ENABLING GEOTECHNICAL DATA FOR BROADER USE BY THE SPATIAL DATA INFRASTRUCTURES

by

Amir Ghasem Zand

A Dissertation Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
DOCTOR OF PHILOSOPHY
(CIVIL ENGINEERING)

August 2011

Acknowledgments

The research presented in this dissertation was carried out under the supervision of Professor Jean-Pierre Bardet in the Sonny Astani Department of Civil and Environmental Engineering at the University of Southern California.

I wish to thanks Professor Bardet for his advice, guidance and support throughout my study at USC and the preparation of this dissertation.

I would also like to thank the following: members of my Ph.D. committee, Professor Geoffrey R. Martin, Professor L. Carter Wellford and Professor Petros Ioannou for their guidance in improving the manuscript; the staff of the Civil and Environmental Engineering Department for their support; the graduate students of the Civil and Environmental Engineering and Computer Science Departments at USC.

Table of Contents

List of Tables

List of Figures

Abstract

Geotechnical data are one of the most prevalent data types in civil engineering projects. The majority of the civil engineering projects that are in use today are designed using site-specific geotechnical data.

The usage of geotechnical data is not limited to construction projects. This data is used in a wide range of applications, including seismic hazard analysis, planning and zoning studies, risk analysis and other infrastructure development projects. Demand for geotechnical data in these types of applications has increased in the past few decades, due to proliferation of geographic information systems (GIS) and a variety of applications that take advantage of GIS and spatial data.

Considering the widespread collection and usage of geotechnical data in various disciplines, one might expect that data are readily available for most developed areas. However, unlike other types of spatial data that are available in spatial data infrastructures (SDI), geotechnical data is often managed using traditional and ineffective methods. Consequently, for a lot of projects it is difficult to find and acquire these data. This issue is frequently encountered in civil engineering projects, and more importantly, in large-scale multi-disciplinary studies that need large volumes of geotechnical data.

In order to address this problem, the current methods used for management, archiving and distribution of geotechnical data need to be improved upon. The most viable solution is to leverage the existing information technology infrastructure and adopt methods that are already in use for other types of spatial data. These technologies include geography markup language (GML), spatial databases and Web services developed for spatial data exchange.

Following this concept, in the subject dissertation development of a spatial data model for geotechnical data is discussed. The discussion includes an overview of the geotechnical data collection, processing and current methods used to archive and exchange data. The proprietary software and data formats that are used for geotechnical data exchange, including the association of geotechnical and geoenvironmental specialists (AGS) data format, are covered in this review. In addition, the current state of information technology for other types of spatial data is evaluated. This background study includes spatial databases, spatial data infrastructures and various standards that are adopted by the industry and regulating agencies for management and dissemination of spatial data.

Based on this framework, a data model is proposed for integration of geotechnical data in SDIs. This data model uses the terminology of the AGS geotechnical data exchange format and combines it with a GML-conformant schema. GML is the industry-standard markup language for modeling spatial data for use in SDIs.

The developed data model is compared with similar proposals from other research groups. The functionality of the data group is verified using several examples involving visualizing the geotechnical data and using it for analyses such as site response analysis and liquefaction hazard assessment. A case study is presented that demonstrates the potential benefits of these analysis scenarios in real-world studies.

Finally, the achievements of the dissertation are summarized and suggestions are made in order to improve the results of the current study. Also, some related research topics are suggested to continue and further expand the concepts presented in this dissertation.

Chapter 1. Introduction

One of the key pieces of information required in civil engineering projects and infrastructure studies is geotechnical data. All structures are supported on the ground; therefore, they transfer their weight and other external loads to the soil layers underneath them. They are also affected by different phenomena that occur within the underlying ground, e.g., earthquakes, settlement and groundwater. Due to this interaction between the structure and the soil, the properties of the soil should to be known to design or analyze the structure.

The application of geotechnical data is not limited to structures and civil engineering. Geotechnical data is widely used in other disciplines, including infrastructure, environmental and risk analysis. Geotechnical data are increasingly used in these studies, where more sophisticated and realistic analyses are utilized that rely on accurate input parameters.

One of the factors that limit even more widespread adoption of geotechnical data is relatively antiquated methods that are used to manage and disseminate the data. This issue is identified by practitioners in various design and research projects. The main motivation for the research presented in this dissertation is to improve upon these methods and explore solutions that can be used to facilitate access and usability of geotechnical data.

In the following sections of this chapter, the current geotechnical data management issues are discussed in further detail, and the objectives of this research and the methodology adopted to deliver them are presented. The chapter also includes a brief description of the subsequent chapters of the dissertation.

1.1 Objectives

Modern geotechnical engineering is generally believed to have started in 1925 with the publication of “Erdbaumechanik” by Karl Terzaghi (Goodman, 2002). In addition to establishing the theoretical principles of geotechnical engineering, Terzaghi also pioneered a large number of the investigation, instrumentation and testing methods used in soil investigations for the next few decades. Some of those methods are still in use today. Since then millions of borings have been made for different types of civil engineering projects. In the past few decades in the United States, rarely has any construction project been done without drilling some soil borings. These facts lead to the conclusion that a huge amount of data on subsurface soil layers has been collected in the past decades.

Traditionally, data was prepared and archived as boring logs in hard-copy and microfilm format. Hard copies are an unreliable archiving method, because they are bulky, hard to retrieve and susceptible to loss and deterioration. While microfilms are less bulky, they share other problems associated with hard copies. These shortcomings have resulted in massive amounts of lost, misplaced or inaccessible geotechnical data.

Expansion of personal computers and computer-aided drawing software provided new means to archive boring logs in electronic format. These storage methods have led to vastly improved archives for agencies that own and manage large amounts of geotechnical data.

Despite these advancements, the accessibility and dissemination of geotechnical data has remained mediocre at best. More importantly, the geotechnical engineering discipline has lagged behind other infrastructure disciplines in utilizing state-of-the-art solutions in information technology and geographical information systems. For example, storing of geotechnical data in databases and providing the data through Internet has gained traction in recent years. However, no standardized method has been adopted by the geotechnical engineering community, and the technologies used by the industry are far from cutting edge. This deficiency continues to be an issue in large geotechnical engineering projects; however the bigger impact is on the large-scale studies that involve data from several disciplines, including geotechnical data. These studies generally require larger amount of data and less time for inspecting and conditioning individual pieces of data, common in geotechnical engineering applications.

The main culprit is the relatively small size of the geotechnical engineering community, which limits the commercial market for expensive products and customized solutions in this field. This problem can be partially solved by adopting standardized solutions from other disciplines that deal with similar types of data usage and customizing them for the particular needs of the geotechnical data.

The objective of this research is to investigate and apply current information-technology solutions in order to improve the accessibility and usage of geotechnical data. This research involves adopting current technologies, which are in use in various disciplines, and customizing them for geotechnical engineering applications. The main driving force behind this research is not only the demand from the geotechnical engineering community but demand from other research and engineering fields, which can use the geotechnical data in combination with other data types.

1.2 The Benefits of the Proposed Research

Geotechnical data is used in a wide range of engineering and scientific applications. Civil engineers use the data in foundation engineering, geotechnical earthquake engineering and structural design applications. In these traditional applications, geotechnical data are currently processed and used as hard copies, PDF files and, more recently, geotechnical database programs. For a typical small- to medium-size civil engineering project, these methods provide a practical and established way to exchange and use geotechnical data. While these methods might be inefficient for larger projects, the industry has continued to use the geotechnical data in the same basic way with gradual revisions and improvements.

In recent years, however, there has been an emerging market for geotechnical data from large-scale, multi-disciplinary projects, such as infrastructure planning, risk management, fragility study of transportation and water/wastewater networks, mining and scientific applications. These analyses take advantage of the computational power of new computers in combination with spatial data served by powerful spatial databases and Spatial Data Infrastructures (SDIs) (Onsrud, 2007).

Some of these analyses encompass areas as large as a state or country and require data from thousands of boreholes. Geotechnical data are valuable for these studies because they provide first-hand data for more realistic simulations of various phenomena impacted by soil conditions. Using more realistic models improves the accuracy of the predictions from these studies. Therefore, these applications will render the most benefit from the research conducted in this dissertation.

The current state of geotechnical data exchange is limiting for these large-scale studies because the majority of the existing geotechnical data is not readily accessible for these applications. Furthermore, there is no standard procedure to access and query this data.

In this dissertation a framework for improved access and exchange of the geotechnical data is proposed. The proposed framework builds upon the standards recommended by the governing bodies in the field of spatial data management and SDIs. This framework provides a common interface with other SDIs and serves the needs of large-scale, multi-disciplinary applications that depend on this kind of spatial data.

1.3 Proliferation of Spatial Data Infrastructure

Spatial data infrastructure is a collection of spatial data, metadata, standards and tools that enables acquiring, processing, distributing and using spatial data. A more formal definition is given by the United States government as “the technology, policies, standards, human resources, and related activities necessary to acquire process, distribute, use, maintain, and preserve spatial data” (The White House, 2002). Kuhn offered another definition for SDIs: “An SDI is a coordinated series of agreements on technology standards, institutional arrangements, and policies that enable the discovery and use of geospatial information by users and for purposes other than those it was created for” (2005).

Nowadays SDIs are important tools for decision-making, planning and operation for national and local governments, agencies and the private sector. Many nations spend considerable resources in order to create, expand and maintain SDIs. Various industries invest time and money providing data for SDIs (Onsrud, 2007).

Development of SDIs has made the spatial data available to the public and private sectors at unprecedented extents. Availability of spatial data in this manner benefits a wide spectrum of disciplines and industries that use the spatial data in large, multi-disciplinary studies. Using more realistic models based on actual spatial data instead of simplified assumptions improve the quality and accuracy of these studies. A growing number of these large-scale, multi-disciplinary applications has resulted in a huge demand for location-based or spatial data from various sources.

1.4 Geotechnical Data Usage in the Spatial Data Infrastructure

Geotechnical data are a type of spatial data because they comprise the soil properties of a certain location within the earth. The spatial nature of geotechnical data is sometimes downplayed because often they are used in a project-based context. Consequently, in small projects geotechnical data are referenced relative to the local project coordinate system and the global spatial location is not established.

While the traditional applications of geotechnical data in civil engineering continue to exist and evolve, there is an emerging demand from multi-disciplinary users who utilize the geotechnical data in combination with other spatial data for risk analysis, hazard mitigation and other large-scale planning studies. These studies often rely on SDIs to discover, analyze and disseminate the spatial data.

1.5 Examples of Applications Using Geotechnical Data

Due to current growth and proliferation of applications relying on location-based data, there is a multitude of ongoing projects in various disciplines dealing with all kind of spatial data types. Table 1-1 lists a number of relevant applications that use geotechnical data and rely on data typically served by SDIs. Figure 1-1 shows the concept of SDIs providing spatial data for several unrelated applications. In order to enable geotechnical data for these kinds of applications, they need to be presented in a standard format, comprehensible to various disciplines. Figure 1-2 shows the interaction between geotechnical engineering and SDIs and the steps involved in adopting the geotechnical data to spatial data that is usable by SDIs.

In the following sections, these applications are further discussed in order to provide a clearer picture of the environments in which geotechnical data usage can be improved.

1.5.1 HAZUS-MH

HAZUS-MH is a nation-wide methodology for estimating potential losses from multiple natural disasters including earthquakes, hurricanes and floods. HAZUS-MH was developed by the Federal Emergency Management Agency (FEMA) under contract with the National Institute of Building Sciences (NIBS) (FEMA, 2010a).

HAZUS uses a GIS interface to display hazard data and damage-analysis results. Government agencies use HAZUS to develop policies, mitigation plans and response-and-recovery operations for natural disasters. The current version can be used to estimate the damage from earthquakes, hurricane winds and floods.

Figure 1-3 and Figure 1-4 show sample applications of HAZUS-MH for earthquake and flood hazard evaluations.

1.5.2 REDARS

Risk due to Earthquake Damage to Roadway Systems (REDARS) is a public-domain methodology and software developed by MCEER for estimation of traffic flow, travel times and economic loss after earthquake events (MCEER, 2010). REDARS was released for public use in March 2006.

REDARS analysis can be performed for both deterministic and probabilistic earthquake scenarios. A REDARS analysis for an earthquake event involves the following steps:

- Estimating the seismic hazard, including ground shaking, liquefaction and surface fault rupture
- The state of damage to each component of the network (damage type, extent, and location)
- Estimating the repair cost, resulting down time and traffic impact of each damaged component
- Based on this information REDARS forms a series of system-states of the transportation network node-link model at various times after the earthquake
- Using network analysis procedures REDARS estimates the economic loss and impact on travel time for key routes in the network

Figure 1-5 shows a screenshot of a REDARS analysis. In Chapter 1 a case study involving REDARS analysis is provided which demonstrates how geotechnical data can improve the results of REDARS analysis.

1.5.3 Seismic Hazard Analysis

1.5.3.1 Liquefaction Hazard Maps

Interactive liquefaction hazard maps are also GIS-based applications, which are prime users of geotechnical data. As an example, liquefaction hazard maps of the state of California are reviewed in this section.

Figure 1-6 shows a liquefaction hazard map for San Francisco-Bay Area (CGS, 2010). The boreholes used for this map are marked by red dots on the map. More information on these maps can be found in CGS Special Publications 117A and 118 (CGS, 2004 and 2008).

1.5.3.2 Kobe-Jibankun

A case study demonstrating usage of geotechnical data in seismic hazard studies is Kobe-Jibankun Geotechnical Database (Kobe City, 1999), a geotechnical database and GIS system for Kobe City in Japan. Jibankun means “Mr Ground” in Japanese. It is a public-private partnership between the city of Kobe, Japan, and the engineering community. As of 2010 it contains the information of more than 7000 borings. It has a GIS interface that can show the damage data from the 1995 Hyogo-ken Nanbu earthquake and provide soil conditions along any desired section in the coverage area. The system was developed to study the damage from the earthquake. Figure 1-7 shows a boring location plan in the Jibankun GIS interface. Figure 1-8 shows a fence diagram presented by Kobe-Jibankun.

Kobe-Jibankun system uses two database engines: the GDBS engine serves the geotechnical data, and the Geo-Base GIS engine provides the geographical information. Figure 1-9 shows the Kobe-Jibankun system architecture.

1.5.4 InfoTerre

Another example of GIS systems containing geotechnical data is InfoTerre (BRGM, 2010). InfoTerre Geomatics is provided by the French National Geological Services (BRGM) data. The system contains geological maps of 1:1,000,000 to 1:50,000, records of the subsurface data logs and geological maps of natural hazards and groundwater data. InfoTerre uses international interoperability standards published by the Open Geospatial Consortium (OGC). Figure 1-10 shows a screen shot of InfoTerre GIS interface.

1.5.5 SCEC Community Modeling Environment

Southern California Earthquake Center (SCEC) has been working on a Community Modeling Environment (CME) with the primary objective of providing a test bed for performing seismological and geophysical simulations (SCEC, 2010a). Figure 1-11 shows a diagram of system workflow and capabilities.

The proposed system architecture includes an improved Seismic Hazard Analysis (SHA). A typical SHA requires the following models (SCEC, 2010b):

- Unified Structural Representation (USR): a self-consistent, 3D characterization of active faults and material properties (e.g., seismic velocities) needed to describe regional deformation and seismogenic processes.
- Fault System Model (FSM): an evolving representation of the regional stress and deformation fields capable of predicting the rupture of individual fault segments.
- Rupture Dynamics Model (RDM): a dynamical description of the nonlinear stress/displacement interactions across a rupturing fault as a function of space and time.
- Anelastic Wave Model (AWM): a computation of the propagation, interference and attenuation of the seismic waves that travel along complex paths from a fault rupture to a target site.
- Site Response Model (SRM): a dynamic description of ground excitation in the near-surface environment at a target site, which, for strong ground motions, often involves significant nonlinearities.

The SCEC CME will use waveform modeling based on principles of physics for ground motion prediction. It will also use actual ground conditions for the SRM model. Therefore, geotechnical data are needed to develop the SRM model.

1.6 Outline of the Dissertation

This dissertation is organized in six chapters, including this introduction. An overview of the subsequent chapters is provided as follows:

Chapter 2: Geotechnical Data Acquisition, Archiving and Usage

In Chapter 2 the current state of geotechnical data usage in the industry is reviewed. It starts with an overview of the geotechnical engineering profession and community. The discussion continues with the process of data collection, exploration and sampling methods, in-situ tests, followed by laboratory tests. The data presentation and archiving method, data transfer formats and common software used for these operations are discussed.

Chapter 3: Spatial Data

Chapter 3 reviews the definition of spatial data, the past developments and current trends to access and use this data type and the technologies used for these purposes. The concept of SDIs for spatial data is discussed, and governing standards and specifications are reviewed. The synergy between geotechnical data and other data types in SDIs is examined to make a case for integrating the geotechnical data in SDIs.

Chapter 4: GML Data Model for Geotechnical Data

Chapter 4 covers the development of a GML-compatible data model for geotechnical data. The methodology used in development of the GML schema is discussed in this chapter. It also provides the details of the GML schema, including sample schema and instance documents. A comparison with similar efforts is provided, and pros and cons for each solution are discussed.

Chapter 5: Implementation and Case Study

In Chapter 5 the implementation details and components of SDI for geotechnical data, including the spatial database, Web services and client applications are discussed. Several pilot applications developed for validation of the concepts developed during this research are presented. The chapter also includes a case study involving REDARS analysis for a transportation network. Through this case study it is demonstrated how access to geotechnical data through SDI architecture can enhance the results and improve the decision-making process using REDARS in real situations.

Chapter 6: Summary and Conclusion

In Chapter 6 achievements of this dissertation are summarized, major obstacles and problems encountered during research are explained, and potential solutions are proposed. Finally, suggestions are made in order to further expand and enhance these outcomes.

At the end a full bibliography for the dissertation is provided.

1.7 Chapter Summary

Based on the introduction presented in this chapter, the author believes that there are substantial benefits in integrating the geotechnical data in the larger scheme of spatial data infrastructures. The examples presented show that this proposition has benefits for both the traditional geotechnical community and, more significantly, for large-scale planning and infrastructure studies that rely on spatial data.

Table 1-1. A list of selected multi-disciplinary applications that use geotechnical data.

Name	Application	Developed By	Comments	Reference
HAZUS-MH	Earthquake, hurricane and flood hazard for buildings	Federal Emergency Management Agency (FEMA)	Geotechnical data are used in seismic hazard analysis and runoff calculations	FEMA, 2010
REDARS	Earthquake hazard for transportation networks	Multidisciplinary Center for Earthquake Engineering Research (MCEER)	Geotechnical data are used in seismic hazard analysis	MCEER, 2010
California Seismic and Liquefaction Hazard Maps	General seismic shaking and liquefaction hazard in California	California Geological Survey (CGS)	Geotechnical data are used in seismic hazard analysis	CGS, 2010
Kobe-Jibankun	Geotechnical database and seismic hazard for the city of Kobe	Research Center for Urban Safety and Security (RCUSS), University of Kobe	Geotechnical database	Kobe City, 1999
InfoTerre	Geology and groundwater data for France	French National Geological Service (BRGM)	Geological and geotechnical database	BRGM, 2010
SCEC Community Modeling Environment	Model for simulating ground shaking in southern California	Southern California Earthquake Center (SCEC)	Geotechnical data are used in seismic hazard analysis	SCEC, 2010

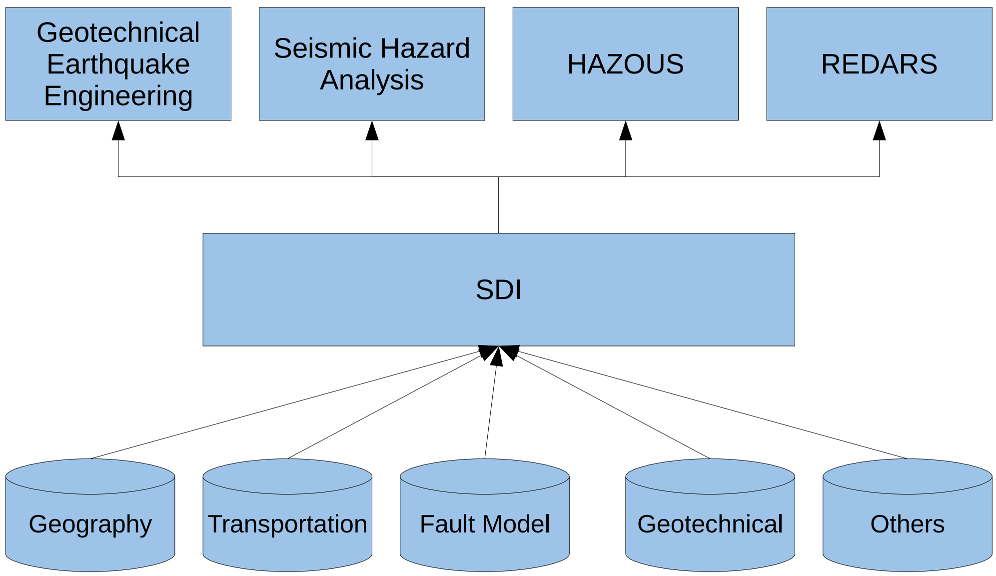


Figure 1-1. A typical SDI providing geographic and transportation network data in combination with geotechnical and earthquake source model data. This SDI can be used for geotechnical earthquake engineering, seismic hazard analysis and transportation network response to earthquake (REDARS) and earthquake risk assessment (HAZOUS).

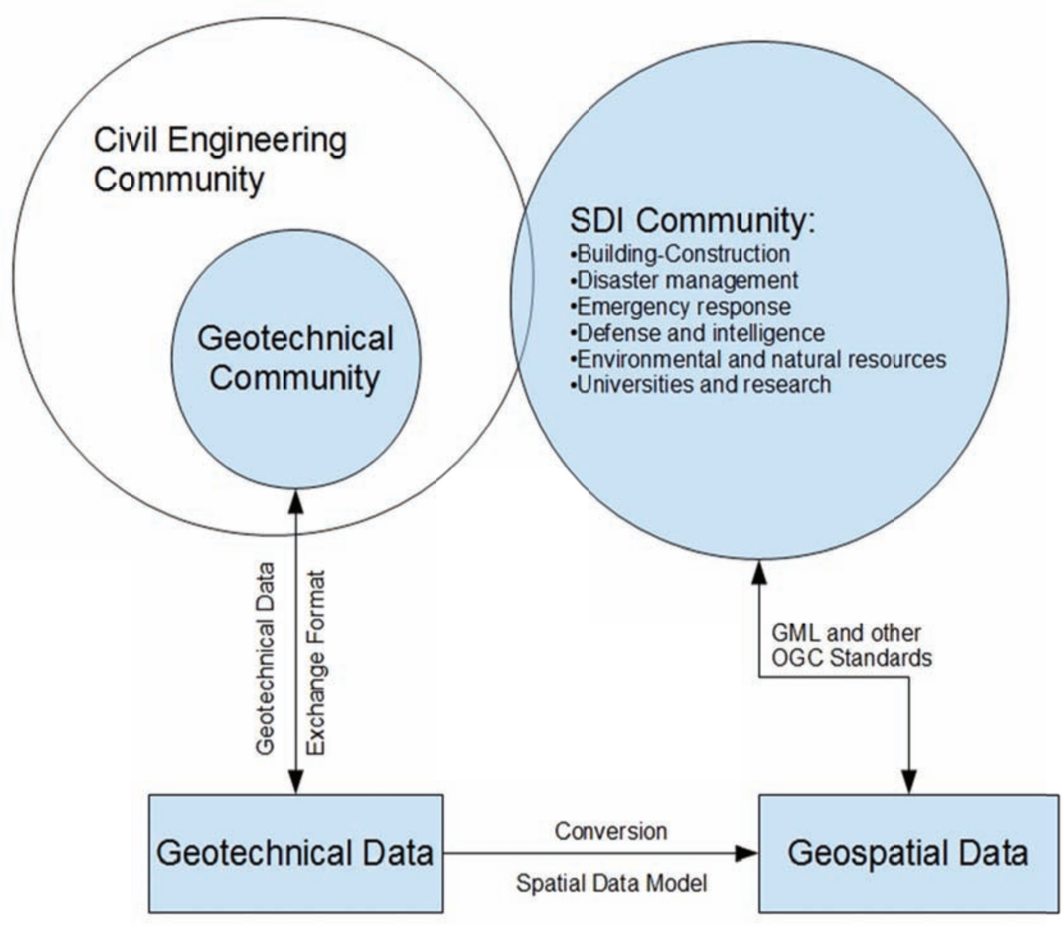


Figure 1-2. This schematic diagram shows the interaction between geotechnical community and SDI community. The geotechnical data needs to be georeferenced for conversion to spatial data. It needs to be presented in a form compatible with SDIs, e.g., GML.

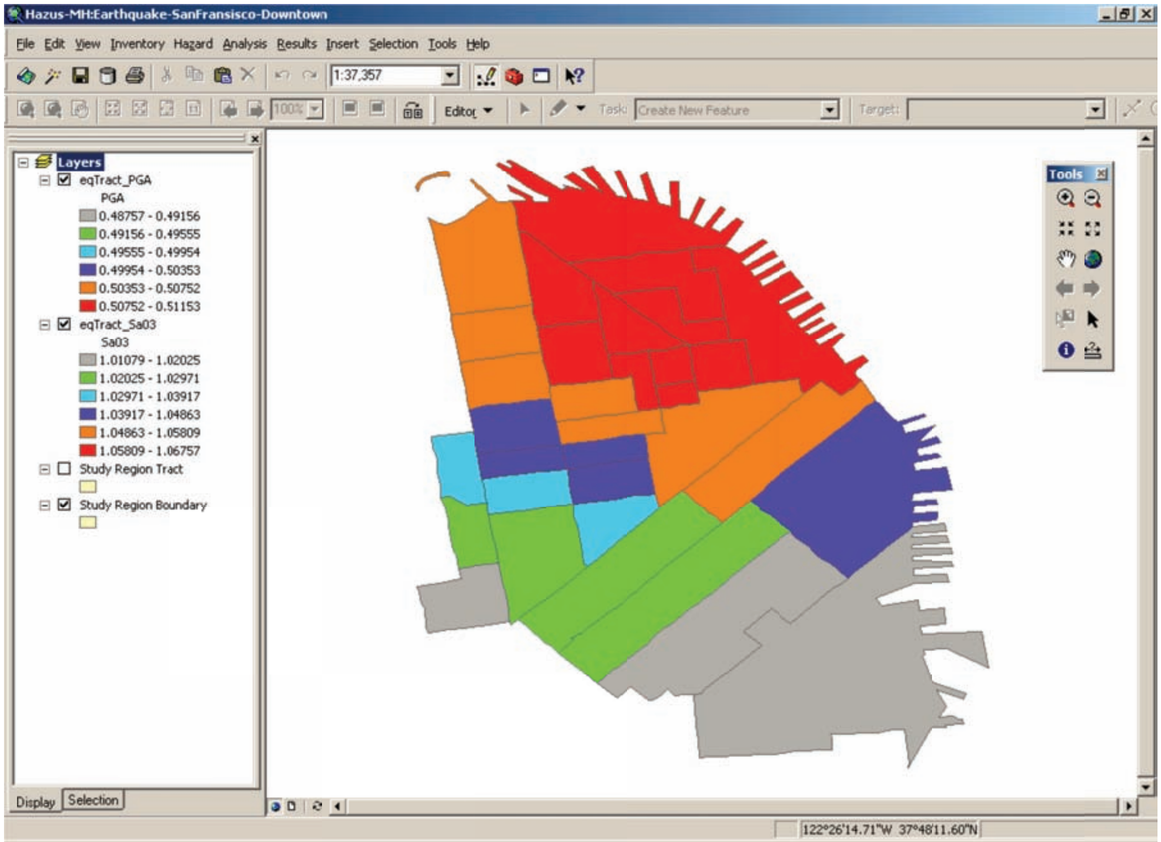


Figure 1-3. A sample HAZUS map, showing peak ground acceleration for census tracts in a study region.

(original figure from FEMA, 2010b)

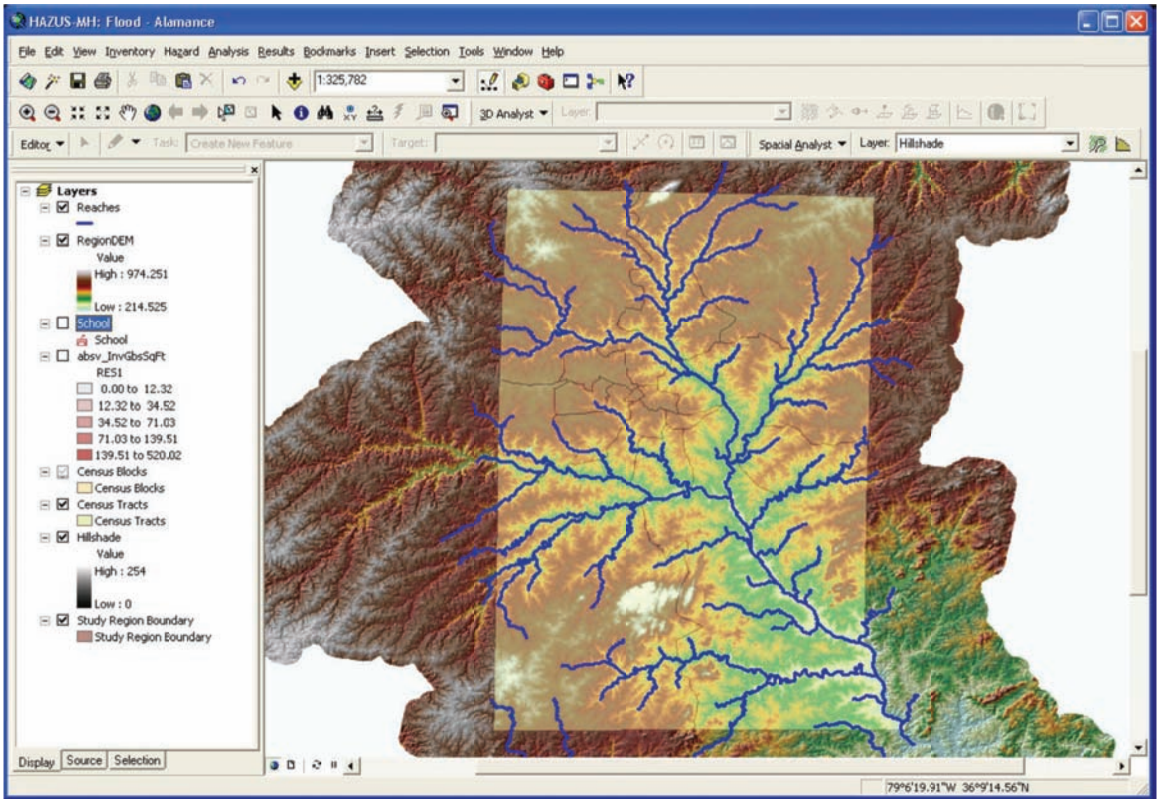


Figure 1-4. A screenshot of ArcMap software, displaying a HAZUS flood model stream network and elevation color shade contours.

