation of data from different custodians using Web services

A common task in carrying out very large network adjustments is the collation of information relating to stations and measurements. Invariably, some of this information is managed by an agency external to the user's organisation and software application. Accordingly, the current practice involves the transfer of files from the central repository (or point-of-truth database) via CD or email and, if necessary, conversion into the required file format of the software application being used. Figure 1 demonstrates the concept of collating data from various custodians of geodetic data into a single dataset using the Web services architecture. At the core of this Web service delivery use case is "Geodetic XML" – a consistent, standards-based data model for managing geodetic data.

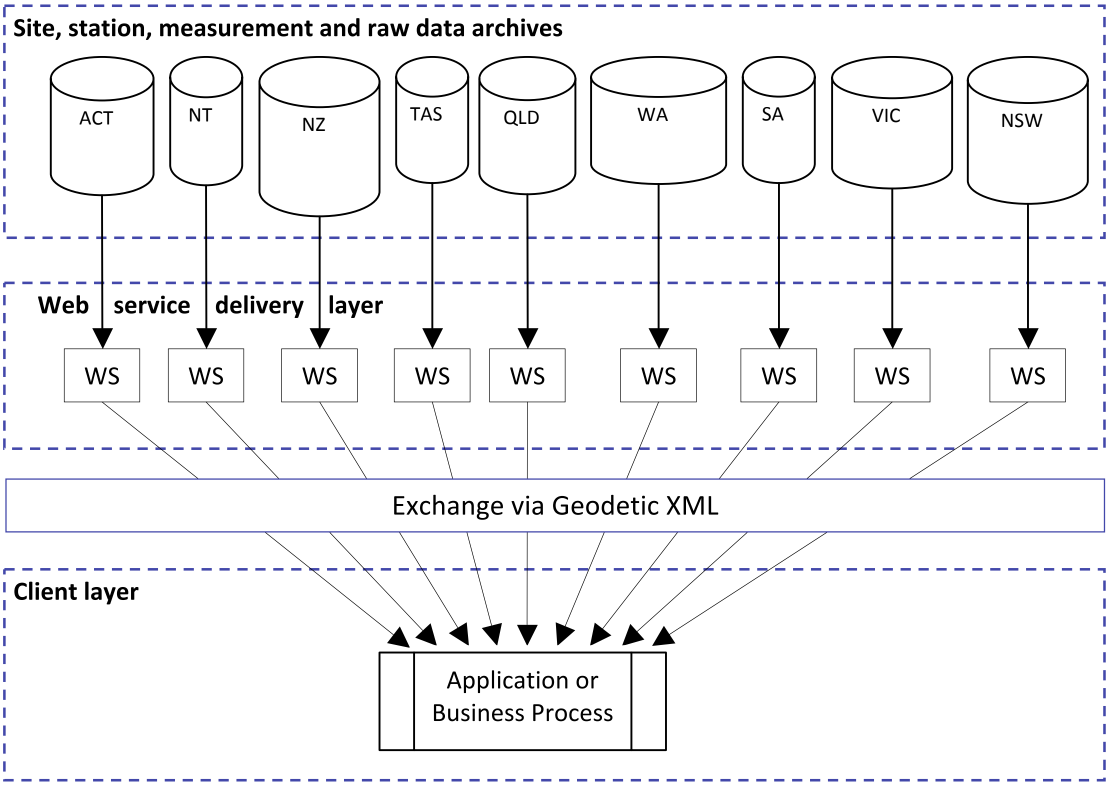

Figure 1 Collation of geodetic data via Web services

5.2 Online submission and validation of geodetic data

The process and formats employed by LINZ for exchanging geodetic data with contractors are currently rather cumbersome. Data is delivered via CD or DVD as a series of csv and other files. LINZ then converts these to the CRS format for loading into Landonline and the Geodetic Database. Due to the complexity of the csv file formats, inability of the csv format to inherently validate contained data and the lack of any off-the-shelf software packages which output them, there are frequently problems exchanging data due to errors in the csv files.