(original figure from FEMA, 2010c)

Figure 1-5. Sample REDARS analysis results showing the system state 7 days after earthquake. The state of roads is shown using lines with different colors. The state of damage to components is shown using colored dots.

(original figure from Werner et al., 2006)

Figure 1-6. CGS interactive seismic hazard map showing liquefaction hazard zones and location of the boreholes used to evaluate the hazard.

(original figure from CGS, 2010)

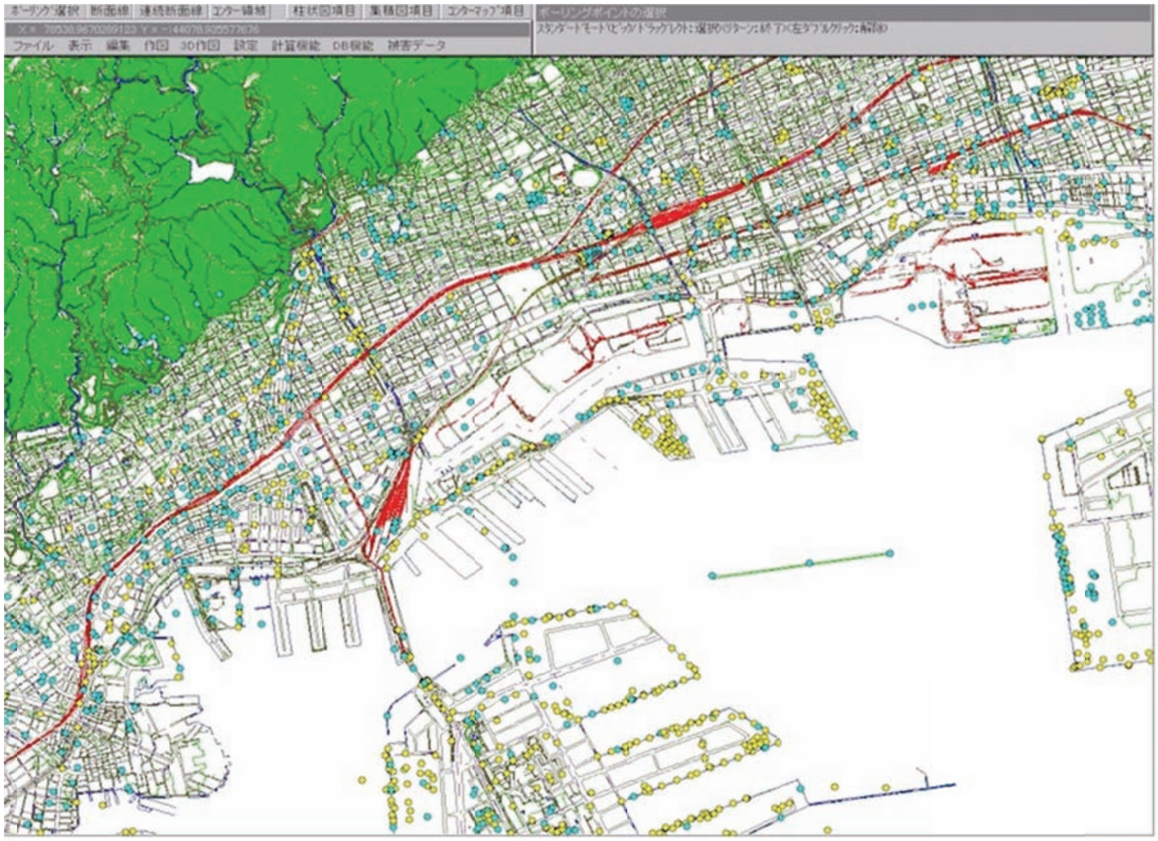


Figure 1-7. Screen shot of Kobe-Jibankun user interface. The figure shows a borehole location plan, where various dot colors show different borehole types.

(original figure from Tanaka and Okimura, 2010)

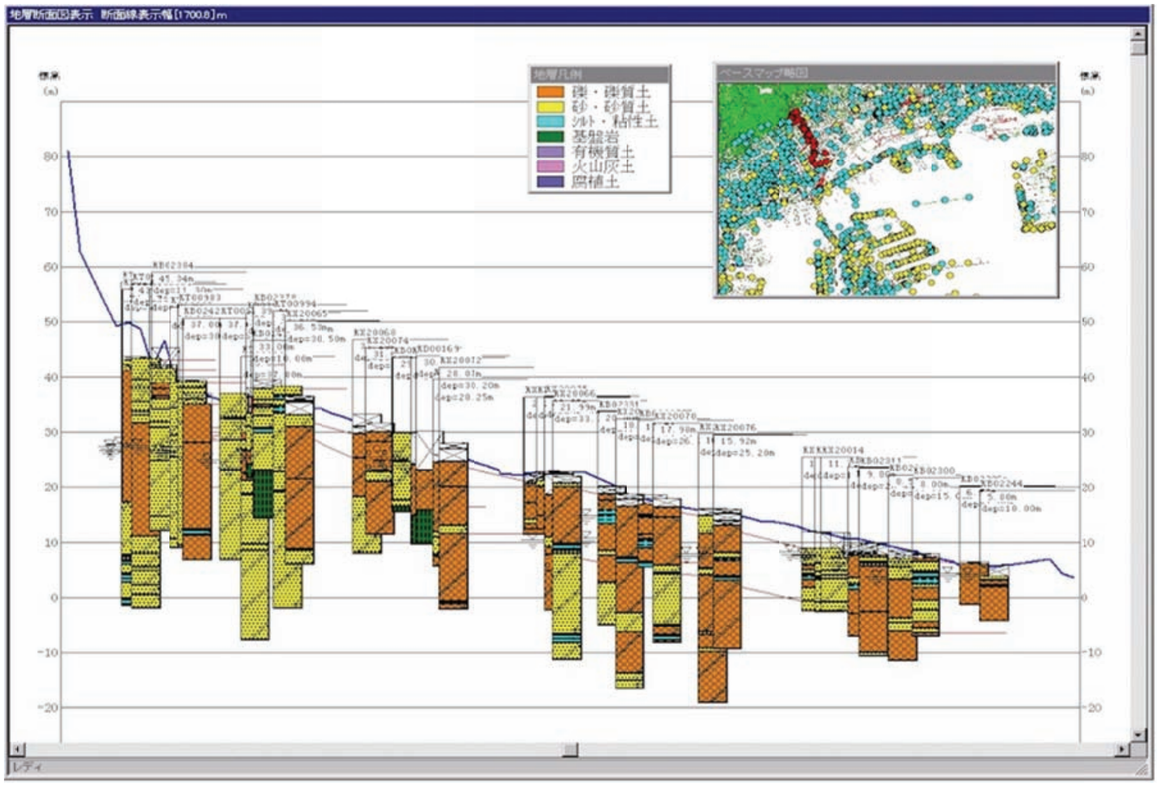


Figure 1-8. A screen shot of a fence diagram generated by Kobe-Jibankun. The inset plan shows the boreholes used to generate the fence diagram.

(original figure from Tanaka and Okimura, 2010)

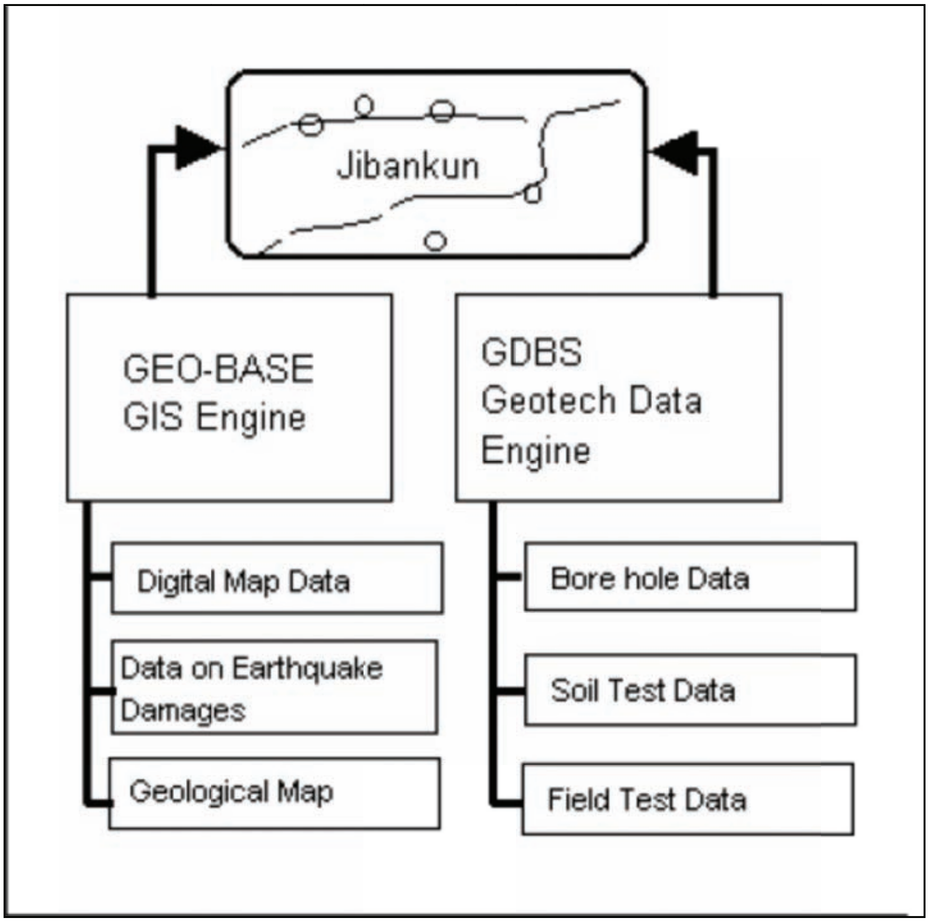


Figure 1-9. Kobe-Jibankun schematic system architecture. The system is based on a GIS engine (Geo-Base) that provides base maps, earthquake damage data and geology data, and a geotechnical database management system (GDBS) providing borehole, lab tests and field test data.

(original figure from Tanaka and Okimura, 2010)

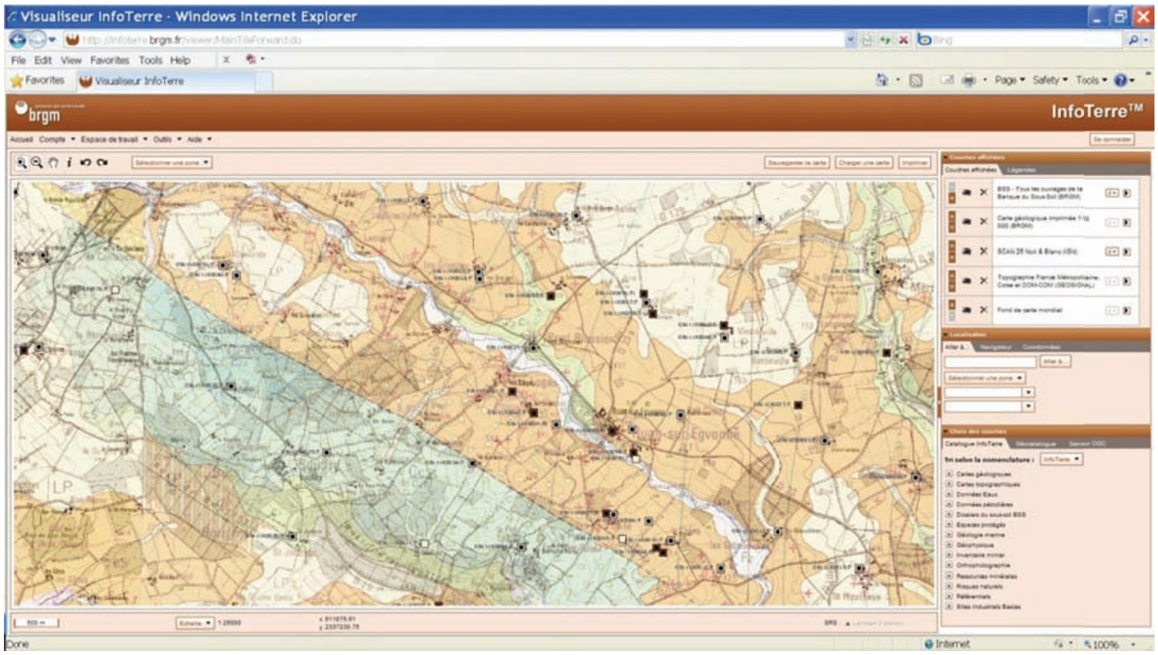


Figure 1-10. A screenshot of InfoTerre. This figure shows surface geology. InfoTerre provides different types of spatial data about earth using this interface.

(original figure from BRGM, 2010)

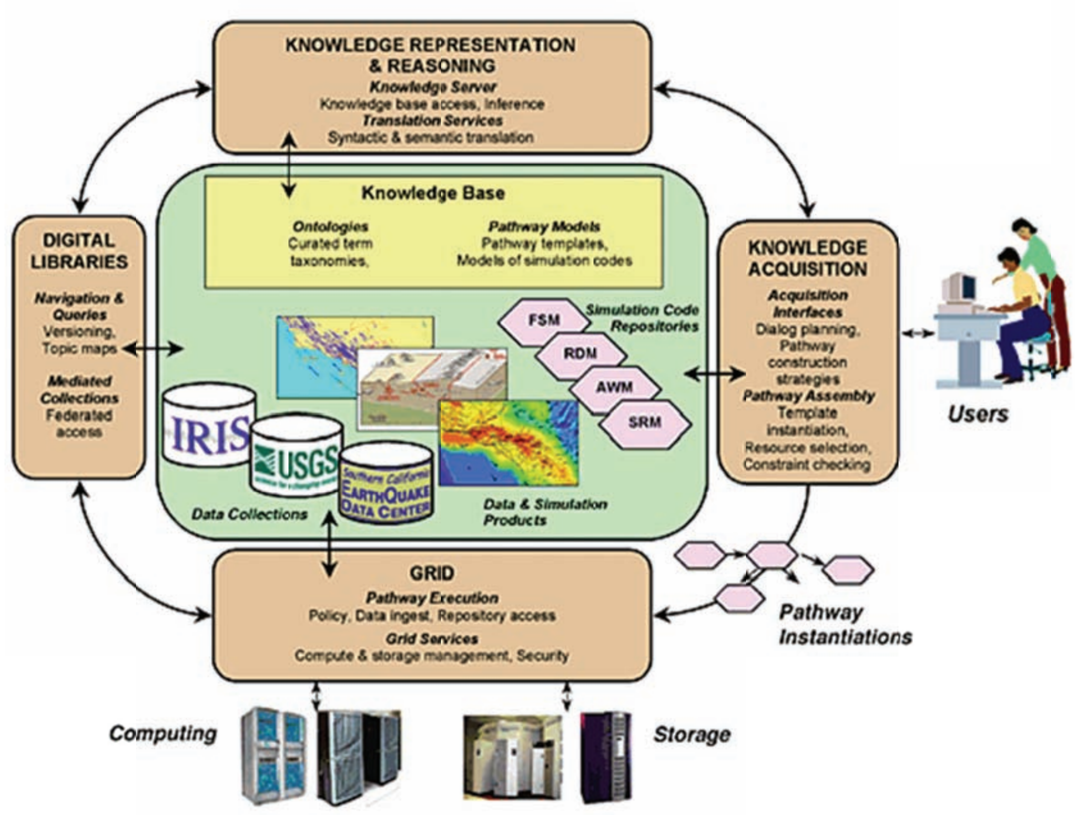


Figure 1-11. SCEC Community Modeling Environment (CME) system chart. CME uses the data from several databases, including Incorporated Research Institutions for Seismology (IRIS), USGS and SCEC, in physical models.

(original figure from SCEC, 2010)

Chapter 2. Data Types in Geotechnical Engineering

In this chapter the type of data that is dealt with in geotechnical engineering is reviewed. Various aspects of the geotechnical data are evaluated. These include the relative size of the data, data sources, as well as methods of data collection, processing and archiving. At the end of this chapter the current methods are evaluated and their shortcomings are discussed.

The chapter starts with an overview of the geotechnical engineering community and size of the field relative to the civil engineering discipline.

2.1 Geotechnical Engineering Field in the United States

Design of any structure requires knowledge of the properties of the subsurface soil, which will support the structure and transfer the loads to the ground. Geotechnical engineering is a discipline in civil engineering, which deals with analysis and design of the parts of a structure that interact with the ground that supports the structure. Geotechnical engineers are responsible for design of structure components such as spread footings, pile foundations, tunnels, earth dams and retaining walls.

There is no detailed information on the number of geotechnical engineers in the United States. According to Grigg (2000), there are about 1.5 million professional engineers currently in the United States (second to teachers only); among them are about 200,000 civil engineers. Table 2-1 shows the breakdown of industries that occupy civil engineers based on the Bureau of Labor Statistics (BLS) data (Ellis, 1998).

Based on the American Society of Civil Engineers (ASCE) division enrollment data, it is believed that about 10% of the civil engineers work in the geotechnical engineering field, which translates to about 20,000 civil engineers who practice geotechnical engineering or work in a geotechnical-engineering related field. This number provides a rough estimate of the number of professional who are somehow involved in production and usage of geotechnical data.

2.2 Data Collection in Geotechnical Engineering

Geotechnical data are collected from three major sources: (a) boreholes, trenches and test pits, (b) in-situ tests including geophysical tests, and (c) laboratory tests performed on samples collected from the boreholes. Some geotechnical data may be collected from trenches, test pits or surface observations; however, they comprise a small percentage of the data in comparison to the borehole data and are generally treated in the same manner as borehole data.

Figure 2-1 shows the main sources of geotechnical data and various disciplines that use the data. In this chapter the types of data commonly collected in geotechnical investigations are reviewed. Understanding the process of collecting and preparing these data is essential to developing the concept of spatial geotechnical data. Various methods of obtaining geotechnical data from boreholes, in-situ tests and laboratory tests are discussed below.

2.2.1.1 Geotechnical Boreholes

A borehole is any vertical, inclined, horizontal or curved hole drilled into the ground with the primary purpose of collecting information and samples on soil layers and their properties. Geotechnical boreholes are made in order to collect information on type, thickness, lateral distribution and geotechnical and geoenvironmental properties of subsurface soil and rock material. They also provide information on groundwater conditions. In addition to direct observation of soil and rock cuttings/cores from the borehole, in-situ measurements are typically performed in the borings, and soil and rock samples are collected for laboratory testing.

Boreholes are made using several methods. Table 2-2 shows the most common geotechnical borehole types. The boring method is significant because it can impact the quality of in-situ test results and samples. Table 2-3 shows various methods used for soil and rock sampling in geotechnical investigations. The sampling is usually performed in 2.5- to 10-feet intervals. It is common to take two types of samples at the same depth, consecutively. The reason is that different sample types can be used for different tests.

In addition to typical boreholes, which are drilled into the ground, there are other types of borings that are used in geotechnical investigations. The most common type is Cone Penetration Test (CPT) soundings. In this investigation method a cone-shaped probe is pushed into the ground and various parameters are measured. This test is discussed under in-situ tests.

2.2.2 In-Situ Tests

In-situ tests are preformed directly on the soil in the field versus laboratory tests, which are performed on samples taken to the lab. Using this definition, geophysical tests are also considered in-situ tests. However, they are usually classified separately. Table 2-4 lists the most common in-situ tests in geotechnical engineering, excluding geophysical tests.

The main advantage of in-situ testing is that in this method the soils are tested under natural conditions, e.g., in-situ density and natural moisture content. Sampling alters soil properties in several ways. Sandy soils with low cohesion are generally difficult to be sampled because the soil density changes due to sample disturbance. Fine-grained soils, on the other hand, are sensitive to disturbance and moisture content, although they are often better represented by laboratory samples. Some sampling methods, e.g., thin-walled tubes, provide relatively undisturbed samples, but these methods are more expensive and limited to certain soil types.

Due to this major advantage, in practice engineers rely heavily on in-situ test results to estimate parameters such as soil density and shear strength. Laboratory tests are used for properties such as corrosion and environmental testing, which are not sensitive to sample disturbance.

Current state of practice in geotechnical engineering bases a lot of design parameters on standard penetration test (SPT) correlations. Many design methods are used that utilize SPT blowcount for designing geotechnical structures. They are used for problems such as bearing capacity, settlement, liquefaction, piles and drilled shafts.

In recent years design methodologies have been developed that use CPT results instead of SPT blowcounts. The advantages of CPT over SPT are discussed later in this chapter. Overall it appears that these two tests will remain the main in-situ tests for geotechnical applications in the near future and will play an important role in any geotechnical engineering SDI. Due to the importance of these two tests, they are discussed in more detail in the following sections.

2.2.2.1 SPT

The standard penetration test (SPT) is the most common in-situ test in geotechnical investigations. It serves the dual purpose of measuring soil strength, as well as obtaining a relatively disturbed soil sample. The standard penetration test is performed using a split-spoon sampler, which is driven into the ground using a free-falling hammer.

In the United States, the test follows the ASTM D1586-08a standard. According to this standard the SPT split-spoon sampler has an inside diameter of 1-3/8 inches, an outside diameter of 2 inches and an approximate length of 25 inches.

The sampler is driven into the ground using a 140-lb hammer with a free-falling distance of 30 inches. The sampler is driven in three 6-inch intervals and the number of hammer blows for each 6-inch penetration is recorded. The number of hammer blows for a total penetration of 12 inches after an initial penetration of 6 inches is reported as SPT N-Value. Figure 2-2 shows a drilling rig equipped with SPT equipment. Such drilling rigs often alternate between SPT sampler and a larger diameter sample (e.g., Modified California Sampler) to acquire both SPT N-Value and a larger diameter for relatively undisturbed samples from the borehole. Figure 2-3 shows typical SPT results reported in a boring log.

Although it is called standard penetration test, deviations from standard procedure are pretty common in practice. One of the major deviations is using a larger diameter sampler than the SPT sampler. A larger diameter sampler is desirable because it can be used to hold a stack of rings and obtain relatively undisturbed samples. These tests are sometimes called large penetration tests (LPT). Another deviation is using larger hammers that deliver higher energy to the sampler. This facilitates sampler driving and is necessary sometimes to obtain usable samples in stiffer soils. Even using standard hammers and samplers, the energy delivered by a certain setup could deviate from the standard tests. This is a well-known problem of SPT test and the combination of these deviations often calls for some sort of correction for the blowcounts. There has been a lot of research on how to make this correction, (e.g., Aggour and Radding, 2001 and Daniel et al., 2003).

2.2.2.2 CPT

The cone penetrometer test (CPT) is a test in which a cone-shaped probe is pushed into the ground. Modern cone penetrometers are pushed into the ground using a hydraulic jack. The jack can be mounted on a truck, track, barge or drilling ships. Figure 2-4 shows several CPT test cones with different diameters. The cone has strain gauges and load cells that measure tip resistance and skin friction on the cone. Most modern cones measure pore water pressure as well. The measurements are usually performed at 5-centimeter intervals and automatically recorded in electronic format. Figure 2-5 shows a typical CPT log with both raw data (side friction, tip resistance and pore pressure) and interpreted normalized soil behavior type (SBT_N).

CPT is a fast and economical investigation method. The test procedure is standardized and measurements are taken without user interpretation. It can be performed on soils and weathered bedrock, but cannot be performed in soils with large boulders and hard rock. Another shortcoming of the test is lack of direct observation of underground soils. Although it is possible to obtain samples from CPT, it is not common because it negates one of the measure advantages of the test, being fast and economical.

Another major advantage of CPT is that the results provide a continuous interpretation of soil properties. In comparison, SPT samples are taken at discrete intervals, and material variations between two consecutive samples generally cannot be determined directly and are subject to human interpretation. Moreover, the standard and continuous nature of CPT measurement can be easily adapted to automatic analysis and design routines. For example, liquefaction analysis can be performed directly using CPT data, and the analysis is subject to much less human interpretation than SPT-based methods.

CPT results have been correlated to a large number of soil parameters. Table 2-5 shows the perceived applicability of CPT test results for estimating other soil parameters (Robertson and Cabal, 2009).

In the past 20 years CPT probes have become more sophisticated. Figure 2-6 shows several advanced cones, which measure additional soil parameters as well as typical tip resistance/side friction. Due to these new CPT techniques and development of new design methods that rely on CPT measurements, it is expected that this test will become more significant in the future and will comprise a larger share in geotechnical data than before.

2.2.2.3 Geophysical Tests

Geophysical tests are an indirect method of collecting data on soil and rock formations. They are often used to collect quick and economical data over large areas. The data obtained from these tests is general and large scale. They are particularly useful in areas with shallow rock formations. The most common geophysical test is seismic refraction survey. Table 2-6 provides a list of a common geophysical test and the pros and cons of each method.

Most geophysical methods are based on elastic wave theory. Figure 2-7 shows the basic theory behind seismic direct, reflection and refraction tests. More information on these tests can be found in Hunt (2005).

2.2.3 Laboratory Soil Tests

Soil samples taken from the field are generally transferred to the soil lab for classification and testing. The most common tests are density (from a relatively undisturbed sample), moisture content, sieve analysis and soil classification. Other soil tests are assigned by the geotechnical engineers based on the requirements of the project. In the United States, soil laboratory tests are standardized by ASTM.

As discussed before, some degree of sample disturbance is inevitable and this fact impacts some of the measured properties in the lab, while other properties are not affected. This problem somehow limits the use of laboratory tests for strength parameters, but they are significant for other tests such as corrosion and environmental tests.

Table 2-7 lists some the most common laboratory soil tests. Most soil labs use Excel spreadsheets to report test data. The results are usually presented by borehole name, sample ID and depth from top of the borehole. The borehole name and depth from the surface allows users to determine the original position of the test sample. Figure 2-8 shows a lab test form for a direct shear test. Such test results may be provided in hard copy, PDF and spreadsheet format.

2.3 Geotechnical Data Reporting and Archiving

After completion of fieldwork and laboratory testing operations, geotechnical data are converted to different formats based on project requirements. These formats include traditional hard copies, PDF and image formats (e.g., jpg and tiff), spreadsheet and database formats, as discussed in the following sections.

2.3.1 Hard Copies

Traditionally, geotechnical data have been prepared as boring logs and laboratory forms in hard copy format. The hard copies used to be prepared by hand drawing. Later, computer software was used to prepare the hard copies (e.g., LogPlot, AutoCad and Microstation). However, during the transition period consultants usually used the software internally to prepare hard copies, while the electronic-format data was left un-submitted to the owner of the data and subsequently was lost.

Agencies such as Caltrans have large archives of soil data in hard copy format. This format is bulky, hard to retrieve and vulnerable to loss and deterioration due to physical damage. Some of these archives were converted to microfilm format, which has the same problem to a lesser degree.

2.3.2 Electronic Image Formats

In recent years imaging and CAD software have been used for storing and reporting the data. Some agencies require consultants to provide boring logs and test results in PDF and/or CAD format. For example, for the past few years Caltrans requires boring logs to be submitted in PDF and Microstation DGN format.

Some agencies have converted their hard copy archives to electronic format (usually scanned to PDF or image format). Unfortunately, during this conversion some data often is lost, either due to lost original hard copies or due to loss of resolution and legibility during scanning.

One example is Caltrans Bridge Inspection Records Information System (BIRIS), which provides an archive of bridge inspection records and plans, including LOTBs. This database is not public and is only accessible by Caltrans employees within the agency network. Figure 2-9 shows a sample LOTB downloaded from BIRIS in PDF format.

2.3.3 Geotechnical Databases

Database programs for geotechnical engineering have been in use for more than 20 years. They proliferated from the mining and oil industry, where the scale of investigations is much larger than in geotechnical engineering. Table 2-8 shows a list of geotechnical database programs from the geotechnical and geoenvironmental engineering software directory website GGSD (2011).

Among these programs gINT, TechBase and Rockware are among the more widely used programs in the United States. These programs are essentially a relational database with a relatively simple built-in schema for boreholes. The schema is customizable and users add new data columns to existing tables, or add new tables, based on their needs.

Hereafter, the gINT program is discussed as an example of the current workflow of geotechnical data processing and archiving in the industry. The reason for selection of this program is: a) it is the standard program for geotechnical data management and reporting in the United States, and b) the author has extensive personal experience in using this program.

2.3.3.1 gINT Geotechnical Software

gINT is the most widely used program in geotechnical and geoenvironmental engineering in the United States for more than 20 years. The program is developed by gINT Software, which is acquired by Bentley Systems, a major infrastructure software company whose main product is Microstation CAD software.

gINT is based on a relational database engine (Microsoft Access) and provides customizable database tables and graphing and reporting templates. The idea behind this architecture is that each client requires a different database schema, based on project needs. Therefore, it is not practical to provide a comprehensive standard schema for all applications.

The data in gINT are entered in tables. Figure 2-10 shows a gINT form for entering borehole data. The relationships between data fields are established using key data fields. For example, the combination of borehole ID and depth from top of boring is used as a composite key for most data fields.

In addition to storing data, gINT provides templates for reports and borehole logs. Figure 2-11 shows a report on SPT N-values, presenting the measured SPT values and statistics on the measured data. A fence diagram is a 2-dimensional cross-section along an alignment, where each borehole is projected on the vertical plane passing though the alignment. Figure 2-12 shows a gINT fence diagram generated along a road alignment.

2.3.4 GIS-Based Archiving of Geotechnical Data

The current trend in geotechnical data dissemination is moving toward Web-based database systems with GIS interface. This method has been used in recent years by some agencies to make their geotechnical data archive available through the Internet.

There are two different types of Web-based geotechnical databases. The first type provides the metadata for querying but stores the main data (e.g., boring log) in an image format (PDF or jpg). The client queries the metadata based on criteria, and the database provides a list of boreholes that meet the criteria and a link for downloading the image file of each borehole log.

The second approach stores the actual borehole data in a database using a comprehensive database schema developed for geotechnical data. The first approach is more practical for legacy data because they are either available or can be easily converted to image format. Therefore the database can be developed at a reasonable timeframe and cost. The second methodology provides much more functionality for new borehole data, because it provides search and query mechanisms on a wide range of data. For example, a user can query all boreholes within a large project area, which has sandy layers with SPT N-Value smaller than 15. The query results give an approximate idea about project areas with high liquefaction potential. As a downside, developing the database and populating it is not practical for legacy data. It is also possible to use a combination of these two methods in the same database to take advantage of new data in database format and legacy data in image format.

Table 2-9 lists some of the existing Web-based geotechnical data management and archiving systems. This list should not be regarded as a complete list; due to the large number of the ongoing projects in this area it was not possible to list all of them in this table.

As an example of this type of geotechnical archive, the COSMOS/PEER-LL Geotechnical Virtual Data Center in the following section discusses some of the Web-based database systems for geotechnical data.

2.3.4.1 COSMOS/PEER LL Geotechnical Virtual Data Center

COSMOS/PEER LL Geotechnical Virtual Data Center (GVDC) is a distributed system for archiving and dissemination of geotechnical data. COSMOS stands for Consortium of Organizations for Strong Motion Observation Systems. It was established in 1999 using an NSF grant with four core members: California Geological Survey (CGS), US Army Corps of Engineers, US Bureau of Reclamation and US Geological Survey (USGS). However, there are other non-core members, including a number of non-governmental companies and agencies who have donated strong motion records for distribution through COSMOS. According to the COSMOS Website the mission statement is as follows:

“To advocate for the establishment of strong-motion measurement systems; to promote development and adoption of verifiable, internationally-ranked standards for the acquisition and processing of earthquake strong-motion measurements; and to promote the global application of strong-motion measurements by design professionals.”

The major project of COSMOS is its Strong Motion Virtual Data Centre (SMVDC), which provides access to a large database of earthquake strong motions from all over the world. In addition to SMVDC, COSMOS and Pacific Earthquake Engineering Research Center Lifelines Program (PEER-LL) have developed the GDVC with support from major agencies and companies, including California Geological Survey, Caltrans, the U. S. Geological Survey, and the Pacific Gas and Electric Company, which provide the geotechnical data. The GDVC does not archive the geotechnical data, it provides the metadata for searching and querying the data, along with a

URL for downloading the actual data, which is maintained by the data providers. The data is archived in a standard format developed by GDVC.

The geotechnical data provided by GDVC include:

- Geophysical tests
- In-situ tests
- Laboratory tests

As part of the GDVC project, COSMOS/PEER-LL performed extensive surveys regarding acquisition and usage of geotechnical data from researchers, academics and practicing professionals.

The workflow in the GDVC project involves making the geotechnical data from providers available to end users through the Internet. A standard Web browser is used as the main user interface. COSMOS developed user interfaces for querying and downloading data via GVDC to end-user PCs. The GVDC is also integrated with COSMOS VDC for earthquake records, so both earthquake records and geotechnical data can be downloaded by a single query. Figure 2-13 shows a screen shot of the GDVC query page.

Other characteristics of the GDVC include plotting boring logs from XML data files using scalable vector graphics (SVG). SVG is an XML-based file format for describing 2-dimensional graphics in vector format. It is mainly used to visualize graphics in Web browsers. Conversion from GDVC XML format to SVG can be done using an XSLT transformation.

The GDVC data format is an XML-based data format, which at this time is still under development. COSMOS GDVC data format is merging with DIGGS GML compatible data format, although the new format is not fully implemented yet. This data format is discussed in detail in Chapter 4.

2.3.5 Data Transfer Formats

In the past 20 years, the acquisition and exchange of geotechnical data has migrated from hard copies to digital format. Digital data collection and transmission is used in field investigation, laboratory testing and report submissions. These operations are performed using various applications including spreadsheets (e.g., Excel), proprietary geotechnical software (e.g., gINT) and CAD programs (e.g., Microstation). This process requires the geotechnical data to be transferred between different applications.

Although most geotechnical applications use proprietary data formats, they also have built-in features to exchange data with other programs. For example, users can import and export data between gINT and Microsoft Excel spreadsheets and Access databases. Despite the high frequency of this operation, the industry has not adopted a standard data exchange format for such purposes. Most data exchange operations are timely and costly, and occasionally result in loss of data integrity and/or conflicting and redundant information.

Existing problems with the exchange of geotechnical data have been recognized since the early 1990s and have led to the introduction of a number geotechnical data formats. Table 2-10 shows a list of file and data formats that are used for data transfer and storage by the geotechnical community. The table also shows the number of existing compatible programs and the current status of usage and development of each format. Some of these file formats are used by proprietary software (e.g., GeoMil and Geopoint) and are not intended to be used between different applications. Other formats (e.g., AGS format) are open standards developed by the industry in order to facilitate data exchange between different entities involved in a project.

Unlike other similar industries, like petroleum engineering with Log ASCII (LAS) format (Heslop et al., 2003, and Canadian Well logging Society, 2008), these efforts have not been widely adopted by the geotechnical engineering community. This issue can be attributed to complexity of data formats, lack of support from mainstream geotechnical software, weak promotion and advertising to the industry, and lack of collaboration and consensus at national and international levels between agencies that own large amounts of data.

Review of the data provided in Table 2-10 clearly shows that the AGS is the data exchange format with the highest adoption in the industry. As part of this background review, the AGS data format has been discussed in the next section.

2.3.5.1 AGS Data Format Structure

The Association of Geotechnical and Geoenvironmental Specialists (AGS) developed one of the most widely used data formats for exchanging geotechnical information. The first edition of AGS was introduced in 1991, and was subsequently updated in a second (1994) and third (1999) edition. The current edition (i.e., 3.1) was published in 2004 and updated in 2005. After undergoing more than 15 years of application and revision, the AGS format is now supported by a large community of practicing engineers and captures the most commonly used types of geotechnical information.

AGS uses a comma separated value (CVS) file format, which can be imported in typical spreadsheet programs, and derives from a relational data model. Table 2-11 shows an excerpt from an AGS data file (AGS, 2008) with one test pit (TP501) and one borehole (BH502). Throughout the remainder of this section, `courier` font is used to identify entities from AGS data format. Figure 2-14 shows the corresponding representation of these data. An AGS file is a succession of tables that contain data from various data groups such as different geotechnical tests. Each data group begins with a line starting with double asterisks (e.g., "`**HOLE`") followed by the data group name. The data group name line is followed by another line containing the headers of the data fields in the data group. Each header starts with a single asterisk (e.g., "`*HOLE_ID`"). The following line contains the mandatory units for each data field. If a data field does not have a unit, an empty unit field is specified (e.g., ""). The actual data are listed in subsequent lines, wherein each line contains one data entry. Due to this tabular structure and use of key fields, AGS data can be imported easily into a relational database.

AGS has a comprehensive data dictionary, which defines the terminology used for characterizing geotechnical information. AGS includes 74 tables, or data groups, as listed in Table 2-12. The data fields in each group are identified by headings. In addition to these 74 data groups, AGS accepts user-defined data groups for introducing new types of geotechnical information. AGS user-defined data groups have to be defined in data files for completeness. The relationships between different types of AGS tables are defined by one or more key fields, which are conceptually similar to foreign keys in relational databases. For example, Table 2-13 shows the data fields in the data group "PROJ", which stores basic information for a geotechnical project. In this table, PROJ_ID is a key field, which uniquely identifies this project.

Through extensive use by engineering professionals and after undergoing several revisions, the AGS format is now mature and well accepted in the geotechnical engineering community. However, in the context of modern information technology and in comparison with what other disciplines in engineering and science nowadays are using, it is a primitive data format and remains difficult to integrate with the Internet, Web services and GIS applications.

2.4 Current Methods' Shortcomings

Based on the review of the current methods used in the acquisition, exchange and submission of geotechnical data, the following shortcoming are identified:

- Data cannot be easily exchanged due to lack of an industry-standard data exchange format.
- Data cannot be easily transformed to a different format or layout (hard copies).
- There is a disconnect between the geotechnical engineering community and other disciplines in terms of geotechnical data exchange, mainly because the geotechnical community cannot agree on a standardized data modeling and exchange format.
- Geotechnical data has poor integration with Internet based applications, Web services and spatial databases.
- The geotechnical engineering community in general is lagging behind similar disciplines in terms of adopting new trends in information technology.

The objective of this research is to explore technologies and solutions that can help in solving these problems, e.g., increasing the efficiency of data usage in geotechnical engineering applications, promoting integration of geotechnical data with other sources of spatial data and facilitating multi-disciplinary geotechnical data exchange.

Table 2-1. Industry-occupation matrix for civil engineers based on Bureau of Labor statistics data.

Employment Category	Number	Percentage
State and Local Government	58,653	32.0
Engineering Services and Management and Accounting	81,340	44.4
Electric and gas utilities and communications	3,905	2.1
Construction	10,852	5.9
All manufacturers	8,236	4.5
Federal government	12,622	6.9
Other services and utilities	3,190	1.7
Other	2,158	1.2
Research and testing	2,119	1.2
Total	183,102	100.0

(after Ellis, 1998)

Table 2-2. Common borehole types in geotechnical engineering investigations.

Boring Type	Uses
Solid stem auger	Used in dry holes in competent materials. May need to use casing for collapsing material.
Hollow stem auger	Similar to solid stem (continuous flight) auger drilling, except hollow stem is screwed into the ground and acts as casing. Sampling and testing from inside of auger. Penetration in strong soils/gravel layers difficult.
Wash boring	Used to advance the borehole and keep the hole open below the water table. Fluid may be mud (polymer) or water depending on the soil conditions. Maintains hydrostatic head.
Rock coring	Hardened cutting bit with a core barrel used to obtain intact rock samples.
Air track probes	Provides a rapid determination of rock quality/depth to rock based on the time to advance the hole. Rock assessment is difficult as rock chippings only obtained.

(after Look, 2007)

Table 2-3. Common sample types in geotechnical engineering investigations.

Symbol	Sample or Test
TP	Test pit sample
W	Water sample
D	Disturbed sample
B	Bulk disturbed sample
SPT	Standard penetration test sample
C	Core sample
U (50)	Undisturbed sample (50mm diameter tube)
U (75)	Undisturbed sample (75mm diameter tube)
U (100)	Undisturbed sample (100mm diameter tube)

(after Look, 2007)

Table 2-4. List of common field tests in geotechnical engineering investigations.

Symbol	Test	Measurement
DCP	Dynamic cone penetrometer	Blows/100mm
SPT	Standard penetration test	Blows/300mm
CPT	Cone penetration test	Cone resistance qc (MPa); friction ratio (%)
CPTu	Cone penetration test with cone resistance	Cone resistance qc (MPa); friction ratio (%); pore pressure measurement pressure (kPa). time for pore pressure (Piezocone); dissipation t (sec)
PT	Pressuremeter test	Lift-off and limit pressures (kPa); volume change (cm3)
PLT	Plate loading test	Load (kN); deflection (mm)
DMT	Dilatometer test	Lift-off and expansion pressures (kPa)
PP	Pocket penetrometer test	kPa
VST	Vane shear test	Nm, kPa
WPT	Water pressure (packer) test	Lugeons

(after Look, 2007)

Table 2-5. Perceived applicability of CPT test for deriving soil design parameters (Robertson and Cabal, 2009).

Soil Type	D_r	ψ	K_0	OCR	S_t	s_u	ϕ'	E, G*	M	G_0*	k	c_h
Sand	2-3	2-3		5			2-3	2-3	2-3	2-3	3	3-4
Clay			2	1	2	1-2	4	3-4	2-3	3-4	2-3	2-3

1 = high, 2 = high to moderate, 3 = moderate, 4 = moderate to low, 5 = low reliability, Blank = no applicability, * improved with SCPT

where:

D_r	Relative density	ψ	State parameter
K_0	In-situ stress ratio	OCR	Over consolidation ratio
S_t	Sensitivity	s_u	Undrained shear strength
ϕ'	Friction angle	E, G	Elastic and shear moduli
M	Bulk modulus	G_0	Small-strain shear modulus
k	Permeability	c_h	Coefficient of consolidation

(after Robertson and Cabal, 2009)

Table 2-6. List of common geophysical tests in geotechnical engineering.

Method	Application	Comments
Seismic refraction from surface	Obtain stratum depths and velocities, land or water. Geologic sections interpreted	Most suitable if velocities increase with depth and rock surface regular
Seismic direct (crosshole, downhole)	velocities for particular strata; dynamic properties; rock mass quality	Requires drill holes. Crosshole yields best results. Costly
Seismic reflection	General subsurface section depicted. Water bodies yield clearest sections	Land results difficult to interpret. Velocities not obtained. Stratum depths comps require other data
Electrical resistivity	Locate saltwater boundaries, clean granular and clay strata, rock depth, and underground mines by measured anomalies	Difficult to interpret. Subject to wide variations. No engineering properties obtained. Probe configurations vary
Gravimeter	Detect faults, domes, intrusions, cavities, buried valleys by measured anomalies	Precise surface elevations needed. Not commonly used, measures density differences
Magnetometer	Mineral prospecting, location of large igneous masses	Normally not used in engineering or groundwater studies
Ground-probing radar	General subsurface section depicted. Most useful to shows buried pipe, bedrock, voids, boulders	Interpretation difficult. Limited to shallow depths. No engineering properties
Thermography	Shallow subsurface section depicted. Useful for water pipeline leaks	Interpretation difficult. Limited to shallow depths. No engineering properties

(after Hunt, 2005)

Table 2-7. List of some of the most common laboratory tests in geotechnical engineering.

Test Name	Comments
Moisture Content	Determine moisture content of the sample
Density	Determine density of the sample
Specific Gravity	Determine specific gravity of the sample
Sieve Analysis	Determine grain size distribution for coarse grains
Hydrometer	Determine grain size distribution for fine grains
Atterberg Limits	Determine liquid limit and plastic limit
Soil Classification	Determine soil classification per one of the classification methods
Compaction	Determine maximum dry unit weight and optimum moisture content
R-Value	Measures the response of a compacted sample of soil or aggregate to a vertically applied pressure under specific conditions
Direct Shear	Measures the shear strength of soils under direct shear
Unconfined Compression	Measures the unconfined compressive strength of cohesive soils
Triaxial Test	Measures the shear strength of soils under triaxial loading condition
Consolidation	Measures the compressibility of fine-grained soils

Table 2-8. Partial list of geotechnical engineering database programs.

Program	Status	Operating system
BLDM	Commercial	DOS, Intergraph Microstation
GDM	Commercial	Win95/98, WinNT, Win2000
GeoBASE	Commercial	Win95/98
GEODASY	Commercial	Win95/98, WinNT
GEO-LOG 2	Commercial	Win3x, Win95/98, WinNT
GEO-LOG 3	Commercial	Win95/98, WinNT, Win2000
gINT Logs Plus	Commercial	Win2000, WinXP
gINT Professional	Commercial	Win2000, WinXP
GIS-Key Winlogs	Commercial	Win95/98, WinNT
HoleBASE III	Commercial	Win95/98, WinNT, Win2000, WinXP
HoleBASE Wizard	Commercial	Win95/98, Excel
Hydro GeoLogger	Commercial	Win2000, WinXP, WinVista
SID	Freeware	DOS, Win3x, Win95/98, WinNT
TECHBASE	Commercial	DOS, Win95/98, WinNT, UNIX, LINUX, Open VMS

(after GGSD, 2011)

Table 2-9. Some of the Web-based geotechnical data management and archiving systems.

Name	Description	Owner/Developer	URL
GeoDOG	Caltrans Geotechnical Services electronic database that houses geotechnical information and documents produced during the course of project development	Caltrans	N/A Caltrans employee access only
COSMOS/PEER-LL GVDC	A distributed system for archiving and dissemination of geotechnical data. Provides access to several data providers' geotechnical data using a common data format.	COSMOS/PEER-LL	http://geodata2.usc.edu/
Kobe-Jibankun	Geotechnical database and GIS system for Kobe City	Kobe City	N/A
National Geotechnical Database	Holds geotechnical information extracted from site investigation records provided by clients, consultants and contractors and from field and laboratory testing carried out by the British Geological Survey. Uses AGS format.	British Geological Survey	N/A

Table 2-10. List of file and data formats for geotechnical engineering applications.

Format Name	File Extension	Description	Status	Ownership	Number of Compatible Programs
AGS	.AGS	Geotechnical field, laboratory and monitoring data transfer format.	Active	Open source	30
AGS-M	.AGS	Monitoring data transfer format. Now included in AGS 3.1	In use	Open source	1
AGSML	.xml	Geotechnical field, laboratory and monitoring transfer format. Incorporated within DIGGS	Development	Open source	0
bch	.bch	Inclinometer monitoring data file type.	Obsolete	Proprietary	1
CivilXML	.xml	Construction records - principally piling, and associated geotechnical data. Incorporated within DIGGS.	Development	Open source	0
COSMOS	.xml	California based geotechnical data transfer format for earthquake studies. Incorporated within DIGGS	Development	Open source	0
DIGGS	.xml	Data interchange for geotechnical and geoenvironmental specialists. Based on AGSML, COSMOS and CivilXML	Development	Open source	0
GEF	.gef	CPT test results	Active	Open source	4
GeoMil		CPT test results	In use	Proprietary	2
Geopoint		CPT test data format	In use	Proprietary	1
GeotechML	.xml	Geotechnical field, laboratory and monitoring transfer format	Inactive	Open source	0
Gorilla!		CPT test results	In use	Proprietary	1
NENGEO		Dutch CPT test results format.	Obsolete	Open source	1
NGES	.xml	Geotechnical data exchange format for scientific research sites	In use	Open source	0
RocProp		Database of Rock Properties	In use	Proprietary	0
RPP	.rpp	Inclinometer monitoring data	In use	Proprietary	1
SGF		CPT test results	In use	Open source	2
SlopesML	.xml	Slope stability case history format	Inactive	Open source	0

(after GGSD, 2010)

Table 2-11. Excerpt from an AGS data transfer format instance file.

```
"**PROJ"
"*PROJ_ID","*PROJ_NAME","*PROJ_LOC","*PROJ_CLNT","*PROJ_ENG","*PROJ_CONT","*PROJ_DATE","*
PROJ_AGS","*FILE_FSET"
"<UNITS>","","","","","","dd/mm/yyyy","",""
"7845","Trumpington Sewerage","Trumpington","Trumpington District Council","Geo-Knowledge
International","Lithosphere Investigations Ltd","23/07/1999","3","FS001"
"**HOLE"
"*HOLE_ID","*HOLE_TYPE","*HOLE_NATE","*HOLE_NATN","*HOLE_GL","*HOLE_FDEP","*HOLE_STAR","*
HOLE_LOG","*FILE_FSET"
"<UNITS>","","m","m","m","m","dd/mm/yyyy","",""
"TP501","TP","523196","178231","61.86","3.25","21/07/1999","ANO","FS002"
"BH502","IP+CP","523142","178183","58.72","15.45","22/07/1999","ANO","FS003"
"**GEOL"
"*HOLE_ID","*GEOL_TOP","*GEOL_BASE","*GEOL_DESC","*GEOL_LEG","*GEOL_GEOL","*GEOL_STAT","*
FILE_FSET"
"<UNITS>","m","m","","","","",""
"TP501","0.00","0.25","Friable brown sandy CLAY with numerous rootlets
(Topsoil)","101","TS","A",""
"TP501","0.25","1.55","Firm brown slightly sandy very closely fissured CLAY with some
fine to coarse subrounded gravel. Medium spaced subhorizontal slightly polished gleyed
shear surfaces. Widely spaced vertical rough desiccat","","","",""
"<CONT>","","","ion cracks with concentrations of rootlets. (Weathered Boulder
Clay)","261","WBC","B",""
"TP501","1.55","3.25","Stiff grey closely fissured CLAY with a little fine to medium
subrounded gravel and rare sandstone cobbles (Boulder Clay)","250","BC","C",""
"BH502","0.00","0.30","Friable brown sandy CLAY with numerous rootlets
(Topsoil)","101","TS","",""
"BH502","0.30","2.60","Firm brown very closely fissured CLAY with a little fine to medium
subrounded gravel (Weathered Boulder Clay)","250","WBC","",""
"BH502","2.60","5.75","Stiff grey slightly sandy closely fissured CLAY with some fine to
coarse subrounded gravel (Boulder Clay)","261","BC","",""
"BH502","5.75","15.45","Dense becoming very dense yellow brown very sandy fine to coarse
subrounded GRAVEL (Glacial Gravels)","307","GG","",""
"**SAMP"
"*HOLE_ID","*SAMP_TOP","*SAMP_REF","*SAMP_TYPE","*SAMP_BASE","*SAMP_DATE","*SAMP_TIME","*
GEOL_STAT","*FILE_FSET"
"<UNITS>","m","","","m","dd/mm/yyyy","hhmmss","",""
"TP501","1.00","1","D","1.00","","","B",""
"TP501","1.00","2","B","1.30","","","B",""
"TP501","2.50","3","B","2.75","","","C",""
…
```

(original data from AGS, 2008)

Table 2-12. List of AGS data transfer format data groups.

No.	Data Group	Description
1	PROJ	Project Information
2	ABBR	Abbreviation Definition
3	?BKFL	Backfill Detail
4	CBRG	CBR Test – General
5	CBRT	CBR Test
6	CDIA	Casing Diameter By Depth
7	CHEM	Chemical Tests
8	CHIS	Chiseling Details
9	CHLK	Chalk Tests
10	CLSS	Classification Tests
11	CMPG	Compaction Tests – General
12	CMPT	Compaction Tests
13	CNMT	Contaminant And Chemical Testing
14	CODE	Chemical Testing Codes
15	CONG	Consolidation Test – General
16	CONS	Consolidation Test
17	CORE	Rotary Core Information
18	DETL	Stratum Detail Description
19	DICT	User Defined Groups And Headings
20	DISC	Discontinuity Data
21	DPRG	Dynamic Probe Test – General
22	DREM	Depth Related Remarks
23	FILE	Associated Files
24	FLSH	Rotary Core Flush Detail
25	FRAC	Fracture Spacing
26	FRST	Frost Susceptibility
27	GAST	Gas Constituents
28	GEOL	Stratum Description
29	GRAD	Particle Size Distribution Analysis Data
30	HDIA	Hole Diameter By Depth
31	?HDPH	Depth Related Hole Information
32	HOLE	Hole Or Location Equivalent
33	HPGI	Horizontal Profile Gauge Installation Details
34	HPGO	Horizontal Profile Gauge Observations
35	ICBR	In Situ CBR Test
36	?ICCT	In Situ Contaminant And Chemical Testing
37	IDEN	In Situ Density Test
38	?IFID	On Site Volatile Headspace Testing Using Flame Ionisation Detector
39	INST	Single Point Instrument Installation Details
40	IOBS	Single Point Instrument Readings
41	?IPID	On Site Volatile Headspace Testing By Photo Ionisation Detector
42	IPRM	In Situ Permeability Test
43	IRDX	In Situ Redox Test
44	IRES	In Situ Resistivity Test
45	ISPT	Standard Penetration Test Results
46	IVAN	In Situ Vane Test
47	MCVG	MCV Test – General
48	MCVT	MCV Test
49	?MONP	Monitor Point
50	?MONR	Monitor Point Reading
51	POBS	Piezometer Readings
52	PREF	Piezometer Installation Details
53	PROB	Profiling Instrument Readings
54	PROF	Profiling Instrument Installation Details

Table 2-12. Continued

No.	Data Group	Description
55	PRTD	Pressuremeter Test Data
56	PRTG	Pressuremeter Test Results, General
57	PRTL	Pressuremeter Test Results, Individual Loops
58	PTIM	Hole Progress By Time
59	PTST	Laboratory Permeability Test
60	PUMP	Pumping Test
61	RELD	Relative Density Test
62	ROCK	Rock Testing
63	SAMP	Sample Reference Information
64	SHBG	Shear Box Testing – General
65	SHBT	Shear Box Testing
66	STCN	Static Cone Penetration Test
67	SUCT	Suction Tests
68	TNPC	Ten Percent Fines
69	?TREM	Time Related Remarks
70	TRIG	Triaxial Test – General
71	TRIX	Triaxial Test
72	UNIT	Definition Of <UNITS> And CNMT_UNIT
73	WETH	Weathering Grades
74	WSTK	Water Strike Details

Table 2-13. AGS data transfer format "PROJ" data group. This data group contains the information for the data file project.

No.	Heading	Unit	Description
1	PROJ_ID*		Project identifier
2	PROJ_NAME		Project title
3	PROJ_LOC		Location of site
4	PROJ_CLNT		Client name
5	PROJ_CONT		Contractors name
6	PROJ_ENG		Project Engineer
7	PROJ_MEMO		General project comments
8	PROJ_DATE	mm/dd/yyyy	Date of production of data
9	?PROJ_CID		Monitoring Contractor Identifier
10	?PROJ_PROD		Data file producer
11	?PROJ_RECV		Data file recipient
12	?PROJ_ISNO		Issue sequence number
13	?PROJ_STAT		Status of data within submission
14	PROJ_AGS		AGS Edition Number
15	FILE_FSET		Associated file reference

* key field.

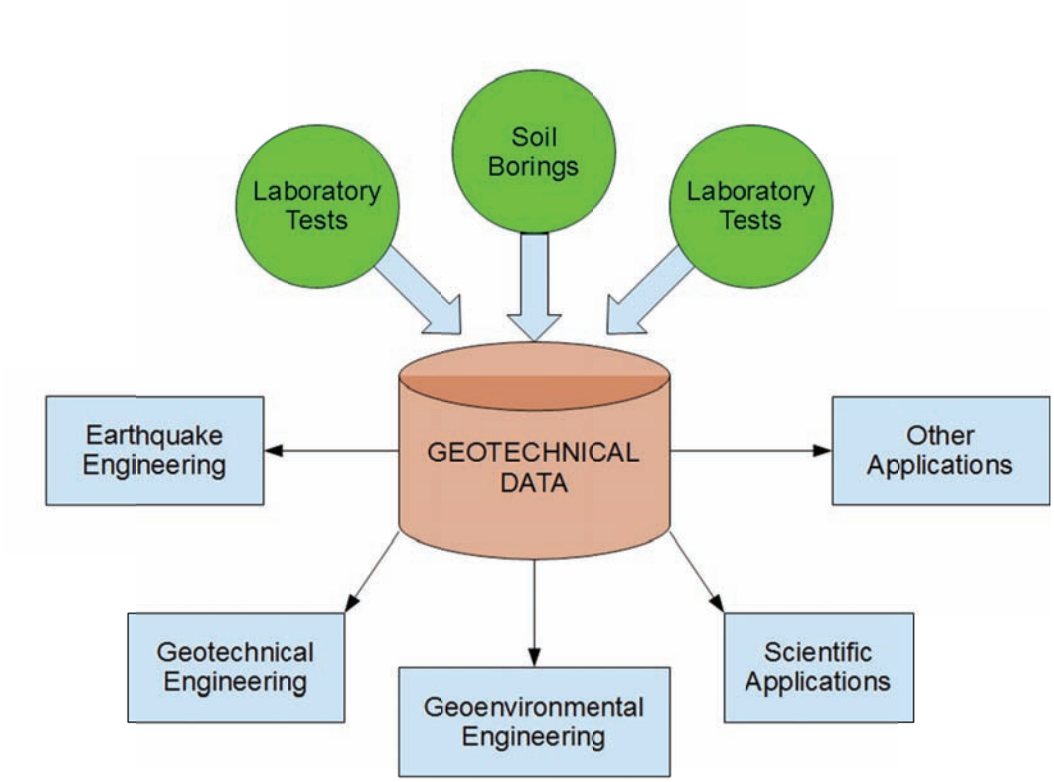
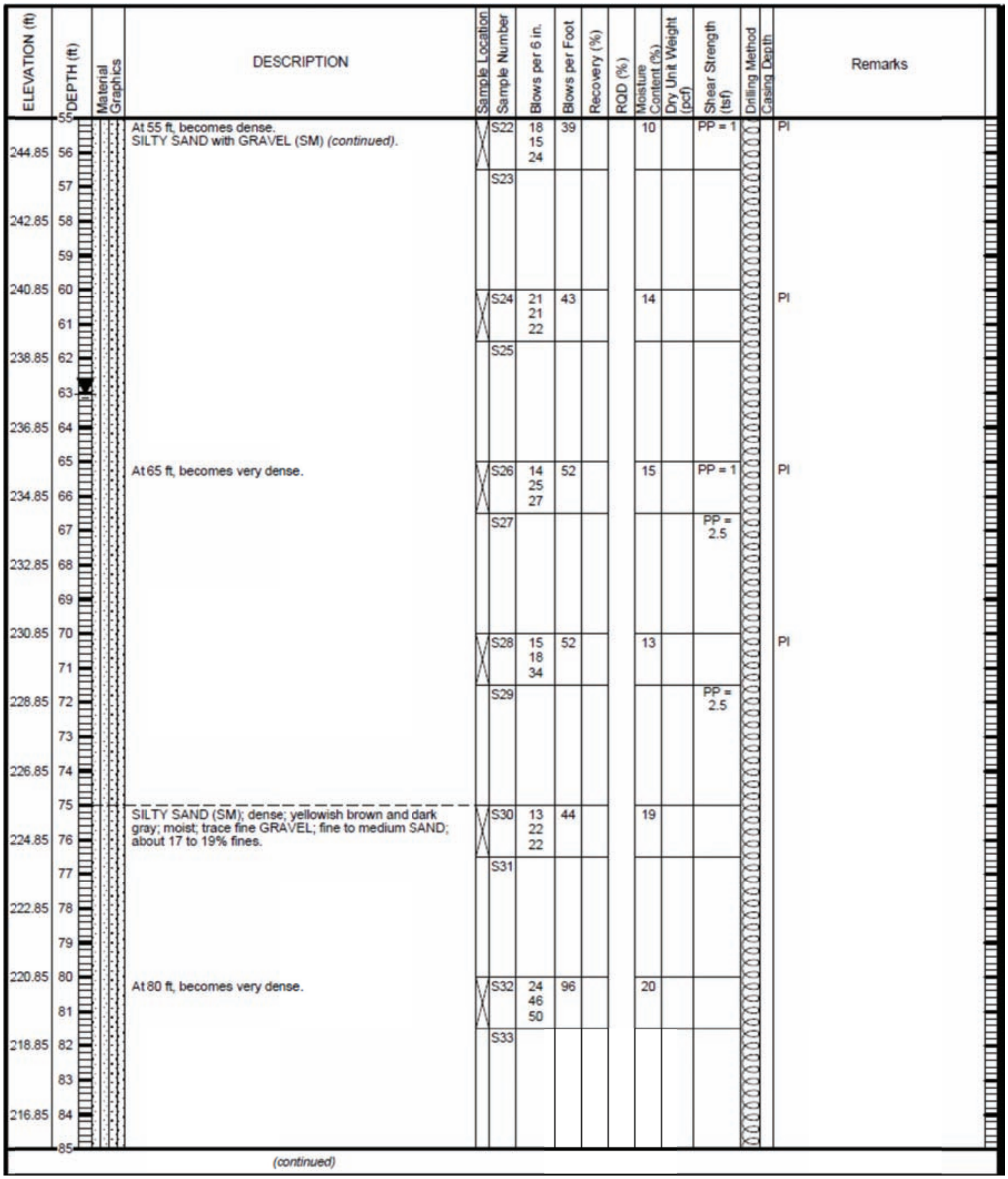


Figure 2-1. Geotechnical data sources include soil borings, field tests and laboratory tests. This data is used in various disciplines. Some of these disciplines are shown in the figure.