Under the proposed process using a Geodetic XML format for data exchange, data loading is simplified through the provision of a Web service via which contractors submit data. The Web service would include a number of automated validation checks to ensure that most errors are picked up prior to the data being formally submitted to LINZ. In the medium to long term, it is anticipated that geodetic software packages would output data directly in the Geodetic XML format, further increasing the ease of data transfer between LINZ and its contractors. Figure 2 shows the current and proposed processes.

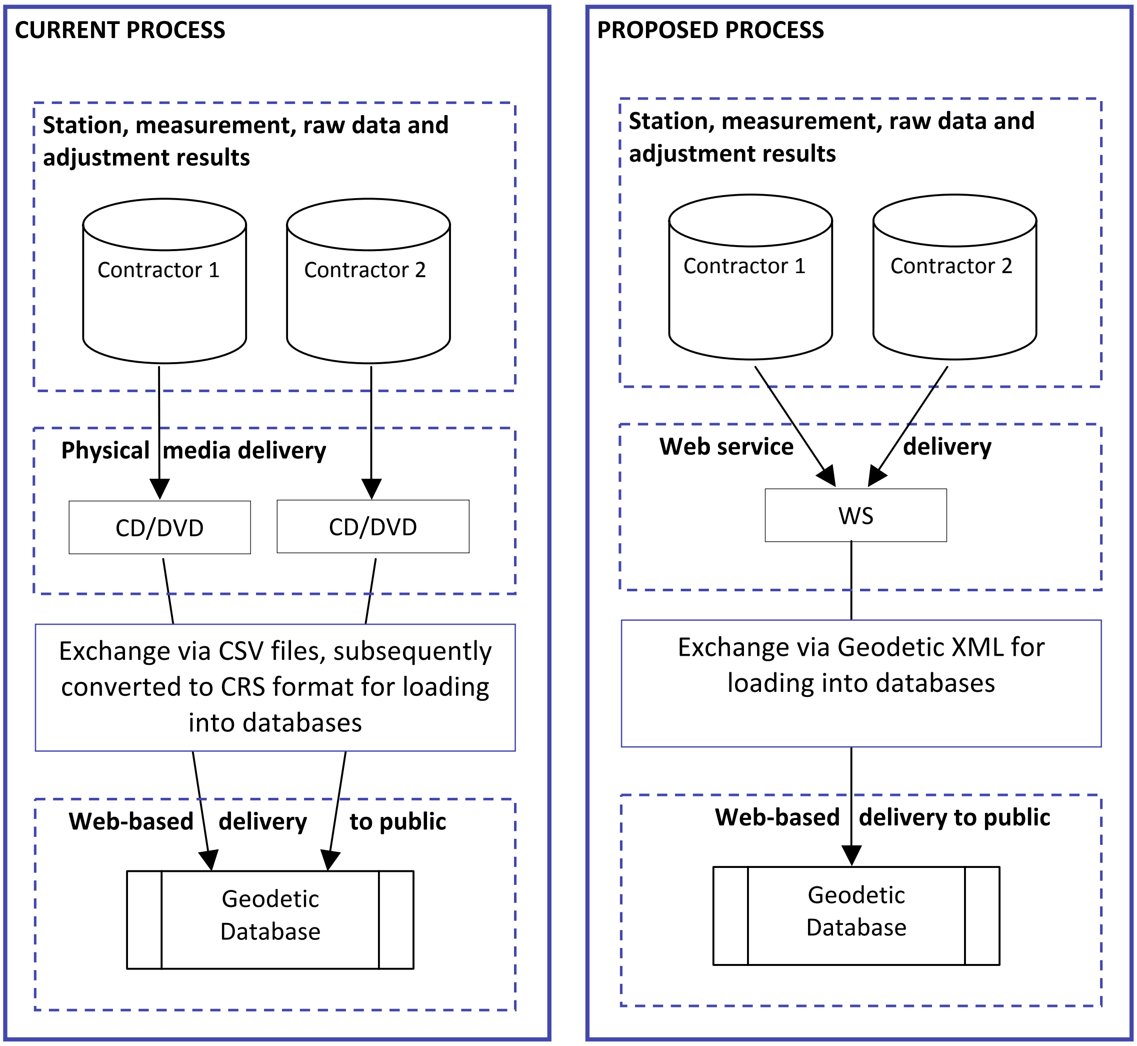

Figure 2 Online submission and validation of geodetic data

6. CONCLUSION

At the time of writing, there is no standardised approach for managing geodetic data in Australia and New Zealand. As such, the current procedures for exchanging geodetic data between software applications, databases and users involves complicated, labour intensive file translation. This paper has reviewed the objectives of the ICSM GTSC eGeodesy project and the work undertaken to produce a comprehensive data model for managing geodetic data and metadata. It presents an open standards based, platform independent way of seamlessly exchanging geodetic information between organisations, users and a wide variety of applications. It is proposed that a first draft of the model will be published on ICSM's GTSC website in the near future for feedback and collaborative development. Stakeholders interested in becoming involved in the development of the model are encouraged to contact the authors of this paper.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the assistance of Jeremy Palmer, who helped develop the geodetic UML model, and ICSM GTSC for financial and technical support.

REFERENCES

ANZLIC (2001). Metadata Guidelines: core metadata elements for geographic information in Australia and New Zealand. Version 2.0. Australia and New Zealand Land Information Council, Canberra, Australia. Available online: http://www.anzlic.org.au/get/2358011755.pdf