Figure 2-2. Picture shows a simple drilling rig conducting SPT. The SPT hammer can be seen in the picture.

(original picture from Towhata, 2008)



ELEVATION (ft)	DEPTH (ft)	Material Graphics	DESCRIPTION	Sample Location	Sample Number	Blows per 6 in.	Blows per Foot	Recovery (%)	RQD (%)	Moisture Content (%)	Dry Unit Weight (pcf)	Shear Strength (tsf)	Drilling Method	Casing Depth	Remarks
	55		At 55 ft, becomes dense. SILTY SAND with GRAVEL (SM) *(continued)*.		S22	18 15 24	39			10		PP = 1			PI
244.85	56														
	57				S23										
242.85	58														
	59														
240.85	60				S24	21 21 22	43			14					PI
	61														
238.85	62				S25										
	63														
236.85	64														
	65		At 65 ft, becomes very dense.		S26	14 25 27	52			15		PP = 1			PI
234.85	66														
	67				S27							PP = 2.5			
232.85	68														
	69														
230.85	70				S28	15 18 34	52			13					PI
	71														
228.85	72				S29							PP = 2.5			
	73														
226.85	74														
	75		SILTY SAND (SM); dense; yellowish brown and dark gray; moist; trace fine GRAVEL; fine to medium SAND; about 17 to 19% fines.		S30	13 22 22	44			19					
224.85	76														
	77				S31										
222.85	78														
	79														
220.85	80		At 80 ft, becomes very dense.		S32	24 46 50	96			20					
	81														
218.85	82				S33										
	83														
216.85	84														
	85														

(continued)

Figure 2-3. SPT test results (Blows per 6 inches and Blows per foot) in a sample borehole log.

Figure 2-4. Several CPT probes with various diameters. The standard test uses the 10 cm^2 or 15 cm^2 cone.

(original picture from Hunt, 2005)

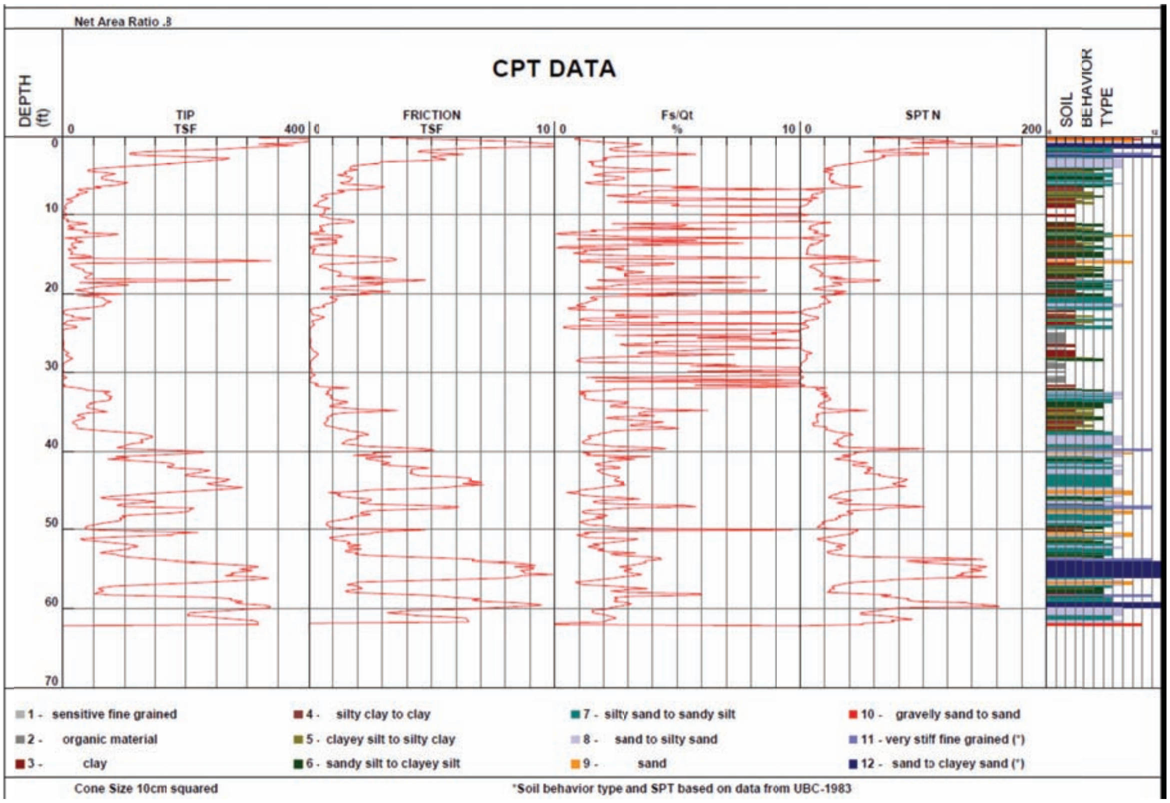


Figure 2-5. A typical CPT log showing field measurements (Side friction, Tip resistance, Friction ratio), as well as interpreted equivalent SPT blowcount (SPT N) and soil behavior type (SBT).

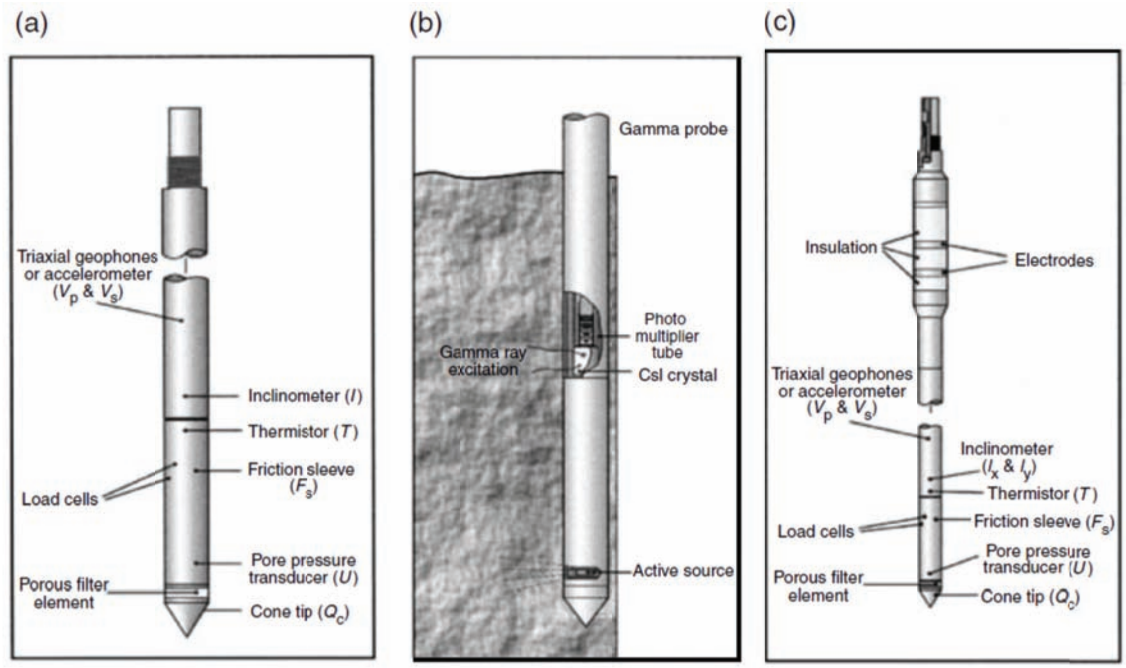


Figure 2-6. Several types of advanced cone penetrometers. (a) Piezo Cone Penetrometer (CPTU): Measures tip resistance, side friction, pore pressure, temperature, inclination and shear wave velocity. (b) Active Gamma Penetrometer (GCPT): Measures in-situ density. (c) Electrical Resistivity Cone (RCPT): Measures relative soil density.

(original figure from Hunt, 2005)

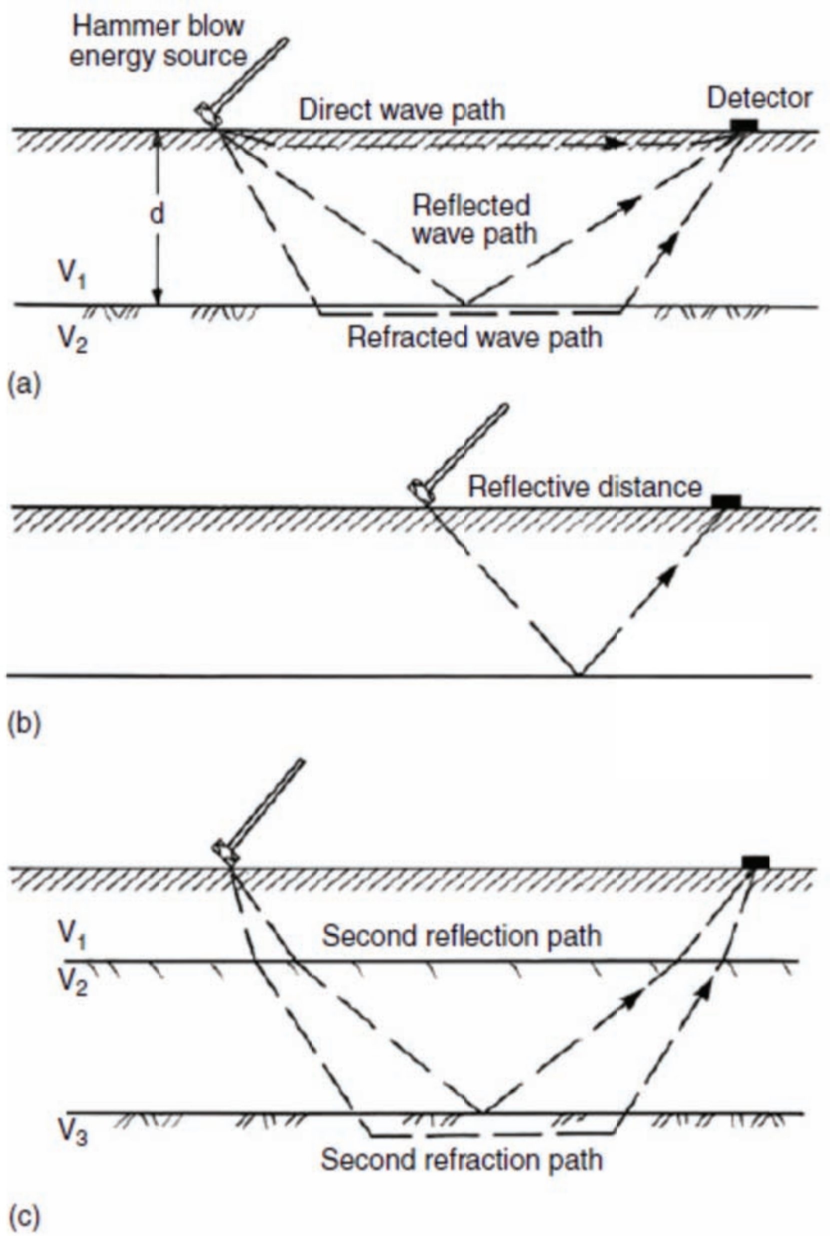
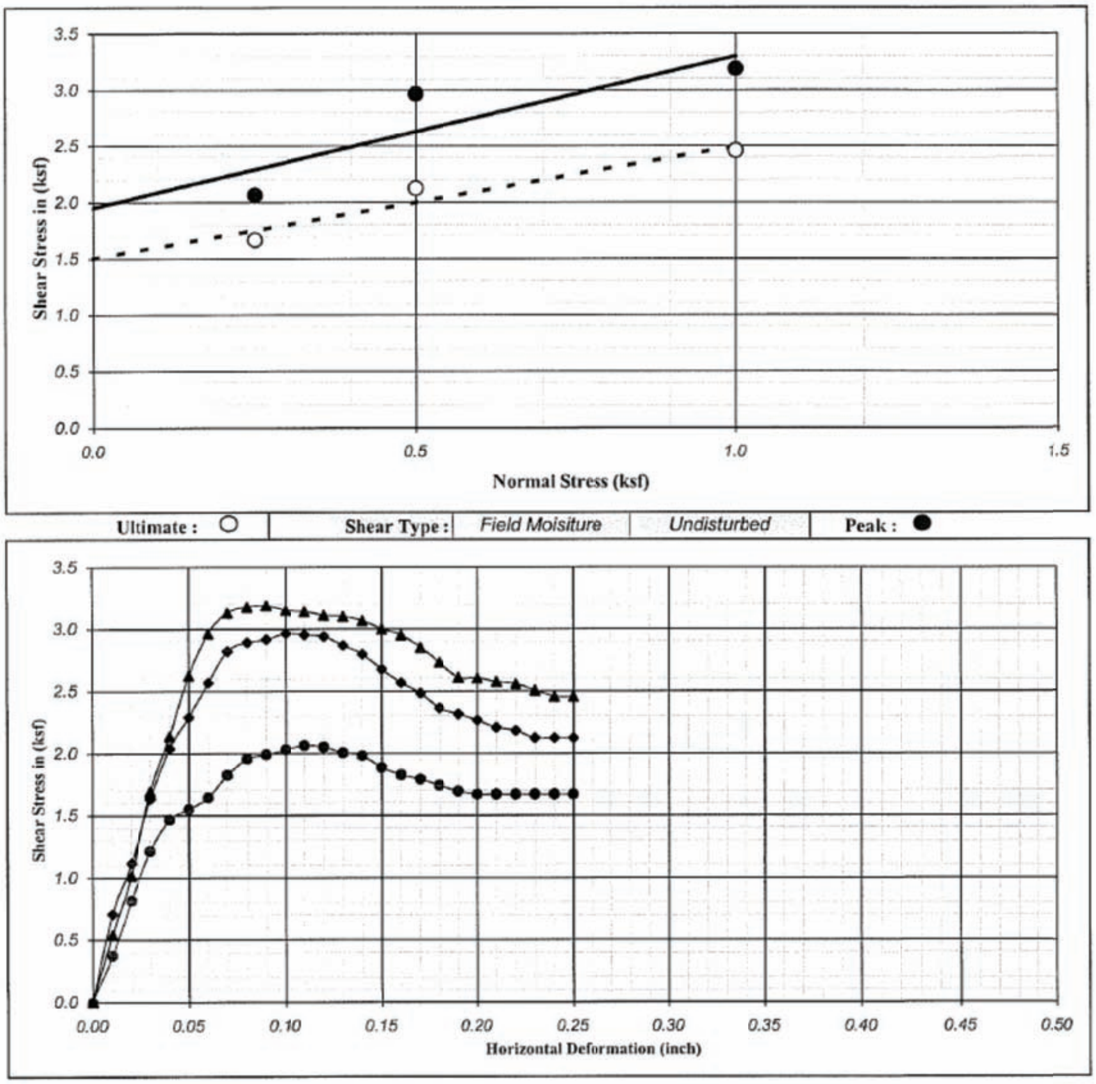


Figure 2-7. The basic concept of geophysical test methods is shown. Seismic direct, reflection and refraction waves through two- or multi-layer strata are shown.

(original figure from Hunt, 2005)



Boring No.	: R-09-032		Strength Intercept (C) :	1.95	(ksf)	Peak	1.50	(ksf)	Ultimate
Sample No.	: D-1			93.46	(kPa)		71.92	(kPa)	
Depth (ft/m)	: 5.0	0.00	Friction Angle (φ) :	53.44	Degree		44.87	Degree	
Description	: Dark brown, Lean CLAY with SAND (CL)					Shear Rate (inch/minute) :		0.02	

SYMBOL	MOISTURE CONTENT (%)	DRY DENSITY (pcf)	DRY DENSITY (kN/m³)	VOID RATIO	NORMAL STRESS (ksf)	NORMAL STRESS (kPa)	PEAK STRESS (ksf)	PEAK STRESS (kPa)	ULTIMATE STRESS (ksf)	ULTIMATE STRESS (kPa)
●	14.96	108.74	17.12	0.55	0.25	11.97	2.06	98.82	1.67	79.86
◆	14.55	107.49	16.92	0.57	0.50	23.94	2.96	141.92	2.12	101.70
▲	14.73	109.68	17.26	0.54	1.00	47.88	3.19	152.64	2.46	117.59

Figure 2-8. A laboratory data form for direct shear test. The lab data sheet may be archived in a spreadsheet, PDF or hardcopy format.

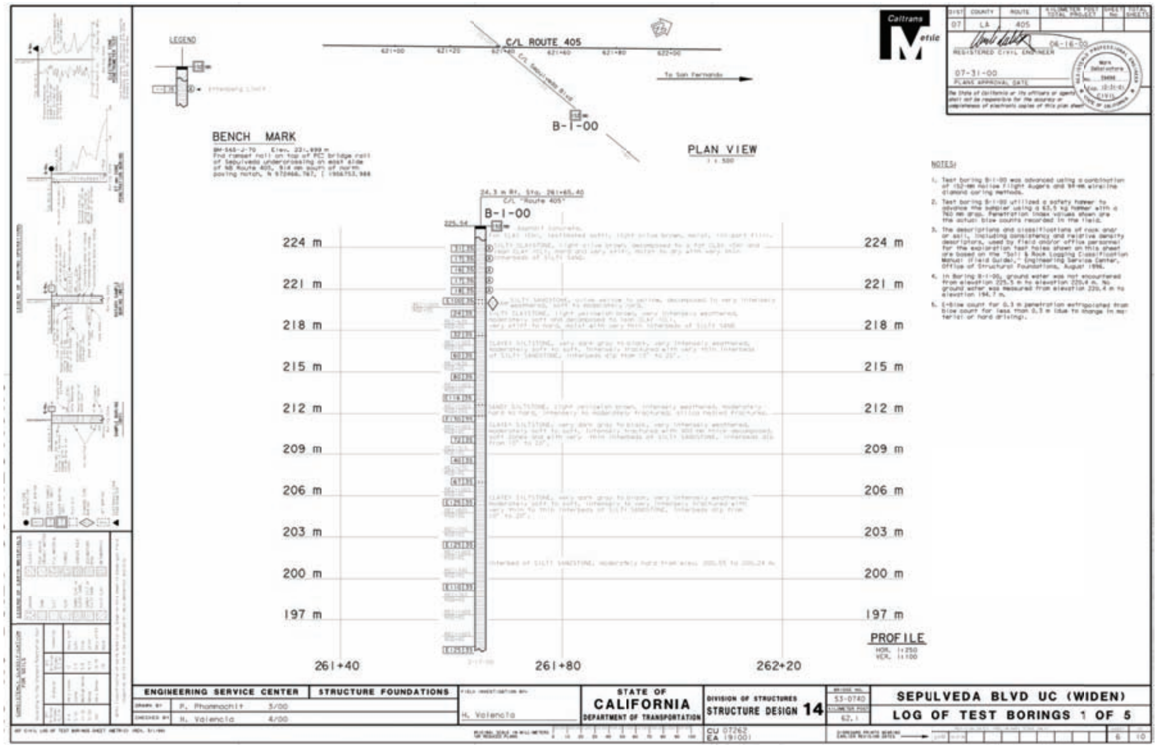
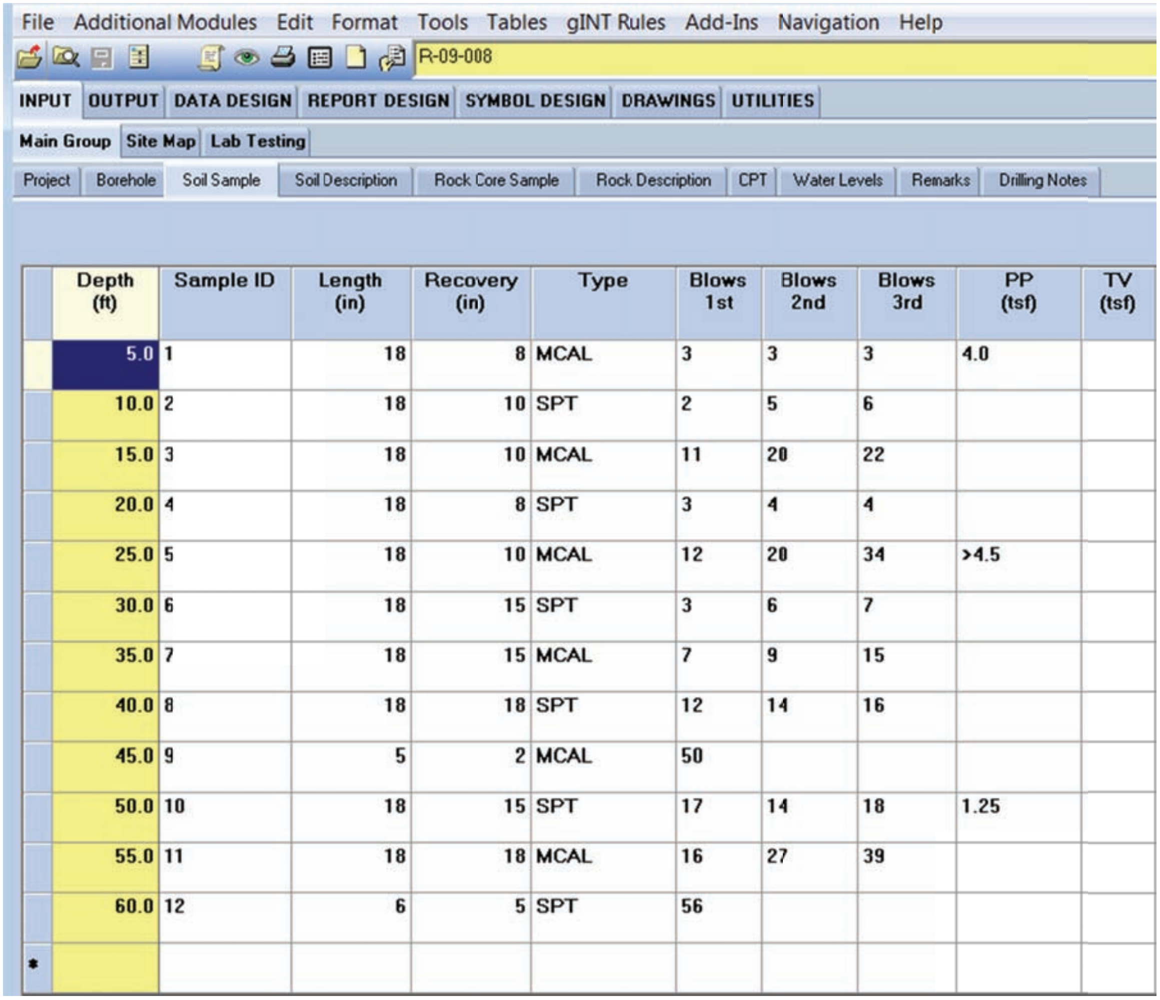


Figure 2-9. Sample LOTB downloaded from Caltrans BIRIS in PDF format. This database contains Caltrans bridge inspection records and as-built plans, including LOTBs.



Depth (ft)	Sample ID	Length (in)	Recovery (in)	Type	Blows 1st	Blows 2nd	Blows 3rd	PP (tsf)	TV (tsf)
5.0	1	18	8	MCAL	3	3	3	4.0	
10.0	2	18	10	SPT	2	5	6		
15.0	3	18	10	MCAL	11	20	22		
20.0	4	18	8	SPT	3	4	4		
25.0	5	18	10	MCAL	12	20	34	>4.5	
30.0	6	18	15	SPT	3	6	7		
35.0	7	18	15	MCAL	7	9	15		
40.0	8	18	18	SPT	12	14	16		
45.0	9	5	2	MCAL	50				
50.0	10	18	15	SPT	17	14	18	1.25	
55.0	11	18	18	MCAL	16	27	39		
60.0	12	6	5	SPT	56				

Figure 2-10. A typical gINT table for entering soil sample data and blowcounts. The samples are referenced through key data fields, e.g., Borehole ID (shown in the menu bar) and Depth.

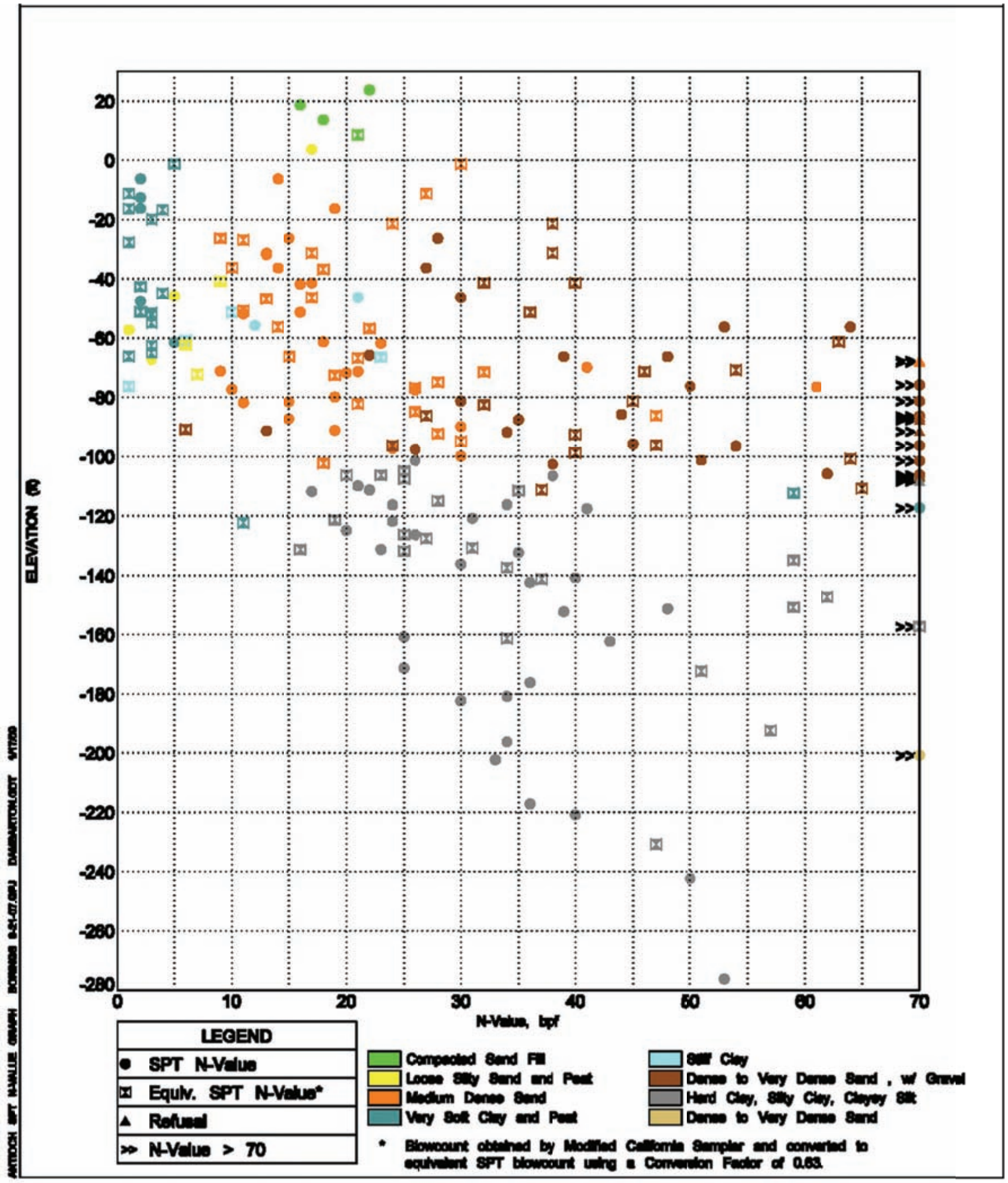


Figure 2-11. A typical gINT Report showing distribution of SPT blowcounts for different soil types from several boreholes.

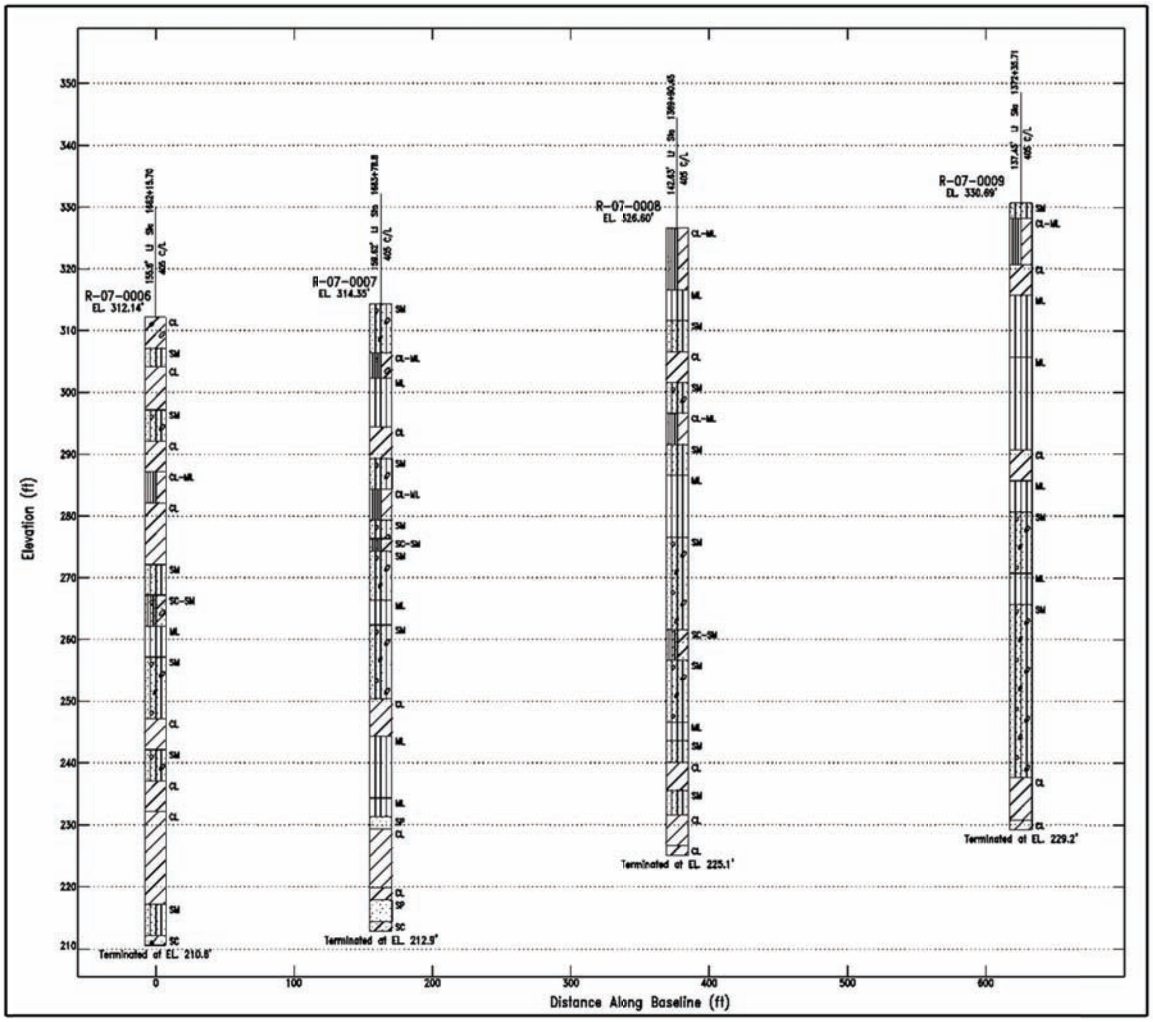


Figure 2-12. A fence diagram, generated by gINT along a road alignment. One advantage of well-formatted digital archiving in a database is that the data can be quickly converted to other forms.

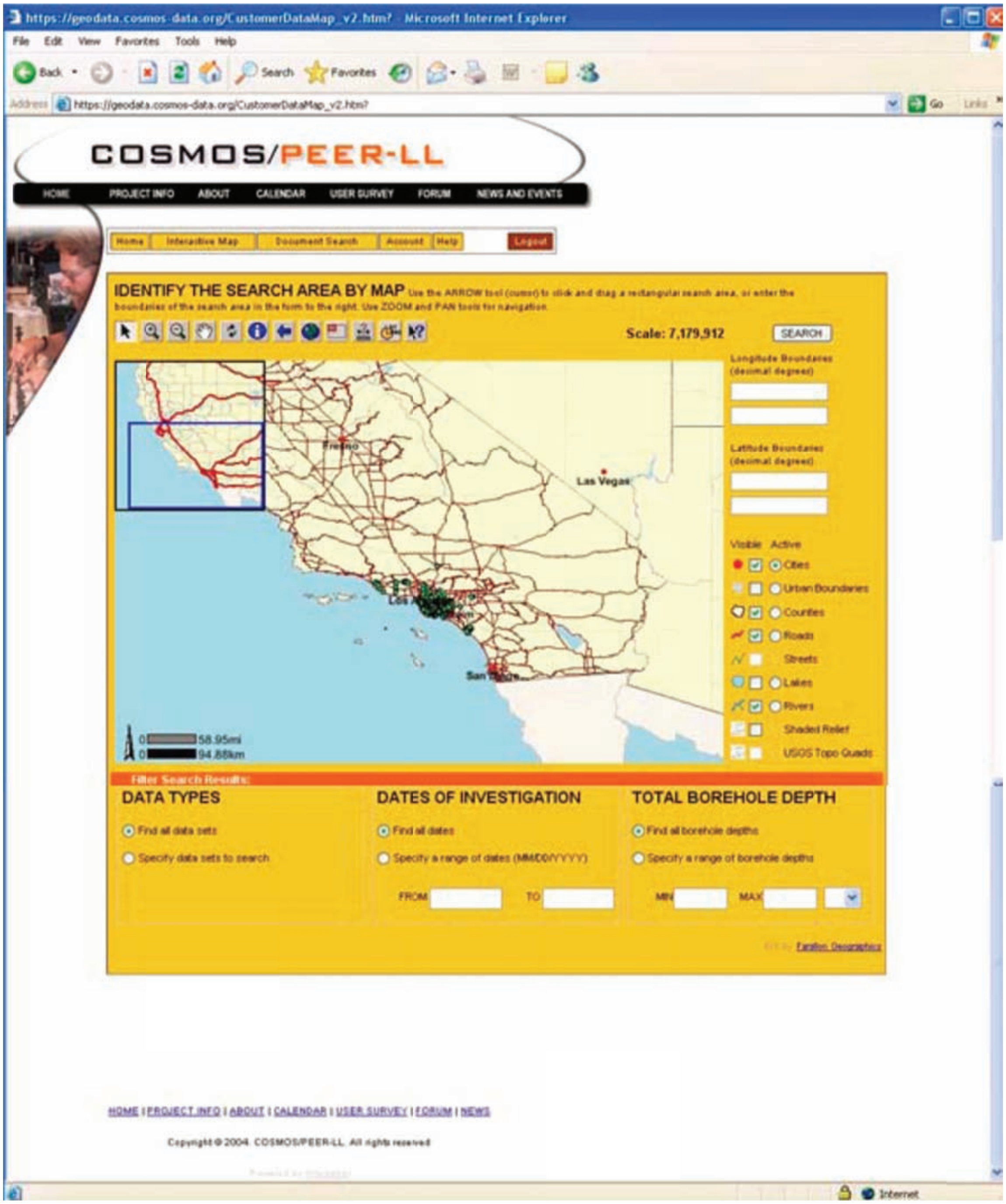


Figure 2-13. A screenshot of COSMOS/PEER-LL Geotechnical Virtual Data Center user interface.

(original figure from COSMOS/PEER-LL, 2004)

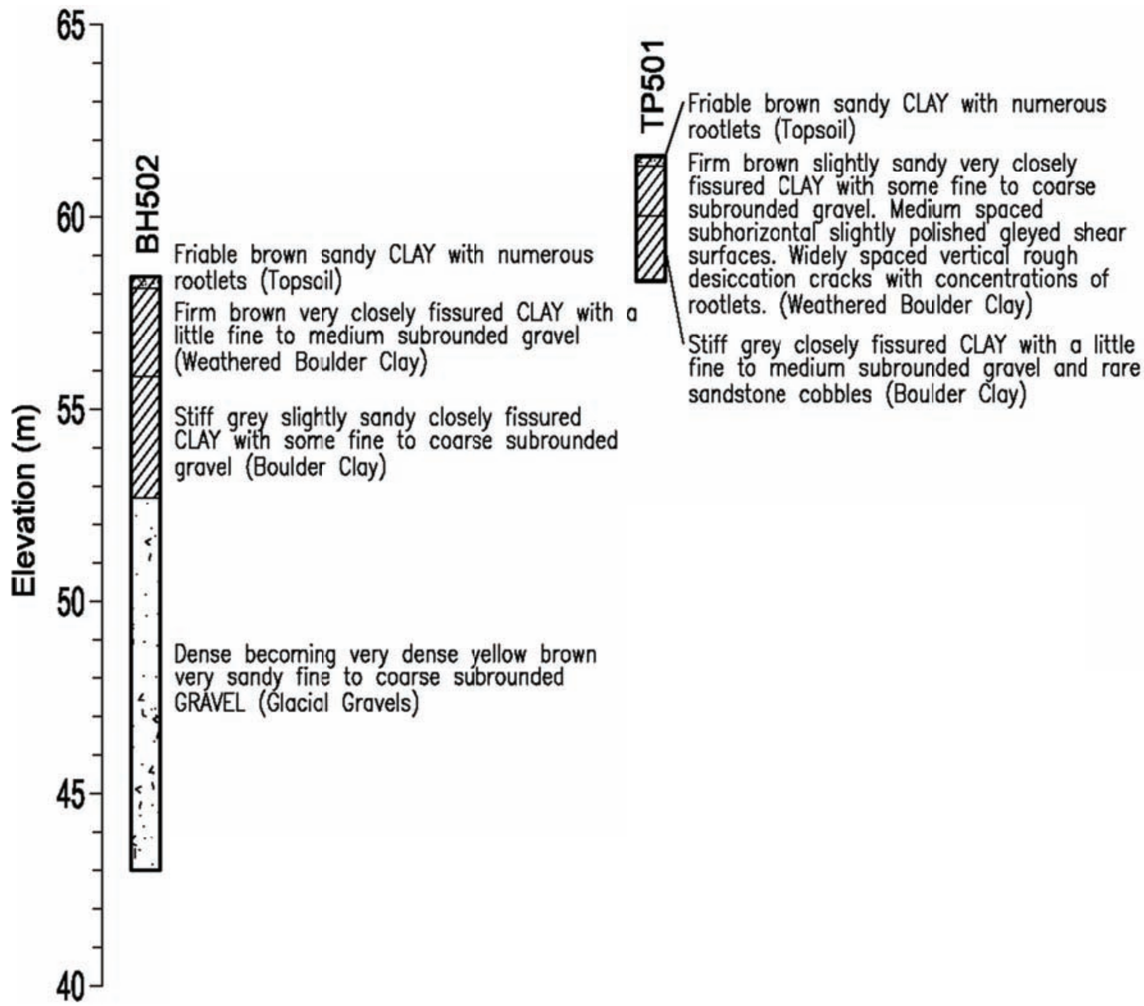


Figure 2-14. Graphical representation of AGS data transfer file shown in Table 2-11.

Chapter 3. Geospatial Data, Spatial Databases and Spatial Data Infrastructure

3.1 Definition of Geospatial Data

Geospatial data, or in short spatial data, are data pertinent to the location of a geographical entity. It is the information that defines the location of abstract or concrete entities or events with respect to the Earth's surface (OGC, 2008a and Blasby 2001). Spatial objects include both abstract data, such as boundaries and property lines, and concrete entities, such as roads and buildings.

Spatial data is a critical piece of information in modern societies. Different types of spatial data are used in high-level decision making in local, national and global levels. They are utilized in urban policy making and land development, business development, risk assessment and mitigation and disaster recovery planning, among other applications.

Spatial data are encountered in various engineering disciplines, particularly those describing geographic features (e.g., rivers, lakes and borders), distributed civil infrastructure networks (e.g., pipelines, bridges and roads) and 3D graphics and virtualization.

Spatial data comprise the backbone of modern geographical information systems, and are increasingly integrated in various disciplines in engineering, science and business.

3.2 Trends in Spatial Data Usage

The origins of GIS go back to the 1950s when computers were first used for mapping and cartography (Yeung and Hall, 2007). From the 1950s until the early 1990s, research and development in the spatial data field were mainly in spatial data management, mapping and cartography, and spatial data analysis.

Until the 1990s, spatial data and GISs were specialized fields, used only by GIS professionals and researchers; therefore, scientists and engineers from other disciplines and the general public had little interest in this field. Starting in the 1990s, the general public gained access to the Internet, and government agencies and private entities started distributing data via the Internet to increase efficiency of their operations and spatial data and GIS became widely available for commercial and non-commercial purposes. Today spatial data and GIS are used for numerous applications including multiple listing services, traffic and direction maps, service finders, etc. In addition, the majority of science and engineering disciplines use spatial data and GIS as powerful tools for improving performance and efficiency of their respective workflows.

3.3 Spatial Relationships and Operations

Spatial data are related through spatial relationships, e.g., adjacency, connectivity and containment. Table 3-1 shows some examples of spatial relationships between two spatial objects.

Spatial operations are used to examine spatial relationships between different spatial objects. Table 3-2 shows the spatial operations on geometry class as defined by OGC (Yeung et al., 2007).

3.4 Spatial Databases

In the past, spatial data were stored in relational databases using regular (non-spatial) data types for coordinates, coordinate reference systems and dimensions. Modern spatial data applications, however, use huge amounts of geometric objects, temporal and regular data, which cannot be handled with traditional systems in an efficient manner. To solve this problem in modern spatial data applications these data are stored in spatial databases using spatial data types. These database systems are usually based on regular commercial database systems but with additional functions and capabilities to handle spatial data types.

The main additions of spatial database systems to regular databases are as follows (Yeung and Hall, 2007):

- Spatial data types: in a spatial database spatial data are stored either as a spatial data type, as defined by Open Geospatial Consortium (OGC, 1999, 2010), or a binary large object (BLOB). A spatial database provides functionality for handling the spatial data types, whereas similar operations on a BLOB requires additional software.
- Spatial indexing: Spatial databases use spatial object coordinates as an indexing mechanism. This permits queries to be performed based on feature location rather value of certain database fields.
- Spatial operators or functions: spatial databases provide spatial functions that perform certain operations or queries on spatial objects. These functions can be invoked using SQL statements. OGC (1999, 2010) provides standards for these SQL queries.

- Spatial application routines: these set of applications support common tasks that need to be performed on a spatial database. Some examples include spatial data loading, long transaction control.

In spatial databases, spatial operations can be performed using SQL statements. For example, commercial spatial databases provide SQL functions that calculate the distance between two geometries, examine whether a particular geometry is within specified distance from others, and check whether two geometries intersect each other. If the data is stored in a non-spatial database, such operations need to be handled by client applications, rather than the database itself.

When used for storing spatial data, spatial databases have certain advantages over regular databases: they organize spatial data in a more logical and consistent manner; they index spatial data using location, rather than name or other properties, resulting in faster queries and calculations; they harbor data with more integrity and less redundancy; they transfer the implementation and enforcement of spatial relationships to the database; and, they facilitate the development of client applications by providing advanced spatial features such as spatial querying and distance calculation (Shekhar and Chawla, 2003).

Oracle is the first database vendor to add spatial support to its commercial database systems, Oracle Spatial (Oracle). A large number of database systems now support spatial data, including DB2 Spatial Extender (IBM), PostGIS (PostgrSQL), ESRI's geodatabase and Informix Spatial DataBlade, which was acquired by IBM in 2001.

3.4.1 Architecture of a Spatial Database Management System

Spatial database management systems (DBMS) architecture and components are similar to general DBMS systems. Figure 3-1 shows a simplified diagram of a DBMS architecture and its main components.

The mapping between the data set and database is established through a data model and the database schema. The data model is a conceptual description of the data and relationships that need to be stored in the database; however, it does not directly translate to the data tables and columns in the database. That task is performed using a database schema, which describes the physical layout of the database tables and relationships.

The database engine is a set of programs that manipulates the stored data and interacts with other components, e.g., user-interface and database development and operation tools. In addition to actual data files reside a data dictionary, a set of stored procedures and integrity rules inside the database. Data dictionary describes the content of database to clients. Stored procedures are a set of Structured Query Language (SQL) statements that define, manage and query the data in the database (Yeung and Hall, 2007). Integrity rules ensure the data integrity and protect the data from corruption due to illegitimate queries.

3.4.2 Classification of Database Management Systems

Database management systems can be classified in terms of their data models, database functions, type of stored data, objectives of the system and hardware platform and configuration. Table 3-3 shows several DBMS classifications based on these criteria (Yeung and Hall, 2007).

In regard to data-model classification, the hierarchical and network databases represent the earlier generation of DBMS and are rarely used today. Nowadays almost all DBMS are relational or object-oriented, or a combination of these two types (object-relational). A relational database system stores the data in a set of tables, which are related to each other using a set of relations, hence a relational database. The main advantage of this method in comparison to the older generation (hierarchical and network DBMS) is that the end user does not need to be aware of the physical implementation of the database.

Object-oriented DBMS database systems are based upon object-oriented programming principles. In this concept the data are observed as a set of identifiable objects, which are categorized in classes based on similar properties.

Another classification method is based on main database function. Data storage or inventory systems mainly serve large quantity of static data. Transaction control systems deal with large numbers of users processing concurrent transactions, which mainly deal with dynamic data. Decision support databases provide data for support of the decision-making process.

In terms of data types database systems can be classified as spatial or non-spatial, as discussed before.

Another important classification is based on the scope of the database. Custodial systems or data warehouses serve the data for a variety of clients within an extended period of time, while project-oriented systems are designed for a specific client and often for a short-term application.

In terms of hardware configuration database systems are either desktop-based or distributed. Desktop-based systems are mainly for smaller databases with a limited number of clients in the same physical location, whereas distributed systems can serve a large number of clients at multiple locations.

3.5 Spatial vs. Pseudo-Spatial Data

There are data forms that have spatial characteristics but cannot be directly used as spatial data, because they are not georeferenced. These data are called pseudo-spatial data. They can be converted to fully functional spatial data by adding metadata that enable users to locate the geographical location where data are pertinent, e.g., georeferencing.

Figure 3-2 shows various spatial data types and the process of converting pseudo-spatial data to spatial data through georeferencing.

A large part of legacy geotechnical data is of this type. In other words, these data should be georeferenced for use as spatial data in a spatial database. While this certainly can be done, the process is time consuming and requires considerable effort. More recent data in databases and Web-based data management systems are typically georeferenced.

3.6 Open Geospatial Consortium Geometry Object Model

The Open Geospatial Consortium (OGC) introduced a formal definition for a geometry object in GIS Simple Feature Specification for SQL (OGC, 1999). Today this definition is universally used in spatial data applications. Figure 3-3 shows the schematic view of a geometric object based on the OGC definition. Based on this definition, a geometry object is an object that has at least one attribute of geometric type.

3.7 Spatial Data Infrastructures (SDIs)

Spatial data infrastructure is a collection of spatial data, metadata, tools and users that enables acquiring, processing, distributing and using spatial data.

In its April 2006 newsletter, the Global Spatial Data Infrastructure Association (GSDI) defines the Spatial Data Infrastructure (SDI) as follows (GSDI, 2006):

“Spatial Data Infrastructures provide a basis for spatial data discovery, evaluation and application and include the following elements:

- Geographic data: the actual digital geographic data and information.
- Metadata: the data describing the data (content, quality, condition and other characteristics). It permits structured searches and comparison of data in different clearinghouses and gives the user adequate information to find data and use it in an appropriate context.
- Framework: includes base layers, which will probably differ from location to location. It also includes mechanisms for identifying, describing and sharing the data using features, attributes, attribute values, as well as mechanisms for updating the data without complete re-collection.
- Services: to help discover and interact with data.

- Clearinghouse: to actually obtain the data. Clearinghouses support uniform, distributed search through a single-user interface; they allow the user to obtain data directly, or they direct the user to another source.
- Standards: created and accepted at local, national and global levels.
- Partnerships: the glue that holds it together. Partnerships reduce duplication and the cost of collection and leverage local/national/global technology and skills.
- Education and communication: allowing individual citizens, scientists, administrators, private companies, government agencies, non-government organizations and academic institutions with local to global interests to communicate with and learn from each other.

The first generation of SDIs was more focused on databases and software. They were mainly developed by public-sector organizations and government agencies. The second generation of SDIs is more focused on involvement from the private sector and client interaction (Craglia and Annoni, 2007).

3.8 Spatial Data Infrastructure Standards

The main driving force behind SDIs is that usage of common agreements and technical specifications and elimination of parallel processes reduce the cost, the development time and the manpower required for acquiring, exchange and utilizing of geospatial data. The concept also ensures interoperability between different SDIs. This goal is achieved through adoption of common standards, tools and algorithms for various types of geospatial data. These efforts are led by several industry-supported entities, including the Open Geospatial Consortium (OGC) and the Global Spatial Data Infrastructure Association (GSDI). These organizations are in charge of developing and adopting technical specifications for use in SDIs.

Although the concept of SDIs implies standardization and interoperability, in practice the majority of current SDIs act as independent applications with minimal interoperability. Each agency or government has a set of standards and best practices for their applications, but in reality little global interoperability has been realized. In order to achieve a higher level of interoperability, not only standards in use but also their current versions and future progress paths should be well defined.

The SDI 1.0 standards suite (Nebert et al., 2007) is the first proposal with this objective in mind. It is basically a collection of standards currently used in different SDIs. The U.S. National Geospatial Intelligence Agency (NGA) approved the SDI 1.0 baseline specifications in 2005 (NGA, 2005). Table 3-4 lists the formal and tentative international standards currently used or proposed for use in SDIs (GSDI, 2010). The SDI 1.0 proposal selects some of these standards as core standards and others as supplemental or future standards. Table 3-5 lists the proposed core, supplemental and future standards in the SDI 1.0 proposal. In the remainder of this chapter some of the key standards from this table are reviewed.

The list of standards in Table 3-5 indicates that most of the standards are defined by OGC. OGC is a consortium by companies, government agencies and universities that work together to develop industry standards for geospatial data and services. The goal of the participating members is to enable complex geospatial data structures and services for various applications that can interoperate using standardized interfaces. OGC was established in 1994 by eight charter members: Camber Corporation, University of Arkansas's Center for Advanced Spatial Technologies, Center for Environmental Design Research at the University of California in Berkeley, Intergraph Corporation, PCI Remote Sensing, QUBA, USACERL (US Army Construction Engineering Research Laboratory), and USDA Soil Conservation Service. Currently more than 400 organizations are members of OGC.

OGC members include various industries. Current working groups include 3-D information management group, architecture-engineering-construction, information models (BIM), defense and intelligence, disaster management, emergency response, environmental and natural resources, geospatial rights management, homeland security, mass-market geospatial, sensor Web enablement and universities and research.

OGC is responsible for multiple standards currently used by the spatial data industry. In the following sections some of these standards are discussed.

3.8.1 Geography Markup Language

Geography Markup Language (GML) is a markup language developed by OGC (OGC 2008a, Galdos Systems 2003) to encode geographical information, including spatial and non-spatial data. Since GML is based upon standards defined by World Wide Web Consortium (W3C), GML data can be easily distributed over the Internet. GML conforms to the eXtensible Markup Language (XML) standards, which implies that GML users can take advantage of technologies available for XML data processing, such as XSLT, XLINK and XQUERY, to generate, query or modify GML data.

The first version of the specifications (GML 1.0) was adopted by OGC in April 2000. Other previous version include GML 2.0 (February 2001), GML 2.1 (January 2002), GML 3.0 (January 2003), and GML 3.1 (February 2004). The current version (Joint OGC/ISO TC 211 Adopted Standard) is GML 3.2 (ISO 19136), which was adopted in September 2007.

GML is adopted by different disciplines as the standard language for data modeling. Table 3-6 shows the current domain-specific standards or initiatives, which are based on the GML.

3.8.2 GML Structure

GML is a set of XML schemas that defines a framework for encoding different aspects of geographical data, such as geographic coordinate systems, data and geometric objects such as points and lines. These schemas constitute the core GML model, which builds upon an object-property model. The data are presented as GML objects with properties characterizing various aspects of features. Throughout the following discussions in this dissertation, the *italic* font is used to distinguish GML elements.

Domain experts can build upon GML core schemas and design GML-compatible application schemas for specific domains. They can define domain-specific *features* using the customary vocabulary of their field of expertise. Each *feature* is a global XML element, which is an extension of the base *gml:featureType*, and is further described by child property elements that might be geometric or non-geometric. Property values in GML can be defined either inline or remotely. An inline property definition is embedded within the parent element, while a remote property is either defined elsewhere in the same document or through a valid URI (Unified Resource Identifier) and is referred to using an xlink attribute.

The root element of the document in GML3 is a *FeatureCollection*, which acts as a container for various *features* or other *FeatureCollections*. Similar to the basic XML structure, the elements can be nested in several levels. GML *features* may not be direct child elements of other GML *features*; their relationship must be established through *featureProperty* elements.

All GML documents have a namespace that acts as an identifier without necessarily pointing to a physical target. The unique namespace averts potential conflicts between elements names and facilitates the inspection and comprehension of GML schemas. For example, base GML schema defines a *Point* in *gml* namespace (e.g., *gml:Point*). If a *Point* feature with different properties has to be used in a schema, it can be defined in a different namespace (e.g. *app:Point*) to avoid ambiguity between core GML *point* and user-defined *point* elements.

Properties are used to define associations or relationships between GML objects. The property name usually indicates the role of the target object in the source object, or vice versa. A relationship name that explains the role of the target object in source object is preferred over describing the contents of the target object.

The first major release of GML, i.e., GML2, supported simple features restricted to two-dimensional geometries (e.g., *Point*, *LineString*, *LinearRing*, *Box* and *Polygon*), as well as aggregate geometries (e.g., *MultiPoint*, *MultiLineString* and *Multi-Polygon*). The second major release (GML3) supports two- and three-dimensional geometries, including geometry types for points, curves, surfaces and solids. Some of the new geometry types are *Arc*, *Circle*, *CubicSpline*, *Ring*, *OrientableCurve*, *OrientableSurface* and *Solid*. In addition to new geometry types, GML3 is more compliant with ISO/TC 211 family of specifications and contains additional features.

3.8.3 Web_ Feature Service Standard

Web Feature Service (WFS) is an OGC standard for Web services providing a set of standard operations on geographic features. WFS is a subset of more general OGC Web Service Specifications (OWS). The WFS standard allows the users to exchange the geographic features directly, rather than exchanging files containing these data. The idea behind this standard is to make the geographic features available for multiple purposes, including applications for which data are not originally produced.

WFS standard supports operations such as INSERT, UPDATE, DELETE, LOCK, QUERY and DISCOVERY on geographic features using HTTP protocol. The WFS 2.0 standard supports 12 operations as follows:

- GetCapabilities (discovery operation)
- DescribeFeatureType (discovery operation)
- GetPropertyValue (query operation)
- GetFeature (query operation)
- GetFeatureWithLock (query and locking operation)
- LockFeature (locking operation)
- Transaction (transaction operation)
- CreateStoredQuery (stored query operation)
- DropStoredQuery (stored query operation)
- ListStoredQueries (stored query operation)
- DescribeStoredQueries (stored query operation)

One of the basic requirements of WFS is that features need to be encoded in GML, although the service may support non-GML encodings in addition to GML. Each feature in a WFS should be identified by a unique resource identifier, which is assigned by the server when the feature is added to the service. This resource identifier is permanent and cannot be reused once the feature is deleted. For GML features the resource identifier is encoded as attribute *gml:id*.

3.8.4 Examples of SDIs

The most extensive examples of SDIs are the national spatial data programs developed by the government. For example, in the United States, the federal government has established the National Spatial Data Infrastructure (NSDI) program, which provides means to share geographic data among users for data collection, use and decision making. According to the NSDI Website (2011) the goal of this program is to “reduce duplication of effort among agencies, improve quality and reduce costs related to geographic information, to make geographic data more accessible to the public, to increase the benefits of using available data, and to establish key partnerships with states, counties, cities, tribal nations, academia and the private sector to increase data availability”.

Another SDI initiative is INSPIRE program in the European Union (Craglia and Annoni, 2007). INSPIRE program is a second-generation SDI based on existing resources and encourages stakeholders’ involvement through organizing them in interest groups. This initiative combines a large spectrum of projects under an umbrella in order to support the research, management and evaluation of European Union environmental policies. The major components of INSIPRE program are summarized as follows by Craglia and Annoni (2007):

- Metadata
- Key spatial data themes and services
- Network services and technologies
- Agreements of sharing and access
- Coordination and monitoring mechanisms
- Process and procedures

The tally of involved entities as of April 29, 2005, was as follows (Craglia and Annoni, 2007):

- Spatial data interest communities: 133
- Legally mandated organizations: 82
- Proposed experts: 180
- Referenced materials: 90
- Identified projects: 91

Some of the agencies and spatial data interest communities involved in INSPIRE program are listed in Table 3-7. This list provides an overview of the type of entities that are involved in this SDI on regional, national and international levels.

3.9 Integration of Geotechnical Data in SDIs

SDIs support geospatial data discovery, access and use in decision-making processes, and can make major contributions to economic development and quality of life of communities. The main motivation for developing SDIs is use of geospatial data in local, regional and global decision making in areas such as earthquake hazard estimation, business development, flood mitigation, environmental restoration, community land use assessment and disaster recovery. Geotechnical data is valuable in some of these decision-making processes. It can be used in analyses such as earthquake hazard estimation, land use and business risk assessment. Therefore, there is a clear demand for integrating the geotechnical data in SDIs used for these processes.

In order to a include geotechnical information in an SDI the spatial reference of data shall be established. As discussed in Chapter 2 geotechnical data are generally obtained from either boreholes in the field, or samples taken from boreholes and tested in the laboratory. It is comprised of different pieces of information that are referenced relative to the top of the borehole. Well-referenced geotechnical data have sufficient geographical attributes to define the location of observations, measurements and samples relative to the borehole top. If the borehole top is spatially referenced, other geometric features (e.g., samples and specimens) are also retroactively referenced. These spatial features provide a framework for the non-geometric geotechnical values that characterize the measured/observed physical properties of the soil.

3.10 Geotechnical Data Model for SDIs

A successful implementation of a spatial data model depends on using a compatible set of standards combined with data model schemas that ensure interoperability and data integrity. Having these characteristics will facilitate implementation and application of the data model for both data providers and clients. Meeting these criteria will improve the functionality and efficiency of the data model and ensure interoperability with other components of a SDI.

Table 3-8 lists some of the requirements of a spatial data model for geotechnical data. The criteria discussed in this table ensure data integrity, interoperability and efficiency in development and usage of the data model.

In order to satisfy these criteria, the data model adopted for the SDI should be compatible with the key pertinent standards used in other SDIs, as show in Table 3-5. In terms of data model this implies that a GML compatible data model should be used. Also, since the data will be distributed mainly as features, the WFS standard compatibility is recommended.

3.11 Comparison between SDI User Community and Geotechnical Engineering Community

One of the characteristics of SDIs is that their data are used by multiple users that are not involved in data acquiring and processing, generally. This user base has different requirements in comparison to the traditional geotechnical community. The geotechnical community has more interest in a project-driven data format, while the SDI community is more interested in a standardized data format that can be used in the same way other data are used. The geotechnical community has more knowledge on the data and the drilling and testing procedures, but they are less sophisticated in terms of using GML, Web services and spatial databases.

Moreover, the geotechnical community generally relies on proprietary geotechnical applications such as gINT and Rockware to process and visualize the data. Therefore, the data format needs to be easily importable to these applications. In comparison the SDI community mainly relies on standard GIS-based applications for visualization and sophisticated geostatistical analysis. Therefore, it makes more sense to use a data format that is similar to what is used by other pieces of data in SDIs.

Based on this background it appears that despite complexity and the required learning curve for new users, a GML compatible data format is appropriate for use in SDI. The clients using these services are sophisticated users who are already exposed to these technologies through other types of data in SDIs.

As a data exchange format for the geotechnical community, however, using GML and related technologies appears to be premature at this time and a simpler data format with better compatibility to existing geotechnical applications, similar to the geotechnical data format proposed by Mokarram (2010), is more appropriate. Those data formats can be used for data exchange between laboratory and engineers, between different companies and also as a submittal format to agencies. Using simple features it can be easily adapted to relational data format which is used by the database programs such as gINT, and also through XSLT translations it can be converted to a more sophisticated GML format comprising of complex features for use in applications schemas.