AS/NZS ISO 19111:2008, 2008, Geographic Information – Spatial Referencing by Coordinates, Standards Australia (Sydney) / Standards New Zealand (Wellington).

ICSM, 2007, Standards and Practices for Control Surveys (SP1). Version 1.7. Intergovernmental Committee on Surveying and Mapping, Canberra, Australia. Available online: http://www.icsm.gov.au/icsm/publications/sp1/sp1v1-7.pdf

ICSM, 2006. Geocentric Datum of Australia Technical Manual. Version 2.3. Intergovernmental Committee on Surveying and Mapping, Canberra, Australia. Available online: http://www.icsm.gov.au/gda/gdatm/gdav2.3.pdf

SOPAC, 2004. XML for Geodesy. Available online: http://sopac.ucsd.edu/projects/xml/

BIOGRAPHICAL NOTES

Roger Fraser, PhD Melb, LS. Since the completion of his undergraduate degrees in 1999, Roger has been involved in various cadastral projects, the maintenance of State geodetic networks and the development of geodetic and GNSS related software. Currently, Roger is Manager of the Geodetic Survey section of the Office of Surveyor-General Victoria within the Department of Sustainability and Environment. His main interests lie in the fields of geodesy, GNSS and in particular, the development of algorithms and software for the adjustment of large geodetic networks.

Nic Donnelly BSurv (Hons), BSc (Otago University). Nic is a Geodetic Surveyor within the Specialist Processing team at Land Information New Zealand (LINZ). His interests and responsibilities include geodetic adjustments and maintenance and enhancement of the geodetic cadastre.

CONTACTS

Dr Roger Fraser
Office of Surveyor General Victoria
Department of Sustainability and Environment
L17, 570 Bourke Street
Melbourne 3000
AUSTRALIA
Tel. +61 3 8636 2551
Fax +61 3 8636 2776
Email: roger.fraser@dse.vic.gov.au

Mr Nic Donnelly
Specialist Processing
Land Information New Zealand
Lambton House, 160 Lambton Quay
Wellington
NEW ZEALAND
Tel. +64 4 460 0191
Fax +64 4 498 3837
Email: ndonnelly@linz.govt.nz