Table 3-1. Spatial relationships between different spatial objects.

Spatial Relationships Type	Example
Adjacency	Two project sites adjacent to each other
Connectivity	Two road alignments connected to each other
Containment	Borings within a project site

Table 3-2. OGC spatial operators defined on the class Geometry.

Classes	Operators	Operator Functions
Basic Operators	Spatial Reference	Returns the reference system of the geometry
	Envelope	Returns the minimum bounding rectangle of the geometry
	Export	Converts the geometry into a different representation
	IsEmpty	Tests if the geometry is the empty set or not
	IsSimple	Returns TRUE if the geometry is simple
	Boundary	Returns the boundary of the geometry
Topological Operators	Equal	Tests if the geometries are spatially equal
	Disjoint	Tests if the geometries are disjoint
	Intersect	Tests if the geometries intersect
	Touch	Tests if the geometries touch each other
	Cross	Tests if the geometries cross each other
	Within	Tests if a geometry is within another geometry
	Contain	Tests if a given geometry contains another geometry
	Overlap	Tests if a given geometry overlaps another given geometry
	Relate	Returns TRUE if the spatial relationship specified by the 9-Intersection matrix holds
Spatial Analysis Operators	Distance	Returns the shortest distance between any two points of two given geometries
	Buffer	Returns a geometry that represents all points whose distance from the given geometry is less than or equal to a specified distance
	ConvexHull	Returns the convex hull of a given geometry
	Intersection	Returns the intersection of two geometries
	Union	Returns the union of two geometries
	Difference	Returns the difference of two geometries
	SymDifference	Returns the symmetric difference (i.e., the logical XOR) of two geometries

(after Yeung and Hall, 2007)

Table 3-3. Classification of database management systems based on different criteria.

Classification Criteria	Database Categories
Data Model	• Hierarchical Systems • Network Systems • Relational Systems • Object-oriented Systems • Object-relational Systems
Primary Database Functions	• Data Storage or Inventory Systems • Transaction Systems • Decision Support Systems
Nature of Data	• Spatial Information Systems • Non-spatial Information Systems
Objectives of Information	• Custodial Systems and Data Warehouses • Project-oriented Systems
Hardware Platforms and System Configurations	• Distributed Systems • Desktop Systems

(after Yeung and Hall, 2007)

Table 3-4. Standards currently used or proposed for use in SDIs.

Standard	Canada CGDI	U.S. NSDI	GDI NRW	Catalonia
Formal	x	x	x	x
OGC Web Map Service	x	x	x	x
OGC Web Feature Service	x	x	x	
OGC Filter Encoding	x		x	
OGC Style Layer Descriptor	x	x	x	x
OGC Geography Markup Language	x	x		x
OGC Web Map Context	x	x		
OGC Catalogue Service 2.0 Z39.50 protocol binding	x	x		
FGDC Content Standard for Digital Geospatial Metadata	x	x		
OGC Web Coverage Service	x	x		x
OGC Catalogue Service 2.0 HTTP protocol binding (CS-W)	x	x	x	
Tentative				
OGC Web Coordinate Transformation Service				
OGC Gazetteer Profile of WFS	x	x		
OGC Web Pricing and Ordering Service			x	
ISO Metadata DTS 19139			x	
OGC Web Processing Service	x			

(after GSDI, 2010)

Table 3-5. Proposed SDI 1.0 Core, Supplemental and Future standards.

SDI 1.0 Core Standards
OGC Web Map Service 1.1.1
OGC Web Feature Service 1.0
OGC Filter Encoding 1.1
OGC Web Coverage Service 1.0
OGC Geography Markup Language 2.1.2
OGC Catalogue Service 2.0 Z39.50 protocol binding
FGDC Content Standard for Digital Geospatial Metadata (CSDGM, 1998)
SDI 1.0 Supplemental Standards
ISO Metadata standard 19115:2003 and ISO DTS 19139:2006
OGC Geography Markup Language 3.1.1
OGC Style Layer Descriptor 1.0
OGC Web Map Context 1.1
OGC Catalogue Service 2.0 HTTP protocol binding (CS-W)
Future Candidate SDI Core Standards
OGC Web Map Service 1.3
OGC Web Feature Service 1.1
OGC GML 3.1.1
OGC Catalogue Service 2.0 HTTP protocol binding, CS-W
ISO DTS 19139:2006 metadata

(from Nebert et al., 2007)

Table 3-6. A list of some of the domain-specific standards/initiatives, which are based on GML specifications.

Standard Name	Domain/Description
AIXM	Aeronautical Information
WXXM	Weather Information
WFS	Web Feature Service
WCS	Web Coverage Service
FPS	Feature Portrayal Service
CSW-ebRIM	Catalog/Registry Service
CityGML	City Planning
O&M	Observation and Measurement
CSML	Climate Science
DIGGS	Geotechnical/Geo-Environmental
GeoSciML	Geo Science
TransXML	Transportation - US
LandGML	Engineering/Construction

(after Burggraf, 2010)

Table 3-7. Some of the organizations and projects involved in the INSPIRE European SDI initiative (Craglia and Annoni, 2007).

Organization/Project Name	Description
ESBN	European soil bureau network, the network of national soil science institutions in Europe
EIONET	European environment information and observation network
EMI	European meteorological infrastructure
GDI NRW	North Rhine-Westphalia, Germany, a spatial data interest community of public/private geoinformation providers and users, including land-surveying office, geological survey, ministry of environment, etc.

Table 3-8. Requirements of a spatially-enabled data model for geotechnical applications.

Number	Question	Comments
1	Is it comprehensive and does it include every common investigation method, in-situ test and laboratory test?	Data model should include every investigation method and in-situ and laboratory test commonly used in the research and practice.
2	Does it support spatial query methods used by current and future clients from various disciplines?	The data model should provide powerful and efficient spatial queries performed by clients from various disciplines.
3	Can the data be easily transformed between different coordinate systems?	In SDIs the data need to be transferable between different coordinate systems.
4	Can the data model be exchanged easily over the Internet by web services?	Some data model formats might be too big to be efficiently exchanged though the internet. Some formats raise security concerns due to potential for malware infection. Some have problems with firewalls and security protocols.
5	Does data model use a data dictionary that is compatible with what developers and clients are currently using?	Using familiar terminology in the data model will shorten the learning curve for developers and clients. Possible solutions are using the terminology from an existing data format, using common English terminology and using a model structure similar to models used in other common applications.
6	Is data model compatible with XML technology to facilitate transformation and conversion?	The data should be easy to convert from existing formats to the new data model. Clients should be able to easily parse the data and adopt their existing applications to work with the data
7	Is data model extensible? Can new tests, attributes and user-defined tests be added to the data model?	Data format should be extensible to accommodate new tests, special cases and user-defined data.
8	Does data model use open standards?	The data model needs to be open to developers and clients in order to allow universal acceptance and utilization; proprietary standards are not recommended due to the limitations they impose on the data and proprietary tools required to work with.
9	Is it compatible with existing industry-approved data model standards?	Data model should be compatibility with existing geotechnical and spatial data formats, e.g., AGS, XML, GML, and WFS. This compatibility facilitates using the model in multi-discipline application.
10	Is it compatible with existing applications?	If data format is compatible with existing standards, existing tools and programs can be used to speed up development time for both developers and clients.

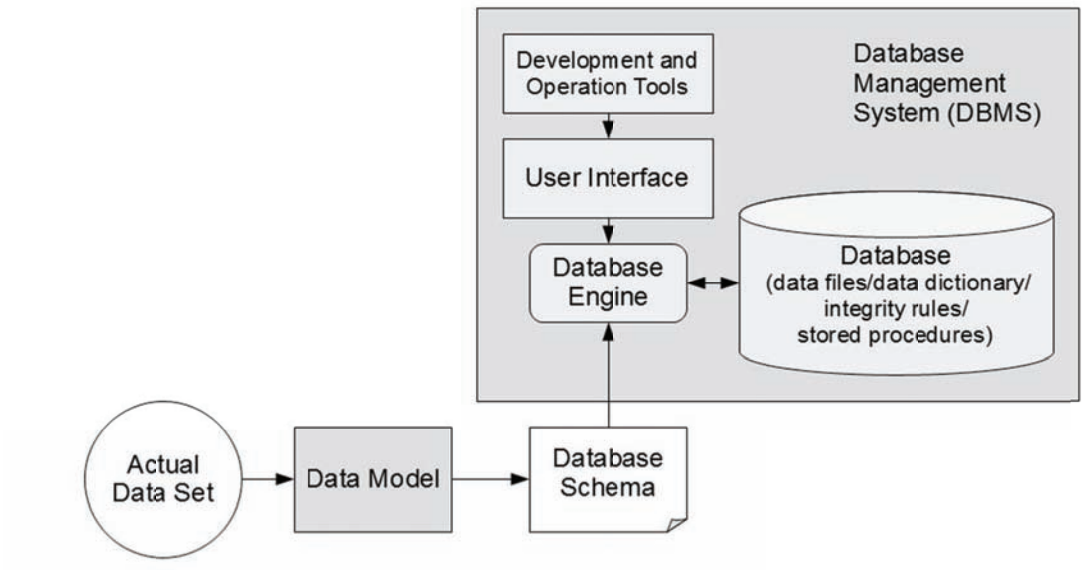


Figure 3-1. The schematic structure and main components of a database management system (DBMS).

(after Yeung and Hall, 2007)

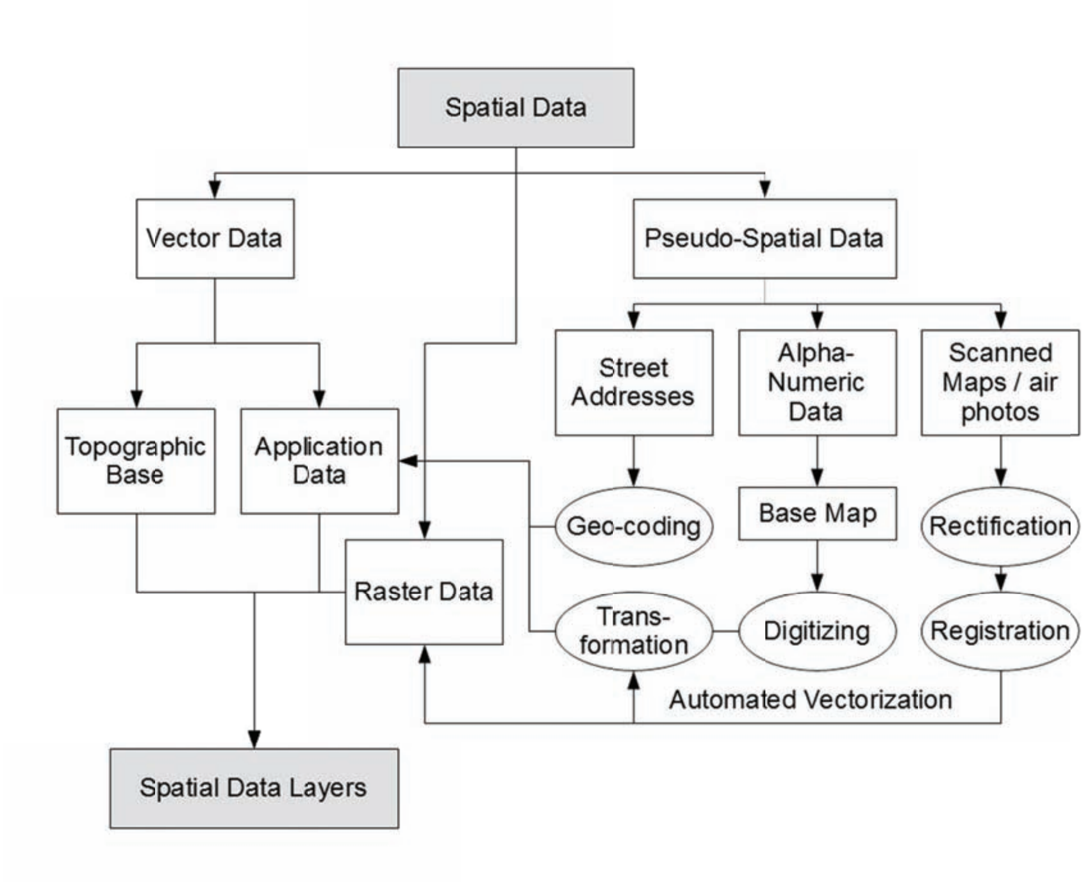


Figure 3-2. Spatial data types and their usage in GIS applications. The figure also shows the conversion of pseudo-spatial data to spatial data for use in GIS.

(after Yeung and Hall, 2007)

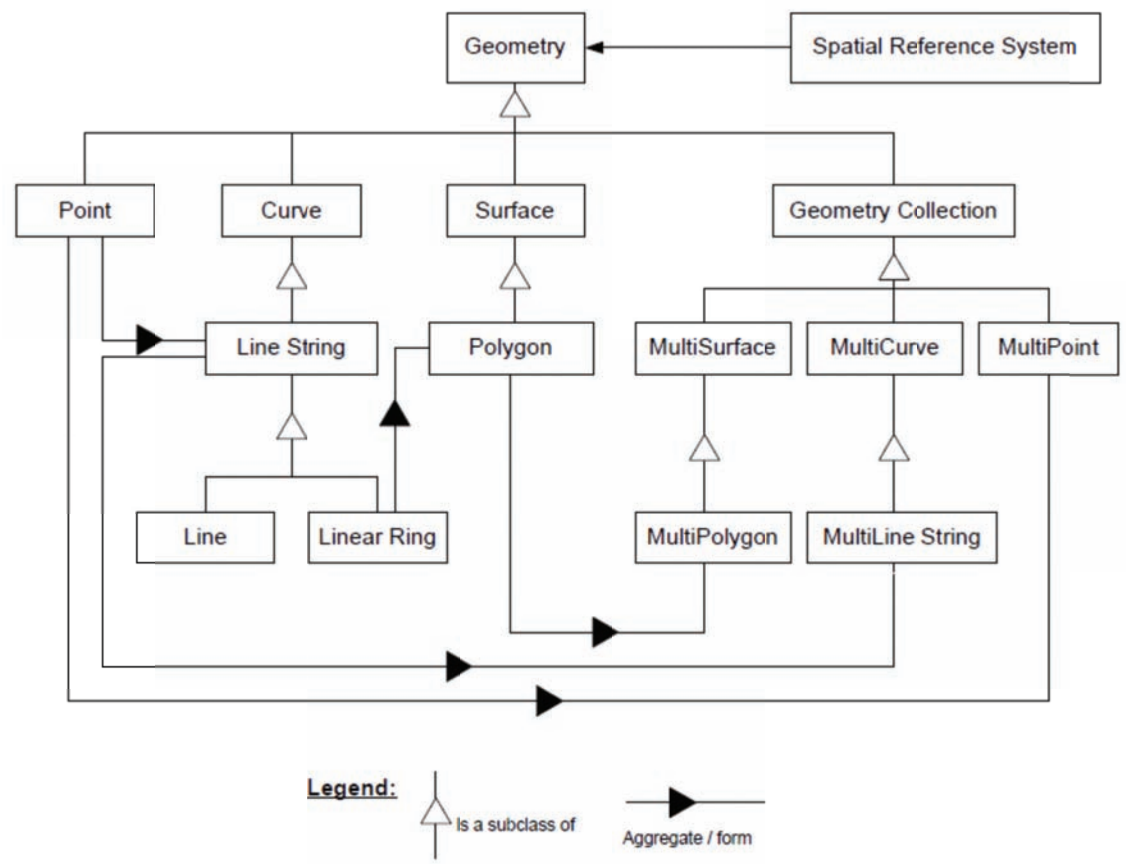


Figure 3-3. The OGC Geometry Object Model.

(after Yeung and Hall, 2007)

Chapter 4. Geotechnical Data Model for Spatial Data Infrastructures

In this chapter, a data format for geotechnical data is presented. This format, named GeotechGML, is developed with the main objective of being used in spatial data infrastructure and large multi-disciplinary applications. The chapter includes the development process of the data format, comparisons with similar formats, validation and implementation.

4.1 GML Encoding of Geotechnical Data

As discussed in Chapter 3, as a markup language, GML is ideally suited for incorporating geotechnical data in SDIs. Some of the advantages of GML for this purpose are: (1) GML is commonly used in other SDIs, which assists in easier integration of the data with existing SDI solutions; (2) GML inherently supports complex issues related to data positioning, e.g., coordinate reference systems, data, and earth curvature; (3) GML is XML compatible, and inherits all the advantages of XML, e.g., platform-independence, using a text-based format, self-documenting schemas and extensibility; (4) GML is a Web-compatible format and is the foundation of Geo-Web; and (5) GML data is easily produced, parsed and manipulated by a large number of available XML processing tools.

The potential benefits of a GML-conformant data format for soil exploration data are recognized by other researchers in geotechnical engineering and other related fields, resulting in similar efforts to develop GML schemas for oil and gas explorations (POSC, 2008) and geotechnical investigations (Ponti et al., 2006).

4.1.1 GML Profiles

One of the problems with the current GML specification is that it is very extensive. It covers more than 1,000 elements, defines many of the geometries for describing features on earth, and also supports the ability to encode objects such as coverage, imagery, topology, time, metadata and dynamic features. GML was designed to be a very general and flexible markup language and covers many needs. This provides challenges in developing GML compatible software, because the software needs to be very comprehensive to be able to read and write general GML files.

One solution to this problem is to support a subset of GML for a specific application. This is called a GML Profile. All user groups can define GML profiles based on their needs. One of the most often used profiles is Simple Feature Profile, which is defined by OGC. The Simple Feature Profile supports simple features exclusively. Only points, lines, and polygons (and collections of these), with linear interpolation between vertices of lines, and planar (flat) surfaces within polygons, are supported.

The motivation for OGC to develop this profile was to encourage software developers to support GML. It is much easier to support Simple Feature Profile than the entire GML schema.

One of the key decisions in developing a schema for new applications is to decide on the extent of GML, which will be included in the schema. A simple profile like simple feature profile will lead to shorter development time for both the data provider and the client. In return, it won’t have the same functionality as a full-featured GML schema because it lacks some of the more advanced GML aspects.

A similar approach can be used by domain-specific GML schemas to reduce the size and complexity of the application schemas and supporting software required to implement the schema. Once a schema is fully developed, it can be easily determined which subset of the GML schema is actually used in the domain-specific schema. Subsequently a GML profile can be defined to include that subset and exclude other unused portions of the GML. Using a GML profile in this manner has several advantages, including the following:

- It reduces the size of scheme
- Increases the speed of instance file validations
- Prevents users from using the schema in unwanted or unpredictable ways
- Reduces the cost and effort needed to develop server and client software, because it does not need to support unused GML features

4.1.2 Guidelines for Development of the GML Schema

Although there are general guidelines on GML schema validation and correctness, there are infinite possibilities when developing a GML schema for geotechnical data, or any other data for that matter. There are no firm guidelines on how to encode different data in GML. Despite this fact, there are several key factors in the process of developing a GML schema for geotechnical data that have major impact on the final outcome and functionality of the resulting GML data. In the following sections these factors are discussed and the selection of these parameters for the current data format is rationalized.

4.1.3 Data Dictionary and Terminology

One of the challenges in developing data standards for any discipline is to reach a consensus on a data dictionary for representing domain-specific objects and properties. The data dictionary should be comprehensive, well understood and widely accepted within the community. This challenge definitely applies to the geotechnical information.

One approach is to deliberately avoid or minimize these challenges by capitalizing on the existing experience and past efforts of the numerous domain experts used in the development of the AGS data exchange format. By utilizing the existing data dictionary of AGS, which has a well-defined and accepted vocabulary, the new efforts are limited to the conversion of AGS objects and properties into the GML rather than the invention of a new data dictionary from scratch. The resulting GML schema adheres to the AGS data dictionary and preserves data nomenclature already accepted and used by the large community of AGS users.

4.1.4 Using Simple vs. Complex GML Features

There are two types of GML features: simple features and complex features. Simple features have non-feature attributes. Complex features can have properties that are features themselves. Using simple versus complex features impacts the functionality and usability of the resulting schema.

Schemas comprised of simple features provide straightforward mapping from a database table or similar structure to a “flat” XML representation, where every column of the table maps to an XML element that usually contains no further structure. This characteristic makes the conversion from columns in a database table to XML elements automatic. The name of the feature type defaults to the name of the database table. The name of each non-feature element can match the name of the corresponding database column.

Complex features provide more possibilities for defining an object-oriented extensible data model. However, the development of the data model, the underlying database and client applications are more involving.

Table 4-1 summarizes the pros and cons of using each feature type in a GML schema. Based on this discussion, the author believes that using complex features provides a better solution for a geotechnical data model and this approach has been used in the GML schema development.

4.2 GML Encoding of the AGS Data Groups

4.2.1 Project Data

The project metadata in AGS are stored in PROJ data group. This data group is encoded as *Proj* feature in GML data model. Figure 4-1 shows the design diagram of this element in GML. The GML data model uses a hierarchical structure, versus the relational data structure of the AGS file format. This can be seen in Figure 4-1, as each borehole in the project is modeled as a *Hole* feature, which is a child element of the *Proj* feature.

4.2.2 Data Groups with Geospatial Data

Close study of AGS data format shows that the spatial data in the 74 AGS data groups can be encoded using only the following GML geometries: (1) a single *Point* geometry; (2) two individual *Point* geometries; (3) a *LineString* geometry; and (4) a *MultiPoint* geometry. The HOLE data group has the highest content of spatial data among AGS data groups. The HOLE data group is encoded as a *Hole* feature in the GML. The *Hole* feature is the top-level element for each soil borehole that contains the basic geometric information of the borehole. This feature is discussed in more detail in the next section. Other data groups have been encoded using one of the common geometric features mentioned above in more or less similar fashion. The GML encoding methodology for these data groups has been demonstrated through several examples involving selected AGS data groups.

4.2.2.1 Hole Feature

Figure 4-2 shows the geometric attributes of the feature *Hole*, which is the top-level element for a borehole object. *Hole* geometry is defined by a point (*Hole_Top*) and an axis (*Hole_Axis*), which are both positioned by longitude, latitude and elevation in a global coordinate reference system. The axis (*Hole_Axis*) is later used to define the position of all other objects along the borehole, using a single parameter—depth. For a straight borehole, the global location of any point is given by its oriented distance to point *Hole_Top* as follows:

$$\mathbf{x} = \mathbf{x}_{Hole\_Top} + d\,\mathbf{n}$$

where $\mathbf{x}_{Hole\_Top}$ are the coordinates of the top point (*Hole_Top*), which is usually located on the ground surface; **n** is a unit vector along the borehole axis; and *d* is the depth along the borehole. The collar point (*Hole_Top*) and the axis (*Hole_Axis*) define a Coordinate Reference System (CRS) in GML (*Hole_Crs*) unique to each individual borehole. *Hole_Crs* is used to position other borehole features (e.g., samples, field tests) relative to the borehole top. As local CRS type for boreholes, *EngineeringCRSType* is preferred to other CRS types (e.g., *GeographicCRSType* and *GeocentricCRSType*). An *EngineeringCRS* is composed of a name identifier (e.g., *srsName*), a

LinearCS (Linear Coordinate System) and a *datum* that defines the CRS origin. A *LinearCS* is a one-dimensional coordinate system that describes the points along a single axis. The only parameter used to locate a point in this coordinate system is the distance from the axis origin. As shown in Figure 4-2, the *LinearCS* is anchored to the *Hole_Top* and aligned along *Hole_Axis*. This representation is applicable to straight boreholes, which include the majority of boreholes in practice. It can be extended to directional boreholes by defining a curve, instead of straight line, as borehole axis. Figure 4-3 illustrates using the *LinearCS* coordinate system to establish the location of soil layer observations, samples and specimens along the borehole.

It should be noted that future versions of GML (GML 3.3) will support a new linear referencing system similar to the method described above. The current proposal for this referencing method is shown in Figure 4-4. Once this referencing method is adopted by GML, it will substitute the user-defined method discussed above.

As shown in Table 4-2, AGS defines boreholes in the HOLE data group with 33 attributes. Table 4 shows the schema for the corresponding GML feature *Hole*, which is defined by extending *AbstractFeatureType*. This data group is encoded as a feature with three geometric properties: (1) *Hole_Top*, (2) *Hole_Axis*, and (3) *Hole_Trv* (optional). *Hole_Top* and *Hole_Axis* are required for all *Hole* features. *Hole_Top* is defined by the AGS coordinates HOLE_NATE, HOLE_NATN and HOLE_GL. *Hole_Trv* is optionally used to define the end point of a traverse (i.e., inclined or vertical surface dug during site investigation). *Hole_Trv* is defined using the AGS coordinates HOLE_ETRV, HOLE_NTRV and HOLE_LTRV. *Hole_Axis* is defined as a *gml:Curve* feature, which is used to encode a general curve in GML. A linear segment (*gml:LineString*) is a special case of a general curve and can be substituted for *curve* feature. Figure 4-5 shows the partial design diagram of the feature *Hole*. Only some of the child elements are shown in this figure. The hierarchical structure of the data model is also evident in this figure.

Table 4-4 shows the schema for *BoringDatum*, an extension of *gml:EngineeringDatum*, which defines the datum for each borehole CRS using *Hole_Axis* and *Hole_Top*. *Hole_Axis* can be defined by specifying either (1) the top and bottom points of the borehole in a global coordinate system, or (2) the bearing angle (measured clockwise from north) and inclination angle (measured positive downward from horizon). A new GML curve feature, *CurveByDirection*, is defined to allow the encoding of borehole axis using bearing and inclination angles. Table 4-5 shows the corresponding schema, and Table 4-6 shows an instance of this feature. The angles of *segmentDirection* are the AGS fields HOLE_ORNT and HOLE_INCL.

All borehole information, e.g., field observations, geology observations, field tests and samples, is nested as child elements within the *Hole* element. For instance, in the schema in Table 4-3, element *Cdias* represents casing diameter vs. depth from the AGS CDIA data group, while *Geols*, *Ispts*, *Stcns* and *Samps* represent soil lithology, SPT test, CPT test and soil samples. These elements correspond to AGS data groups GEOL, ISPT, STCN and SAMP. This approach deviates from AGS methodology, which stores these data in independent data groups and relates them to the HOLE data group through external key fields. Other data fields in *HOLE_GROUP* are encoded as non-geometric properties. Appropriate GML non-geometric data types, e.g., *gml:lengthType* and *gml:CalDate,* are used for non-geometric data. The former data type applies to measured length, and the latter one to dates. Table 4-7 lists an instance of GML *Hole* feature, which corresponds to the AGS data of shown in Table 2-11 and Figure 2-14 in Chapter 2. Table

4-8 shows how to use the *srsName* attribute to define a soil layer (*Geol* feature) in the local coordinate system of a borehole (corresponding original data is in Table 2-11).

4.2.2.2 Other Geospatial Features

All the geometrical features of a borehole (e.g., samples, soil layer observations and field tests) can be positioned using the parameter depth (*d*) in the local coordinate systems *Hole_Crs*. All features fall into four classes based on their spatial data content, namely features with a single *Point* geometry, features with two individual *Point* geometries, features with *LineString* geometry and features with *MultiPoint* geometry.

A single *Point* is positioned by a singular value of *d* in borehole local coordinate system. The feature with a single *Point* geometry carries the information that concerns that point in the soil strata.

A feature with two individual *Point* geometries is positioned by two depth values d_1 and d_2. A *LineString* along the borehole axis is also positioned by the same two depth values. The difference between two features with double *Point* geometries and *LineString* geometry is that the two individual *Point* geometries represent two distinct locations while a *LineString* concerns the interval between its beginning and end points.

A *MultiPoint* geometry is specified by a series of *n* depth values, d_i, $i = 1,\ldots,n$, each one positioning a point in the borehole local coordinate system.

The following sections illustrate the GML encoding methodology for these feature classes through examples for each type.

4.2.2.3 Data Groups with Single- Point Geometry

Table 4-9 lists all the AGS data groups with a single *Point* geometry and identifies the AGS data fields that establish the point location. Table 4-10 and Table 4-11 illustrate how to encode single *Point* geometries through the example ISPT, the AGS data group for the Standard Penetration Test (SPT). As discussed previously in Chapter 2 the SPT estimates the soil strength by counting the number of hammer blows required to drive a split-barrel sampler one foot into the ground. Table 4-10 illustrates the AGS fields of ISPT, and Table 4-11 shows the corresponding GML encoding. ISPT is encoded as *Ispt* GML feature. The only geometric property, *Ispt_Top*, refers to a *gml:Point* geometry, which determines the position of the top of the SPT test. Other fields in the ISPT data group are encoded using non-geometric attributes.

4.2.2.4 Data Groups with Two Point Geometries

Table 4-12 lists all the AGS data groups with two distinct *Point* geometries. These data groups contain the laboratory test results on soil samples. AGS defines the position of soil samples using two depths (Figure 4-3). The first depth corresponds to the top of the sample in the field, and the second depth corresponds to the soil specimen that is tested in the laboratory. For instance, as shown in Table 4-13, the GRAD data group reports the results of particle size distribution analysis, and Table 4-14 shows the corresponding GML encoding for this data group. GRAD is encoded as a *Grad* feature, with two feature properties, *Samp_Top* and *Spec_Dpth*, both referring to *gml:Point* geometries.

4.2.2.5 Data Groups with LineString Geometry

Table 4-15 lists all the AGS data groups with a *LineString* geometry, as well as the AGS data fields containing the geometry of the linear segment. The *LineString* geometry represents an interval along the borehole where the information is applicable. For instance, as shown in Table 4-16, the GEOL data group, used to report the lithology along the borehole. Table 4-17 shows the corresponding GML encoding for this data group. GEOL is modeled as a *Geol* feature, and the geometry of the soil layer is encoded as a *LineString* feature, *Geol_Axis*.

Table 4-18 shows the AGS data group SMAP, which is used to store soil sample information. The GML encoding for this data group is shown in Table 4-19. The sample geometry is encoded as a *LineString* feature.

4.2.2.6 Data Groups with MultiPoint Geometry

The GML *MultiPoint* geometry is useful for establishing the position a series of point observations at various depths along the boreholes. This type is used for the Cone Penetration Test (CPT), which reports the variation of tip resistance and sleeve friction versus depth, as discussed in Chapter 2. In the AGS data format the CPT results are stored in the STCN data group. Following the methodology used for other tests each measurement could have been encoded independently as a single Point feature. However, this approach would have resulted in a large data file due to the large number of CPT measurements. A cleaner approach is to use a GML *Coverage* feature, which is a distribution function defined in a number of discrete points over a domain. A *Coverage* typically has a *DomainSet*, a *RangeSet* and a *CoverageFunction*. In case of CPT, the *DomainSet* is the collection of points (*MultiPoint*) along the borehole where the tip resistance and sleeve friction are measured. The *RangeSet* is a set of elements that contain the test results (e.g., values of tip resistance and sleeve friction at each point). The values in *RangeSet* are presented in the same order as the corresponding locations in *DomainSet*. The *CoverageFunction*, which is optional, defines how the values in the *RangeSet* are mapped to the corresponding entities in the *DomainSet*. By default this function is a linear function (one-to-one relationship).

Table 4-20 lists the attributes of the AGS data group STCN, and Table 4-21 shows the corresponding GML encoding. The data group is encoded as *Stcn* feature, which is an extension of the *AbstractCoverageType*. The *Stcn_Depths* property is a *MultiPointType.* Figure 4-6 illustrates a graphical representation of the *Stcn* feature.

4.2.3 GML Encoding for Non-Geospatial Data Fields

Although the GML core schema is mainly intended for geospatial data, it also has extensive support for encoding non-spatial information. In the following sections some GML capabilities that are used for the GML encoding of the non-geometric data fields in the AGS data format are discussed.

4.2.3.1 Non-Geospatial Data types

The GML core schemas support a large number of non-spatial data types. A number of these data types are used extensively in the AGS-based GML schema discussed here. These GML feature types are as follows:

gml:measureType: this data type is a double number with a *uom* attribute that is applicable to a large number of engineering measurements. The *uom* attribute defines the unit which is used to measure the quantity, and is a required attribute.

gml:length, *gml:angle*: both extensions of *gml:measureType*, these data types are used when the measured value is a length or angle, respectively. The *uom* attribute, which is inherited from *gml:measureType*, should be used to specify the unit of measurement.

gml:CalDate and *gml:dateTime* are used to specify dates and times in standard GML format.

gml:CodeType: this type is an extension of string data type and is used for terms, keywords or names, with an additional optional attribute for specifying the code space of the data.

gml:DirectionVectorType: this data type is used to define a vector. A *srsName* attribute can be used to specify the corresponding coordinate reference system.

Additional information on these and other data types can be found in the GML specifications document (OGC, 2008a).

4.2.4 Data Dictionaries

GML *dictionaries* collect sets of definitions, e.g., CRS definitions, units of measurement definitions, standard values and other user-defined values. A GML dictionary provides a mechanism for GML data to access the *dictionary* entries through an identification attribute (*gml:id*). Table 4-22 shows a sample GML *dictionary*, containing definitions for SI units; other GML document instances can refer to the units defined in this *dictionary* by using a *gml:id* attribute. GML *dictionaries* are used to encode two mandatory AGS data groups ABBR and

`UNIT`, as well as optional groups `CODE`, `FILE` and `DICT`. GML *dictionaries* are stand-alone documents and can be included either directly in the instance data files, or through external references, which allow the same *dictionary* to be used by several GML instance documents. GML *dictionaries* may contain standard definitions, which can be reused as external documents by multiple instance data files. GML *dictionaries* are more flexible than AGS dictionaries, which need to be embedded in the instance files. Bobbitt (2004) provides some useful recommendations on the usage of GML *dictionaries*.

4.2.5 Units of Measurement

Geotechnical data are reported using various units of measurement, which may even combine metric and British units. AGS requires geotechnical data to be reported explicitly with their units of measurement. This is a good practice because it prevents misinterpretation of data by using wrong units. Similarly GML has provisions for values that have a unit of measurement. GML describes units using the *uom* attribute for *gml:measureType* and other data types that inherit from the *gml:measureType* (e.g. *gml:lengthType*). This attribute typically refers to a unified resource identifier (URI) that defines the unit of measurement. In most cases, the unit is defined in a GML dictionary, either in the same document or in a remote resource. Bobbitt (2003) provides useful recommendations on the usage of units of measurement in GML.

4.2.6 Validation and Implementation

The newly developed AGS-based GML application schemas were validated to ensure conformance with both XML and GML specifications at two different levels. At the first level, a XML schema editor, e.g., XMLSpy (Altova, 2008), was used to ensure that schemas were XML compliant and all element types are well-defined based on XML criteria. At the second level, for GML compatibility, a GML schema validator (OGC, 2008b) was used. These validations were only required once during schema development. In routine exchanges of data, well-formed data files can be validated against their GML schema using any general purpose XML editor. This should guarantee that instance documents are compliant to the GML schemas and can be displayed using any GML compliant viewer.

Figure 4-7 illustrates the integration potential of the GML-conformant data files. In this example, the GML data file was visualized using a two-dimensional GML viewer software, Gaia (The Carbon Project, 2008). The GML data file was directly read by software without any transformation, resulting in the display of the borings on top of the Microsoft Virtual Earth map.

Figure 4-8 shows a more complicated three-dimensional representation of the same data as in Figure 4-7 using Google Earth (Google Inc., 2008a). Because currently Google Earth in not GML-conformant this exercise requires the GML data to be first transformed to KML, the native language of Google Earth (Google, 2008b). Since KML is also an XML-based data format, the conversion can be performed easily using an XSLT transformation (W3C, 2008). During this transformation an XSLT style sheet converts the GML compatible data format to KML, and adds the styling instructions for various data fields (e.g., font sizes, symbols and colors for various soil types).

As shown in Figure 4-8, the two boreholes have been represented upside down as buildings are usually represented. This subterfuge was forced due to Google Earth limitations for rendering buried structures. Using this technique one may represent the location and depth of soil layers, soil samples, SPT field tests and lab tests in Google Earth.

It is anticipated that many more programs will natively support the GML in the future, rendering the visualization of geotechnical data even more vivid and powerful.

4.3 Other Spatial Geotechnical Data Models

Due to the current demand for a state-of-the-art data format for geotechnical data, there have been other initiatives to develop a GML-compatible data format for geotechnical data. The major ongoing effort is supported by the Data Interchange for Geotechnical and Geoenvironmental Specialists (DIGGS). According to the DIGGS website, "DIGGS is a coalition of government agencies, universities and industry partners whose focus is on the creation and maintenance of an international data transfer standard for transportation related data".

DIGGS was started as a project supported by Federal Highway Administration (FHWA). It later evolved with the goal to develop an international standard for geotechnical data interchange through merging data formats created by the Association of Geotechnical and Geoenvironmental Specialists in the United Kingdom (AGS), Consortium of Organizations for Strong-Motion Observation Systems (COSMOS) and Florida Department of Transportation (FDOT) created by the University of Florida (UF). Initially it was supporting the geotechnical data including borehole, soil testing, site information, etc. The first Special Interest Group (SIG) is extending the schema to include geoenvironmental testing and more SIGs are expected to join the coalition.

4.3.1 DIGGS Model Structure

The root element of the DIGGS format for geotechnical data is “Diggs”. Figure 4-9 shows a diagram of the “Diggs” element schema, generated by oXygen. This particular schema includes the following elements at top level: “associatedFiles”, “businessAssociates”, “equipment”, “groupings”, “projects”, “specifications” and “transmissionInformation”. The “projects” element includes one or several “project” sub elements, each containing the geotechnical data pertaining to them, as shown in Figure 4-10. These data include “locations”, “samples” and “laboratoryTesting”.

The “locations” element contains the boreholes where geotechnical investigation and testing has taken place. Boreholes are encoded as “HOLE” elements, as shown in Figure 4-11. In-situ tests are encoded as child elements in “HOLE”. Figure 4-12 shows the “InsituTest” element, which is used to encode field tests. These tests include SPT and CPT tests, as shown in Figure 4-13 and Figure 4-14.

In DIGGS format, the samples and laboratory tests are encoded as parallel elements to locations. Figure 4-15 shows the schema diagram of the “Sample” element. Figure 4-16 shows the same diagram for the “LaboratoryTest” element. As an example of a laboratory test, the “MoistureContent” element is shown in Figure 4-17.

Figure 4-18 and Figure 4-19 show the schematic diagram of two DIGGS instance documents. Figure 4-18 shows the upper level elements of the data format, Including “businessAssociates”, “equipment”, “projects” and “transmissionInformation”. Figure 4-19 also shows the in-situ and laboratory test elements.

4.3.2 Comparison between GeotechML and Diggs

Currently the scope of GeotechML is limited to the AGS data format. Therefore, it covers geotechnical and geoenvironmental data. Diggs, on the other hand, currently covers several SIGs in addition to geotechnical and geoenvironmental data. For example, one of these SIGs is piling, which covers data from pile driving and construction. These models, however, are extensible and it is possible to add SIGs to both formats down the road.

In addition to the scope of current data model there are other differences between two models. Table 4-23 summarizes the major differences between GeotechML and Diggs.

The most fundamental difference between DIGGS and GeotechML is the way samples and laboratory test results are encoded. In DIGGS these data are encoded at the same level as the borehole, as shown in Figure 4-10, while in GeotechML the samples and laboratory tests are encoded in a hierarchical structures.

The DIGGS review document (reference) states that its encoding methodology allows the composite specimens (specimens that are made by combining soils from several boreholes) to be modeled, as well as laboratory tests performed on such samples. The problem with this encoding method is that it is much less effective in terms of querying in the database.

The author believes that while these types of samples are frequently used in practice, they are relatively uncommon. They also have little value as spatial data because they cannot be associated with a single well-defined location. Therefore, it doesn’t make sense to jeopardize the entire data model to accommodate this infrequent situation. Rather than using a general sample feature at the same level as *Hole* feature, this situation can be addressed by a different

CompositeSamp feature, similar to DIGGS sample element. This concept is demonstrated in the design diagram presented in Figure 4-20. Since this type of sample is not associated with a single borehole, it is encoded at the same level as boreholes in the hierarchical GML model. In comparison, the regular sample feature is encoded as a child property under the borehole feature. By using this approach the spatial integrity of the data model is maintained.

A similar approach could be used for lab tests performed on unreferenced soil samples or samples not associated with a borehole.

Table 4-1. Pros and cons of simple features versus complex features in GML.

Feature Type	Pros	Cons
Simple Features	• Easy to implement • Fast development time • Support queries on properties, including spatial queries on geometries	• The GML schema is tied to the database schema • To share data on simple features, either the same database schema must be used or data must be transformed between different schemas • Interoperability is difficult because simple features do not allow modification of only part of the schema • As more data owners with different data are added to a community, the number of columns in the table need to be adjusted
Complex Features	• Can define information model as an object-oriented structure, e.g., an application schema • Information is modeled not as a single table but as a collection of related objects whose associations and types may vary from feature to feature (polymorphism), permitting rich expression of content • By breaking the schema into a collection of independent types, communities need only extend those types they need to modify. This simplifies governance and permits interoperability between related communities that can agree on common base types but need not agree on application-specific subtypes.	• More complex to implement • Could result in slower queries • The user community needs to agree on a data model/application schema

(after GeoServer, 2010)

Table 4-2. AGS data transfer format for HOLE data group. This data group is used to store the data for boreholes, e.g., borehole type, location and geometry.

No.	Heading	Unit	Description
1	HOLE_ID*		Exploratory hole or location equivalent
2	HOLE_TYPE		Type of exploratory hole
3	HOLE_NATE	m	National grid Easting of hole or start of traverse
4	HOLE_NATN	m	National grid Northing of hole or start of traverse
5	HOLE_GL	m	Ground level relative to Datum of hole or start of traverse
6	HOLE_FDEP	m	Final depth of hole
7	HOLE_STAR	dd/mm/yyyy	Date of start of excavation
8	HOLE_LOG		The definitive person responsible for logging the hole
9	HOLE_REM		General remarks on hole
10	HOLE_ETRV	m	National grid Easting of end of traverse
11	HOLE_NTRV	m	National grid Northing of end of traverse
12	HOLE_LTRV	m	Ground level relative to Datum of end of traverse
13	HOLE_LETT		Ordinance Survey letter grid reference
14	HOLE_LOCX	m	Local grid x co-ordinate
15	HOLE_LOCY	m	Local grid y co-ordinate
16	HOLE_LOCZ	m	Level to local datum
17	HOLE_ENDD	dd/mm/yyyy	Hole end data
18	HOLE_BACD	dd/mm/yyyy	Hole backfill data
19	HOLE_CREW		Name of driller
20	HOLE_ORNT	deg	Orientation of hole or traverse
21	HOLE_INCL	deg	Inclination of hole or traverse (measured positively down from horizontal)
22	HOLE_EXC		Plant used
23	HOLE_SHOR		Shoring/support used
24	HOLE_STAB		Stability
25	HOLE_DIML	m	Trial pit or logged traverse length
26	HOLE_DIMW	m	Trial pit or logged traverse width
27	HOLE_LOCM		Method of location
28	HOLE_LOCA		Location sub-division within project
29	HOLE_CLST		Hole cluster reference number
30	HOLE_OFFS		Offset
31	HOLE_CNGE		Chainage
32	HOLE_STAT		Status of hole information
33	FILE_FSET		Associated file reference

* key field.

Table 4-3. GML encoding for HOLE data group.

Schema[1]	Comments
<xs:element name="Hole" type="app:HoleType"	
substitutionGroup="gml:_Feature"/>	
<!-- === -->	
<xs:complexType name="HoleType">	
<xs:annotation>	
<xs:documentation>A single instance of borehole, test pit or	
traverse</xs:documentation>	
</xs:annotation>	
<xs:complexContent>	
<xs:extension base="gml:AbstractFeatureType">	
<xs:sequence>	
<xs:element name="Hole_id" type="xs:string"/>	
<xs:element name="Hole_Type" type="app:Hole_TypeType"/>	
<xs:element name="Hole_Top" type="gml:PointPropertyType"/>	
<xs:element name="Hole_Axis" type="gml:CurvePropertyType"/>	-Hole top
<xs:element name="Hole_Crs" type="gml:EngineeringCRSType"/>	-Hole axis
<xs:element name="Hole_Fdep" type="gml:LengthType"	-Hole CRS
minOccurs="0"/>	
<xs:element name="HOLE_STAR" type="gml:CalDate" minOccurs="0"/>	
<xs:element name="HOLE_LOG" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_REM" type="xs:string" minOccurs="0"/>	
<xs:element name="Hole_Trv" type="gml:PointPropertyType"	
minOccurs="0"/>	-Optional
<xs:element name="HOLE_LETT" type="xs:string" minOccurs="0"/>	Point for
<xs:element name="HOLE_ENDD" type="gml:CalDate" minOccurs="0"/>	traverse
<xs:element name="HOLE_BACD" type="gml:CalDate" minOccurs="0"/>	
<xs:element name="HOLE_CREW" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_EXC" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_SHOR" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_STAB" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_DIML" type="gml:LengthType"	
minOccurs="0"/>	
<xs:element name="HOLE_DIMW" type="gml:LengthType"	
minOccurs="0"/>	
<xs:element name="HOLE_LOCM" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_LOCA" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_CLST" type="xs:string" minOccurs="0"/>	
<xs:element name="HOLE_OFFS" type="gml:LengthType"	
minOccurs="0"/>	
<xs:element name="HOLE_CNGE" type="gml:LengthType"	
minOccurs="0"/>	
<xs:element name="HOLE_STAT" type="xs:string" minOccurs="0"/>	
<xs:element name="FILE_FSET" type="app:FileType"	
minOccurs="0"/>	
<xs:element ref="app:Cdias" minOccurs="0"/>	
<xs:element ref="app:Geols" minOccurs="0"/>	-Casing
<xs:element ref="app:Ispts" minOccurs="0"/>	info
<xs:element ref="app:Stcns" minOccurs="0"/>	-Lithology
<xs:element ref="app:Samps" minOccurs="0"/>	-SPT test
…	-CPT test
</xs:sequence>	-Sample
</xs:extension>	info
</xs:complexContent>	
</xs:complexType>	

1. New schema elements are defined in namespace ***app***. GML schema elements are defined in namespace ***gml***. XML schema elements are defined in namespace ***xs***.

Table 4-4. GML schema for BoringDatum element. BoringDatum defines a local datum for each borehole, which is used to reference soil layers and samples along the borehole.

Schema	Comments
`<!-- ================================================ -->`	
` <xs:element name="BoringDatum"`	
`     type="app:BoringDatumType"`	
`     substitutionGroup="gml:EngineeringDatum"/>`	
`<!-- ================================================ -->`	
`<xs:complexType name="BoringDatumType">`	
` <xs:annotation>`	
`  <xs:documentation>`	
`     Extends EngineeringDatum for boreholes by adding`	
`     definition of a curve and a point (origin) to the`	
`     base definiton`	
`  </xs:documentation>`	
` </xs:annotation>`	
` <xs:complexContent>`	
`  <xs:extension base="gml:EngineeringDatumType">`	-Extends
`   <xs:sequence>`	EngineeringDatumType
`    <xs:element name="usesCurve"`	-Boring axis
`     type="gml:CurvePropertyType" minOccurs="0"/>`	
`    <xs:element name="origin"`	-Boring origin
`     type="gml:PointPropertyType" minOccurs="0"/>`	
`   </xs:sequence>`	
`  </xs:extension>`	
` </xs:complexContent>`	
`</xs:complexType>`	

Table 4-5. GML encoding for a CurveByDirection element defined by Origin and Length/Direction pairs.

Schema	Comments
`<!-- ================================================== -->`	
` <xs:element name="CurveByDirection"`	
`     type="app:CurveByDirectionType"`	
`     substitutionGroup="gml:_Curve"/>`	
`<!-- ================================================== -->`	
`<xs:complexType name="CurveByDirectionType">`	
` <xs:annotation>`	
`  <xs:documentation>`	
`    A curve defined by an origin and one or more distance-`	
`     direction pairs</xs:documentation>`	
` </xs:annotation>`	
` <xs:complexContent>`	-Extends
`  <xs:extension base="gml:AbstractCurveType">`	AbstractCurveType
`   <xs:sequence>`	-Origin point
`    <xs:element name="origin" type="gml:PointPropertyType"/>`	-Segment(s)
`    <xs:element ref="app:segment" maxOccurs="unbounded"/>`	
`   </xs:sequence>`	
`  </xs:extension>`	
` </xs:complexContent>`	
`</xs:complexType>`	
`<!-- ================================================== -->`	
`<xs:element name="segment" type="app:SegmentType"/>`	
`<!-- ================================================== -->`	
`<xs:complexType name="SegmentType">`	
` <xs:annotation>`	
`  <xs:documentation>`	
`    A segment on a curve defined by direction`	
`     (CurveByDirection)`	
`  </xs:documentation>`	
` </xs:annotation>`	-Length of
` <xs:sequence>`	segment
`  <xs:element name="segmentLength" type="gml:LengthType"`	-Direction of
`     minOccurs="1" maxOccurs="1"/>`	segment (vector)
`  <xs:element name="segmentDirection"`	
`     type="gml:DirectionVectorType" minOccurs="1"`	
`     maxOccurs="1"/>`	
` </xs:sequence>`	
`</xs:complexType>`	

Table 4-6. An instance of CurveByDirection defined by Origin and a single Length-Direction pair.

GML

```
<CurveByDirection gml:id="Hole_BH502_Axis">
 <origin xlink:href="#Hole_BH502_Collar"/>
 &lt;segment&gt;
  <segmentLength uom="#degrees">15.45</segmentLength>
  <segmentDirection>
   <gml:horizontalAngle uom="#degrees">0.0</gml:horizontalAngle>
   <gml:verticalAngle uom="#degrees">90.0</gml:verticalAngle>
  </segmentDirection>
 &lt;/segment&gt;
</CurveByDirection>
```

Table 4-7. A sample instance of Hole feature, based on the schema shown in Table 4-3.

Schema	Comments
`<Hole>`	
` <Hole_id>BH502</Hole_id>`	
` <Hole_Type>RO</Hole_Type>`	
` <Hole_Top>`	
`  <gml:Point gml:id="Hole_BH502_Collar">`	-Hole top
`   <gml:pos>-122.4266 37.7603 58.72</gml:pos>`	coordinate
`  </gml:Point>`	
` </Hole_Top>`	
` <Hole_Axis>`	
`  <CurveByDirection gml:id="Hole_BH502_Axis">`	-Hole axis
`   <origin xlink:href="#Hole_BH502_Collar"/>`	defined by
`   <segment>`	length and
`    <segmentLength uom="m">15.45</segmentLength>`	direction
`    <segmentDirection>`	
`     <gml:horizontalAngle uom="#degrees">0.0</gml:horizontalAngle>`	
`     <gml:verticalAngle uom="#degrees">90.0</gml:verticalAngle>`	
`    </segmentDirection>`	
`   </segment>`	
`  </CurveByDirection>`	
` </Hole_Axis>`	
` <Hole_Crs gml:id="Hole_BH502_CRS">`	-Hole CRS,
`  <gml:srsName>BH502 CRS</gml:srsName>`	defined by
`  <gml:usesCS>`	hole top
`   <gml:LinearCS gml:id="Hole_BH502_LCS">`	point and
`    <gml:csName>BH502 CS</gml:csName>`	hole axis
`    <gml:usesAxis>`	
`     <gml:CoordinateSystemAxis gml:id="HOLE_BH502_CSA"`	
`     gml:uom="#m">`	
`      <gml:name>Depth along Borehole 502</gml:name>`	
`      <gml:axisAbbrev>D</gml:axisAbbrev>`	
`      <gml:axisDirection>Down</gml:axisDirection>`	
`     </gml:CoordinateSystemAxis>`	
`    </gml:usesAxis>`	
`   </gml:LinearCS>`	
`  </gml:usesCS>`	
`  <gml:usesEngineeringDatum>`	
`   <BoringDatum gml:id="HOLE_BH502_Datum">`	
`    <gml:datumName>BH502 Ground Level</gml:datumName>`	
`    <usesCurve xlink:href="#Hole_BH502_Axis"/>`	
`    <origin xlink:href="#Hole_BH502_Collar"/>`	
`   </BoringDatum>`	
`  </gml:usesEngineeringDatum>`	
` </Hole_Crs>`	
` <HOLE_STAR>1999-07-22</HOLE_STAR>`	
` <HOLE_LOG>ANO</HOLE_LOG>`	
` <FILE_FSET>#FS003</FILE_FSET>`	
` …`	-Other
`</Hole>`	Hole data
	not shown

Table 4-8. Application of srsName attribute to refer to local CRS for a Point feature. The "Hole_TP501_CRS" refers to the coordinate reference system for borehole TP501, defined in Table 4-7.

Schema	Comments
<Geol> <Geol_Top> <gml:Point gml:id="Hole_TP501_Geol2_Top" srsName="Hole_TP501_ CRS"> <gml:pos>0.0</gml:pos> </gml:Point> </Geol_Top> <Geol_Base> <gml:Point gml:id="Hole_TP501_Geol2_Base" srsName=" Hole_TP501_ CRS "> <gml:pos>1.31</gml:pos> </gml:Point> </Geol_Base>	-Points are defined in hole CRS using single parameter (depth)
<Geol_Axis> <gml:LineString> <gml:pointProperty xlink:href="#Hole_TP501_Geol2_Top"/> <gml:pointProperty xlink:href="#Hole_TP501_Geol2_Base"/> </gml:LineString> </Geol_Axis> <Geol_Desc>Firm brown slightly sandy very closely fissured CLAY with some fine to coarse subrounded gravel. Medium spaced subhorizontal slightly polished gleyed shear surfaces. Widely spaced vertical rough desiccation cracks with concentrations of rootlets. (Weathered Boulder Clay) </Geol_Desc> <Geol_Leg>261</Geol_Leg> <Geol_Geol>WBC</Geol_Geol> <Geol_Geo2/> <Geol_Stat>B</Geol_Stat> </Geol>	-Geol axis is defined by referring to top and bottom points

Table 4-9. AGS data groups that are defined using a single Point geometry.

Table/ Data Group	Data Group Description	Geometry Field Heading	Geometry Field Description
CDIA	Casing Diameter by Depth	CIDA_CDEP	Depth achieved at CDIA_HOLE
DPRB	Dynamic Probe Test	DPRB_DPTH	Depth to start of dynamic probe increment
DREM	Depth Related Remarks	DREM_DPTH	Depth of DREM_REM
HDIA	Hole Diameter by Depth	HDIA_HDEP	Depth achieved at HDIA_HOLE
HPGI	Horizontal Profile Gauge Installation Details	HPGI_DLN	Level of datum point relative to HOLE_GL or HOLE_LOCZ
HPGO	Horizontal Profile Gauge Observations	HPGO_DIS	Distance from datum point to reading point
ICBR	In Situ CBR Test	ICBR_DPTH	Depth to top of CBR test
?ICCT	In Situ Contamination and Chemical Test	?ICCT_DIS	Distance from reference point
IDEN	In Situ Density Test	IDEN_DPTH	Depth of in situ density test
?IFID	On Site Volatile Headscape Testing using Flame Ionisation Detector	?IFID_DPTH	Depth of headscape test sample
INST	Single Point Instrument Installation Details	INST_TDEP	Depth of reference level of instrument from HOLE_GL or HOLE_LOCZ
IOBS	Single Point Instrument Readings	IOBS_TDEP	Depth of reference level of instrument from HOLE_GL or HOLE_LOCZ
ISPT	Standard Penetration Test Results	ISPT_TOP	Depth to top of test
?PID	On Site Volatile Headscape Testing by Photo Ionisation Detector	?PID_DPTH	Depth of headscape test sample
IRDX	In Situ Redox Test	IRDX_DPTH	Depth of redox test
IRES	In Situ Resistivity Test	IRES_DPTH	Depth range to which in situ resistivity test relates
IVAN	In Situ Vane Test	IVAN_DPTH	Depth of vane test
POBS	Piezometer Readings	POBS_TDEP	Depth to reference level of piezometer tip
PREF	Piezometer Installation Details	PREF_TDEP	Depth to reference level of piezometer tip
PRTD	Pressuremeter Test Data Details	PRTD_DPTH	Depth of test
PRTG	Pressuremeter Test Results - General	PRTG_DPTH	Depth of test
PRTL	Pressuremeter Test Results, Individual Loops	PRTL_DPTH	Depth of test
PTIM	Hole Progress by Time	PTIM_DEP	Hole depth at PTIM_TIME
WSTK	Water Strike Details	WSTK_DEP	Depth to water strike

Table 4-10. AGS data transfer format "ISPT" data group. This data group is used to store the data for SPT.

No.	Heading	Unit	Description
1	HOLE_ID*		Exploratory hole or location equivalent
2	ISPT_TOP*	m	Depth to top of test
3	ISPT_SEAT		Number of blows for seating drive
4	ISPT_MAIN		Number of blows for main test drive
5	ISPT_NPEN	mm	Total penetration for seating drive and test drive
6	ISPT_NVAL		SPT 'N' value
7	ISPT_REP		SPT reported result
8	ISPT_CAS	m	Casing depth at time of test
9	ISPT_WAT	m	Depth to water at time of test
10	ISPT_TYPE		Type of SPT test
11	?ISPT_SWP	mm	Self-weight penetration
12	ISPT_REM		Remarks relating to the test
13	ISPT_INC1		Number of blows for 1st Increment (Seating)
14	ISPT_INC2		Number of blows for 2nd Increment (Seating)
15	ISPT_INC3		Number of blows for 1st Increment (Test)
16	ISPT_INC4		Number of blows for 2nd Increment (Test)
17	ISPT_INC5		Number of blows for 3rd Increment (Test)
18	ISPT_INC6		Number of blows for 4th Increment (Test)
19	ISPT_PEN1	mm	Penetration for 1st Increment (Seating Drive)
20	ISPT_PEN2	mm	Penetration for 2nd Increment (Seating Drive)
21	ISPT_PEN3	mm	Penetration for 1st Increment (Test)
22	ISPT_PEN4	mm	Penetration for 2nd Increment (Test)
23	ISPT_PEN5	mm	Penetration for 3rd Increment (Test)
24	ISPT_PEN6	mm	Penetration for 4th Increment (Test)

* key field.

Table 4-11. Proposed GML encoding for AGS ISPT data group, shown in Table 4-10. Each test is positioned using a Point feature at the top of the SPT sample.

Schema

```
<!-- ============================================================-->
<xs:element name="Ispt" type="app:IsptType"
    substitutionGroup="gml:_Feature"/>
<!-- ============================================================-->
<xs:complexType name="IsptType">
 <xs:annotation>
  <xs:documentation>
   Standard penetration test
  </xs:documentation>
 </xs:annotation>
<xs:complexContent>
 <xs:restriction base="gml:AbstractFeatureType">
 <xs:sequence>
  <xs:element name="ISPT_TOP" type="gml:PointPropertyType"
    minOccurs="1"/>
    <xs:element name="ISPT_SEAT" type="xs:string" minOccurs="0"/>
    <xs:element name="ISPT_MAIN" type="xs:string" minOccurs="0"/>
    <xs:element name="ISPT_NPEN" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_NVAL" type="xs:string" minOccurs="0"/>
    <xs:element name="ISPT_REP" type="xs:string" minOccurs="0"/>
    <xs:element name="ISPT_CAS" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_WAT" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_TYPE" type="xs:string" minOccurs="0"/>
    <xs:element name="ISPT_SWP" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_REM" type="xs:string" minOccurs="0"/>
    <xs:element name="ISPT_INC1" type="xs:int" minOccurs="0"/>
    <xs:element name="ISPT_INC2" type="xs:int" minOccurs="0"/>
    <xs:element name="ISPT_INC3" type="xs:int" minOccurs="0"/>
    <xs:element name="ISPT_INC4" type="xs:int" minOccurs="0"/>
    <xs:element name="ISPT_INC5" type="xs:int" minOccurs="0"/>
    <xs:element name="ISPT_INC6" type="xs:int" minOccurs="0"/>
    <xs:element name="ISPT_PEN1" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_PEN2" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_PEN3" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_PEN4" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_PEN5" type="gml:LengthType"
    minOccurs="0"/>
    <xs:element name="ISPT_PEN6" type="gml:LengthType"
    minOccurs="0"/>
   </xs:sequence>
  </xs:restriction>
 </xs:complexContent>
</xs:complexType>
```

Comments

-Location of test is defined by this point defined in hole CRS

Table 4-12. AGS data groups with geometry established using two distinct Point geometries.

Data Group[1]	Data Group Description
CBRG	CBR Test – General
BRT	CBR Test
CHEM[2]	Chemical Tests
CHLK	Chalk Tests
CLSS	Classification Tests
CMPG	Compaction Test - General
CMPT	Compaction Tests
CNMT	Contamination and Chemical Testing
CONG	Consolidation Test - General
CONS	Consolidation Test
FRST	Frost Susceptibility
GAST[2]	Gas Constitutes
GRAD	Particle Size Distribution Analysis Data
MCVG	MCV Test – General
PTST	Laboratory Permeability Test
RELD	Relative Density Test
ROCK	Rock Testing
SHBG	Shear Box Testing - General
SHBT	Shear Box Testing
SUCT	Suction Tests
TNPC	Ten Percent Fines
TRIG	Triaxial Test - General
TRIX	Triaxial Test

1) These lab tests are positioned by AGS fields SAMP_TOP and SPEC_DPTH.
2) Data group will be deleted from the future version of AGS specifications.

Table 4-13. AGS data transfer format GRAD data group, used to store the data from gradation or sieve analysis.

No.	Heading	Unit	Description
1	HOLE_ID*		Exploratory hole or location equivalent
2	SAMP_TOP*	m	Depth to TOP of test sample
3	SAMP_REF*		Sample reference number
4	SAMP_TYPE*		Sample type
5	SPEC_REF*		Specimen reference number
6	SPEC_DPTH*	m	Specimen Depth
7	GRAD_SIZE*	mm	Sieve or particle size
8	GRAD_PERP	%	Percentage passing/finer
9	GRAD_TYPE		Grading analysis test type

* key field.

Table 4-14. Proposed GML encoding for GRAD data group, shown in Table 4-13. Each test location is defined using two points, one at the top of the sample used in the test (Samp_Top), and the other one at top of the specimen used in the test (Spec_Dpth).

Schema	Comments
<!-- == --> <xs:element name="Grad" type="app:GradType" substitutionGroup="gml:_Feature"/> <!-- == --> <xs:complexType name="GradType"> <xs:complexContent> <xs:restriction base="gml:AbstractFeatureType"> <xs:sequence> <xs:element name="Samp_Top" type="gml:PointPropertyType"/> <xs:element name="Samp_Ref" type="xs:string" minOccurs="0"/> <xs:element name="Samp_Type" type="xs:string" minOccurs="0"/> <xs:element name="Spec_Ref" type="xs:string" minOccurs="0"/> <xs:element name="Spec_Dpth" type="gml:PointPropertyType"/> <xs:element name="Grad_size" type="gml:MeasureType"/> <xs:element name="Grad_perp" type="gml:MeasureType"/> <xs:element name="Grad_type" type="xs:string" minOccurs="0"/> </xs:sequence> </xs:restriction> </xs:complexContent> </xs:complexType>	-top of sample (Samp_Top) -specimen depth (Spec_Dpth)

Table 4-15. AGS data groups where geometry is defined using a single Line geometry.

Data Group	Data Group Description	Geometry Field Heading	Geometry Field Description
?BKFL	Backfill Details	?BKFL_TOP	Depth to top of section
		?BKFL_BASE	Depth to base of section
CHIS	Chiseling Details	CHIS_FROM	Depth at start of chiseling
		CHIS_TO	Depth at end of chiseling
CORE	Rotary Core Information	CORE_TOP	Depth to top of core run
		CORE_BOT	Depth to bottom of core run
DETL	Stratum Detail Descriptions	DETL_TOP	Depth to top of detail description
		DETL_BASE	Depth to base of detail description
DISC	Discontinuity Data	DISC_TOP	Depth to top in hole, or distance to start on traverse, of discontinuity zone, or discontinuity
		DISC_BASE	Depth to base in hole, or distance to start on traverse, of discontinuity zone, or discontinuity
FLSH	Rotary Core Flush Details	FLSH_FRM	Depth to top of flush zone
		FLSH_TO	Depth to bottom of flush zone
FRAC	Fracture Spacing	FRAC_TOP	Depth to top in hole, or distance to start on traverse, of the zone
		FRAC_BASE	Depth to base in hole, or distance to start on traverse, of the zone
GEOL	Stratum Descriptions	GEOL_TOP	Depth to the top of stratum
		GEOL_BASE	Depth to the base of stratum
?HDPH	Depth Related Hole Information	?HDPH _TOP	Depth to top of section
		?HDPH_BASE	Depth to base of section
IPRM	In Situ Permeability Test	IPRM_TOP	Depth to top of test zone
		IPRM_BASE	Depth to base of test zone
SAMP	Sample Reference Information	SAMP_TOP	Depth to top of sample
		SAMP_BASE	Depth to base of sample
WETH	Weathering Grades	WETH_TOP	Depth to top of weathering subdivision
		WETH_BASE	Depth to base of weathering subdivision

Table 4-16. AGS data transfer format GEOL data group, which is used for soil strata data.

No.	Heading	Unit	Description
1	HOLE_ID*		Exploratory hole or location equivalent
2	GEOL_TOP*	m	Depth to the TOP of stratum
3	GEOL_BASE*	m	Depth to the BASE of description
4	GEOL_DESC		General description of stratum
5	GEOL_LEG		Legend code
6	GEOL_GEOL		Geology code
7	GEOL_GEO2		Second Geology code
8	GEOL_STAT		Stratum reference shown on trial pit or traverse sketch
9	FILE_FSET		Associated file reference

* key field.

Table 4-17. Proposed GML encoding for GEOL data group, shown in Table 4-16.

```
Schema                                                                  Comments

<!-- ======================================================== -->
<xs:element name="Geol" type="app:GeolType"
     substitutionGroup="gml:_Feature"/>
<!-- ======================================================== -->
<xs:complexType name="GeolType">
 <xs:complexContent>
  <xs:restriction base="gml:AbstractFeatureType">
   <xs:sequence>
    <xs:element name="Geol_Top" type="gml:PointPropertyType"            -top of
     minOccurs="0"/>                                                    stratum
    <xs:element name="Geol_Base" type="gml:PointPropertyType"           -bottom of
     minOccurs="0"/>                                                    stratum
    <xs:element name="Geol_Axis" type="gml:LineStringPropertyType"      -observed
     minOccurs="0"/>                                                    axis of
    <xs:element name="Geol_Desc" type="xs:string" minOccurs="0"/>       stratum
    <xs:element name="Geol_Leg" type="xs:string" minOccurs="0"/>
    <xs:element name="Geol_Geol" type="xs:string" minOccurs="0"/>
    <xs:element name="Geol_Geo2" type="xs:string" minOccurs="0"/>
    <xs:element name="Geol_Stat" type="xs:string" minOccurs="0"/>
    <xs:element name="File_Fset" type="app:FileType"
     minOccurs="0"/>
   </xs:sequence>
  </xs:restriction>
 </xs:complexContent>
</xs:complexType>
```

Table 4-18. AGS data transfer format SAMP data group, which is used for sample data.

No.	Heading	Unit	Description
1	HOLE_ID*		Exploratory hole or location equivalent
2	SAMP_TOP*	m	Depth to the TOP of sample
3	SAMP_REF*		Sample reference number
4	SAMP_TYPE*		Sample type
5	SAMP_DIA	mm	Sample diameter
6	SAMP_BASE	m	Depth to the BASE of description
7	SAMP_DESC		Sample description
8	SAMP_UBLO		Number of blows required to drive sampler
9	SAMP_REM		Sample remarks
10	SAMP_DATE	dd/mm/yyyy	Data sample taken
11	SAMP_TIME	Hhmmss	Time sample taken
12	SAMP_BAR	kPa	Barometer pressure at time of sampling
13	SAMP_WDEP	m	Depth to water below ground surface at time of sampling
14	SAMP_TEMP	DegC	Sample temperature at time of sampling
15	SAMP_PRES	kPa	Gas pressure (above barometric)
16	SAMP_FLOW	l/min	Gas flow
17	?SAMP_PREP		Details of sample preparation
18	GEOL_STAT		Stratum reference shown on trial pit or traverse sketch
19	FILE_FSET		Associated file reference

* key field.

Table 4-19. Proposed GML encoding for SAMP data group, shown in Table 4-17.

Schema	Comments
<!-- === -->	
<xs:element name="Samp" type="app:SampType"	
substitutionGroup="gml:_Feature"/>	
<!-- === -->	
<xs:complexType name="SampType">	
<xs:complexContent>	
<xs:restriction base="gml:AbstractFeatureType">	
<xs:sequence>	
<xs:element name="Samp_Top" type="gml:PointPropertyType"	-top of
minOccurs="1"/>	sample
<xs:element name="Samp_Ref" type="xs:string"	
minOccurs="1"/>	
<xs:element name="Samp_Type" type="gml:CodeType"	
minOccurs="1"/>	
<xs:element name="Samp_Dia" type="gml:LengthType"	
minOccurs="0"/>	
<xs:element name="Samp_Base" type="gml:PointPropertyType"	-bottom of
minOccurs="0"/>	sample
<xs:element name="Samp_Axis"	-axis of
type="gml:LineStringPropertyType" minOccurs="0"/>	sample
<xs:element name="Samp_Desc" type="xs:string"	
minOccurs="0"/>	
<xs:element name="Samp_Ublo" type="xs:string"	
minOccurs="0"/>	
<xs:element name="Samp_Rem" type="xs:string"	
minOccurs="0"/>	
<xs:element name="Samp_Date" type="xs:date"	
minOccurs="0"/>	
<xs:element name="Samp_Time" type="xs:time"	
minOccurs="0"/>	
<xs:element name="Samp_Bar" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="Samp_Wdep" type="gml:LengthType"	
minOccurs="0"/>	
<xs:element name="Samp_Temp" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="Samp_Pres" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="Samp_Flow" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="Samp_Prep" type="xs:string"	
minOccurs="0"/>	
<xs:element name="Geol_Stat" type="xs:string"	
minOccurs="0"/>	
<xs:element name="File_Fset" type="app:FileType"	
minOccurs="0"/>	
</xs:sequence>	
</xs:restriction>	
</xs:complexContent>	
</xs:complexType>	

Table 4-20. AGS data transfer format STCN data group (Static Cone Penetration Test), used for CPT results.

No.	Heading	Unit	Description
1	HOLE_ID*		Exploratory hole or location equivalent
2	STCN_DPTH*		Depth of result for static cone test
3	STCN_TYP		Cone test type
4	STCN_REF		Cone identification reference
5	STCN_RES	MN/m2	Cone resistance
6	STCN_FRES	kN/m2	Local unit side friction resistance
7	STCN_PWP1	kN/m2	Porewater pressure
8	STCN_PWP2	kN/m2	Second porewater pressure
9	STCN_PWP3	kN/m2	Third porewater pressure
10	STCN_CON	uS/cm	Conductivity
11	STCN_TEMP	DegC	Temperature
12	STCN_PH		pH reading
13	STCN_SLP1	deg	Slope Indicator no. 1
14	STCN_SLP2	deg	Slope Indicator no. 2
15	STCN_REDX	mV	Redox potential reading
16	STCN_FFD	%	Fluorescence intensity
17	STCN_PMT	counts/s	Photo-multiplier tube reading
18	STCN_PID	uV	Photo ionization detector reading
19	STCN_FID	uV	Flame ionization detector reading
20	FILE_FSET		Associated file reference

* key field.

Table 4-21. Proposed GML encoding for STCN data group, shown in Table 4-20.

Schema	Comments
<!-- ==-->	
<xs:complexType name="StcnType">	
<xs:annotation>	
<xs:documentation>STCN Test</xs:documentation>	
</xs:annotation>	
<xs:complexContent>	
<xs:restriction base="gml:AbstractCoverageType">	
<xs:sequence>	
<xs:element ref="app:Stcn_Depths"/>	-coverage
<xs:element ref="app:Stcn_Values"/>	domain
<xs:element ref="gml:coverageFunction" minOccurs="0"/>	-coverage range
</xs:sequence>	-coverage
</xs:restriction>	function
</xs:complexContent>	
</xs:complexType>	
<xs:element name="Stcn_Depths" type="gml:MultiPointDomainType"	-MultiPoint
substitutionGroup="gml:domainSet"/>	geometry type
<xs:element name="Stcn_Values" type="app:Stcn_ValuesType"	
substitutionGroup="gml:rangeSet"/>	
<!-- ==-->	
<xs:complexType name="Stcn_ValuesType">	
<xs:complexContent>	
<xs:restriction base="gml:RangeSetType">	
<xs:sequence>	
<xs:element ref="app:Stcn_Val" maxOccurs="unbounded"/>	
</xs:sequence>	
</xs:restriction>	
</xs:complexContent>	
</xs:complexType>	
<!-- ==-->	
<xs:element name="Stcn_Val" type="app:StcnValType"/>	-cone
<xs:complexType name="StcnValType">	penetration
<xs:sequence>	test measured
<xs:element name="STCN_TYPE" type="xs:string"	values
minOccurs="0"/>	
<xs:element name="ATCN_REF" type="xs:string"	
minOccurs="0"/>	
<xs:element name="STCN_RES" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_FRES" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_PWP1" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_PWP2" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_PWP3" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_CON" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_TEMP" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_PH" type="gml:MeasureType"	
minOccurs="0"/>	
<xs:element name="STCN_SLP1" type="gml:AngleType"	
minOccurs="0"/>	

Table 4-21. Continued

Schema	Comments
<xs:element name="STCN_SLP2" type="gml:AngleType" minOccurs="0"/>	
<xs:element name="STCN_REDX" type="gml:MeasureType" minOccurs="0"/>	
<xs:element name="STCN_FFD" type="gml:MeasureType" minOccurs="0"/>	
<xs:element name="STCN_PMT" type="gml:MeasureType" minOccurs="0"/>	
<xs:element name="STCN_PID" type="gml:MeasureType" minOccurs="0"/>	
<xs:element name="STCN_FID" type="gml:MeasureType" minOccurs="0"/>	
<xs:element name="FILE_FSET" type="app:FileType" minOccurs="0"/>	
</xs:sequence>	
</xs:complexType>	

Table 4-22. Example of encoding for a GML Dictionary containing units of measurement.

```
<gml:Dictionary gml:id="unitsDictionary">
 ...
 <gml:dictionaryEntry>
  <gml:DefinitionCollection gml:id="SIBaseUnits">
   <gml:description>The Base Units from the SI units system.</gml:description>
   <gml:name>SI Base Units</gml:name>
   <gml:dictionaryEntry>
    <gml:BaseUnit gml:id="metre">
     <gml:description>...</gml:description>
     <gml:name
      codeSpace="http://www.bipm.fr/en/3_SI/base_units.html">metre</gml:name>
     <gml:name xml:lang="en/US">meter</gml:name>
     <gml:quantityType>length</gml:quantityType>
     <gml:catalogSymbol
     codeSpace="http://www.bipm.fr/en/3_SI/base_units.html">m</gml:catalogSymbo
     l>
     <gml:unitsSystem xlink:href="http://www.bipm.fr/en/3_SI"/>
    </gml:BaseUnit>
   </gml:dictionaryEntry>
   ...
  </DefinitionCollection>
 </gml:dictionaryEntry>
</gml:Dictionary>
```

Table 4-23. Comparison between GeotechGML and DIGGS GML compatible data formats for geotechnical data.

Test	GeotechGML	DIGGS	Comments
Element and property names	Follow AGS abbreviations	Full English name (no abbreviation)	GeotechGML uses AGS data dictionary whenever possible, Diggs uses full English words for each element or property
Schema complexity	Simpler schema	More complex schema	Diggs has more levels of element extensions comparing to GeotechGML, the scope of schema is more comprehensive and includes SIGs (pile driving, etc.)
“Sample” encoding	Child element of Hole feature	Same level as Hole	In GeotechGML samples are under borehole feature and can be exchanged together in one operation but in Diggs they are parallel elements so finding all samples of a borehole needs a query operation, however in Diggs samples are independent elements.
Laboratory tests	Child elements of sample and hole	Same level as hole and sample elements	Since in GeotechGML the samples are under holes, the lab test under samples are hierarchically under Hole elements, however in Diggs they are under Samples but not under Hole elements.
CPT Test	Encoded using Coverage feature	Encoded using Table element	Coverage is a spatial feature, table is more compact but not a spatial feature

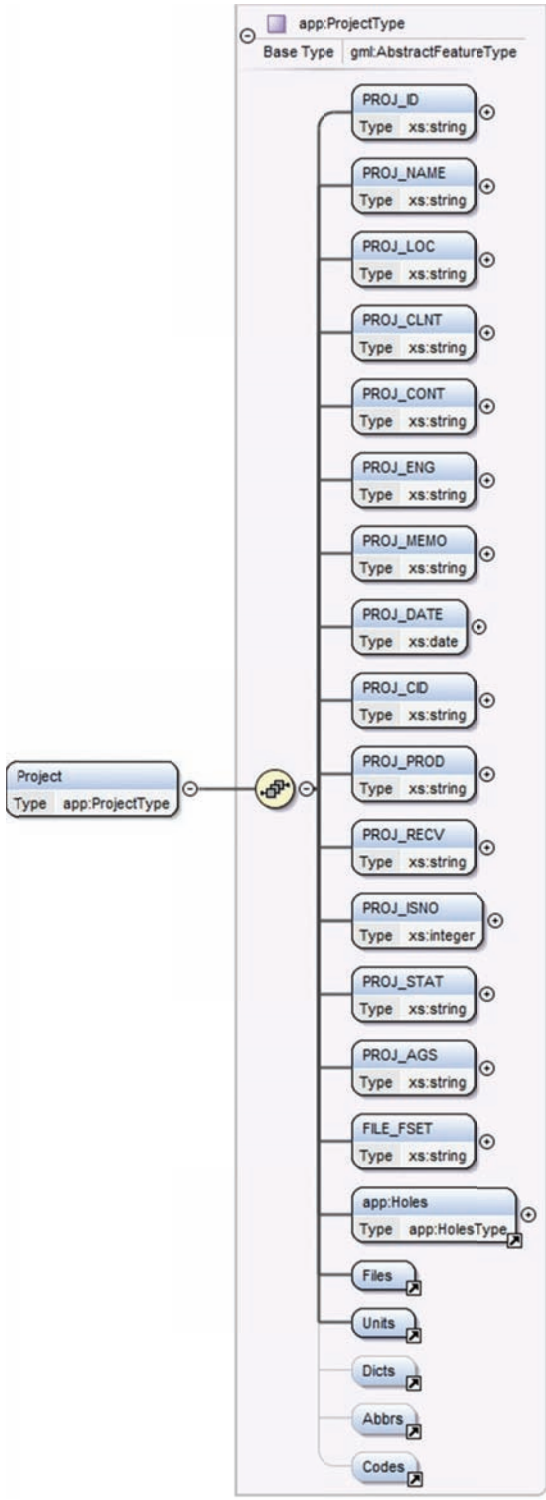


Figure 4-1. Design diagram of GML feature Project, corresponding to AGS PROJ data group.

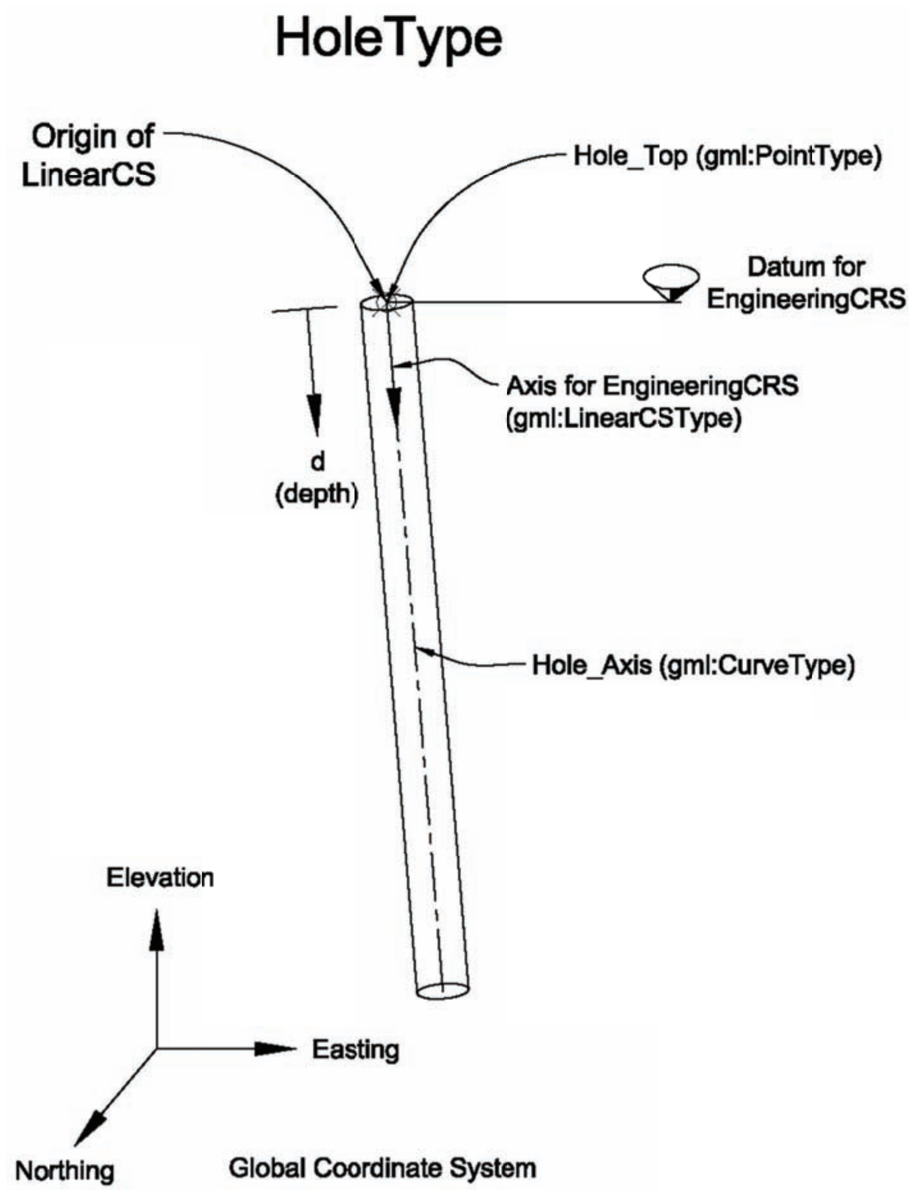


Figure 4-2. Geometric representation of GML feature HOLE, based on AGS HOLE data group.

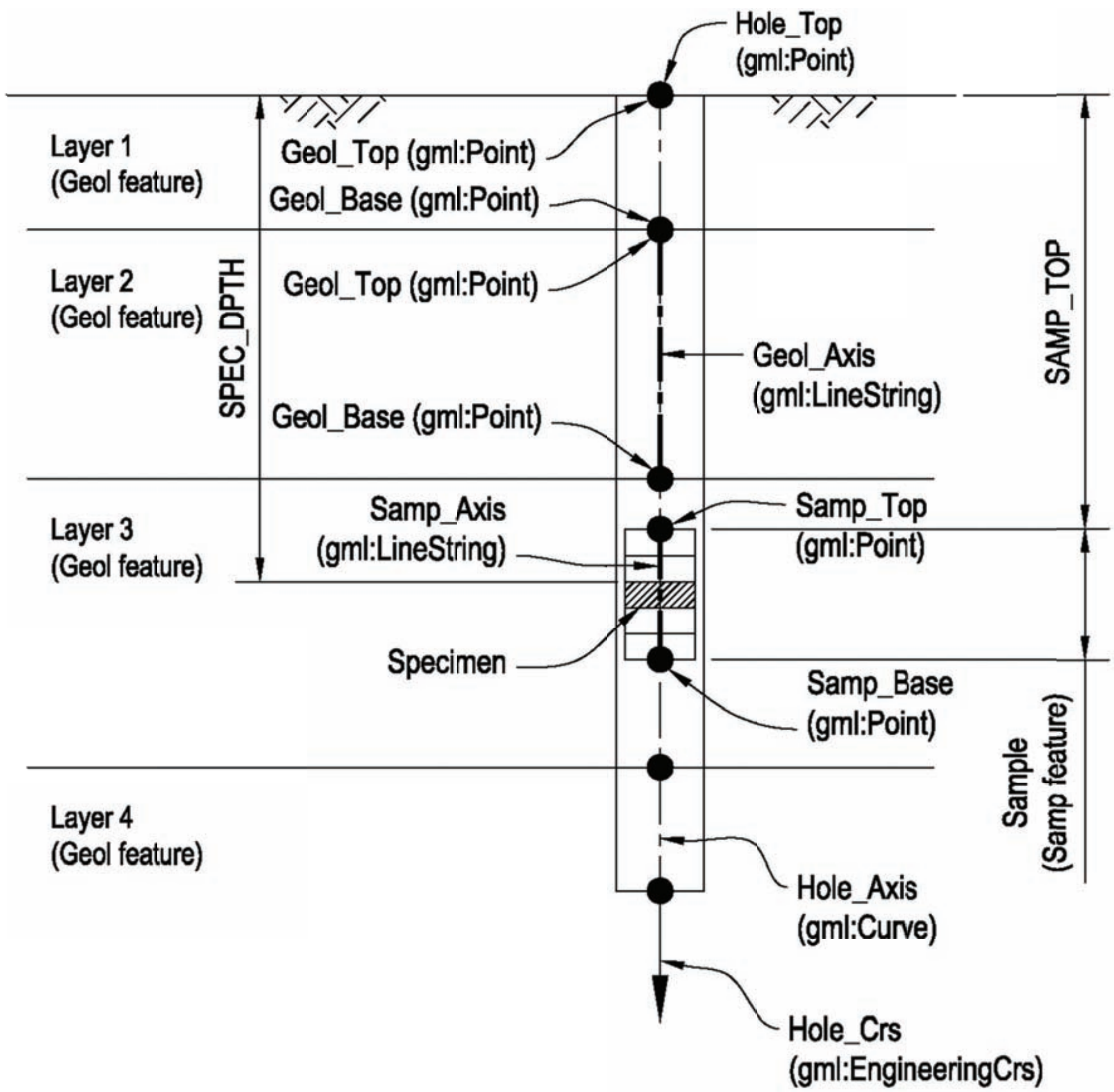


Figure 4-3. Schematic diagram of spatial referencing method in a borehole. Soil layers, samples and specimens are spatially referenced relative to the top of the borehole, using a local coordinate system (Hole_Crs) that is anchored to a point at hole top (Hole_Top) and measured along the borehole axis (Hole_Axis).

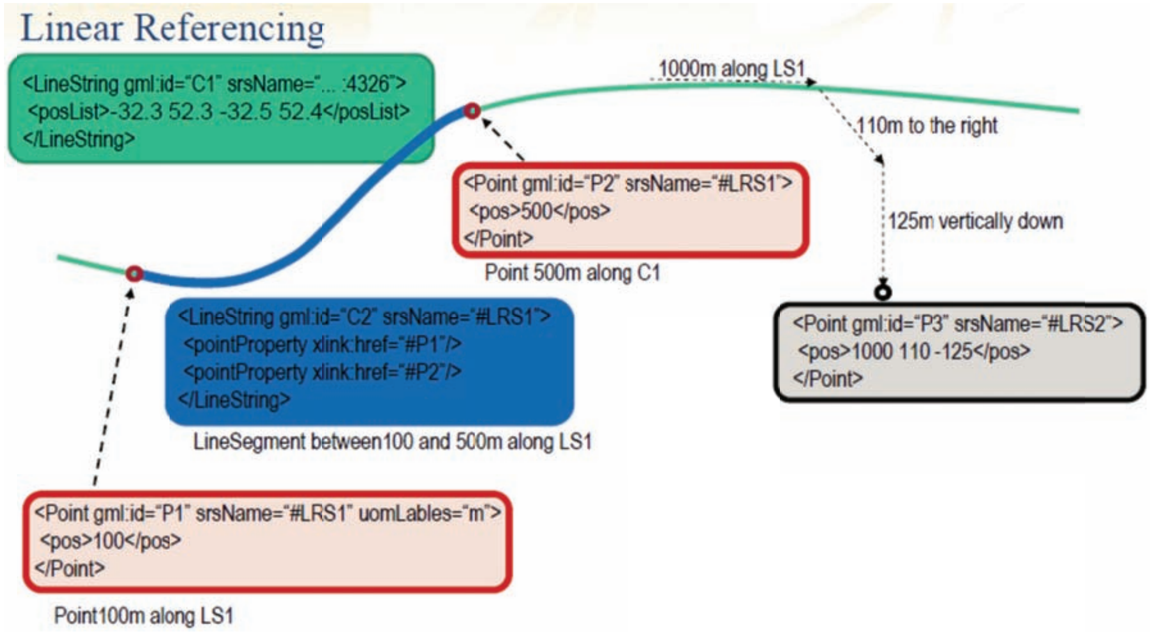


Figure 4-4. GML 3.3 proposed linear referencing methodology, which is similar to the methodology described in Figure 4-3.

(from Burggraf, 2010)

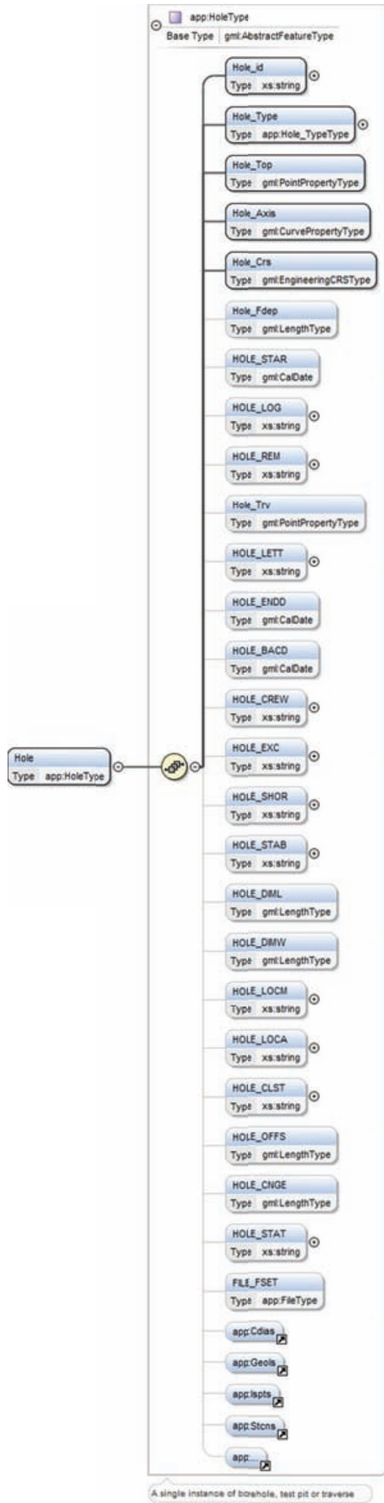


Figure 4-5. Design diagram of the GML feature Hole, based on the AGS HOLE data group (Table 4-2).

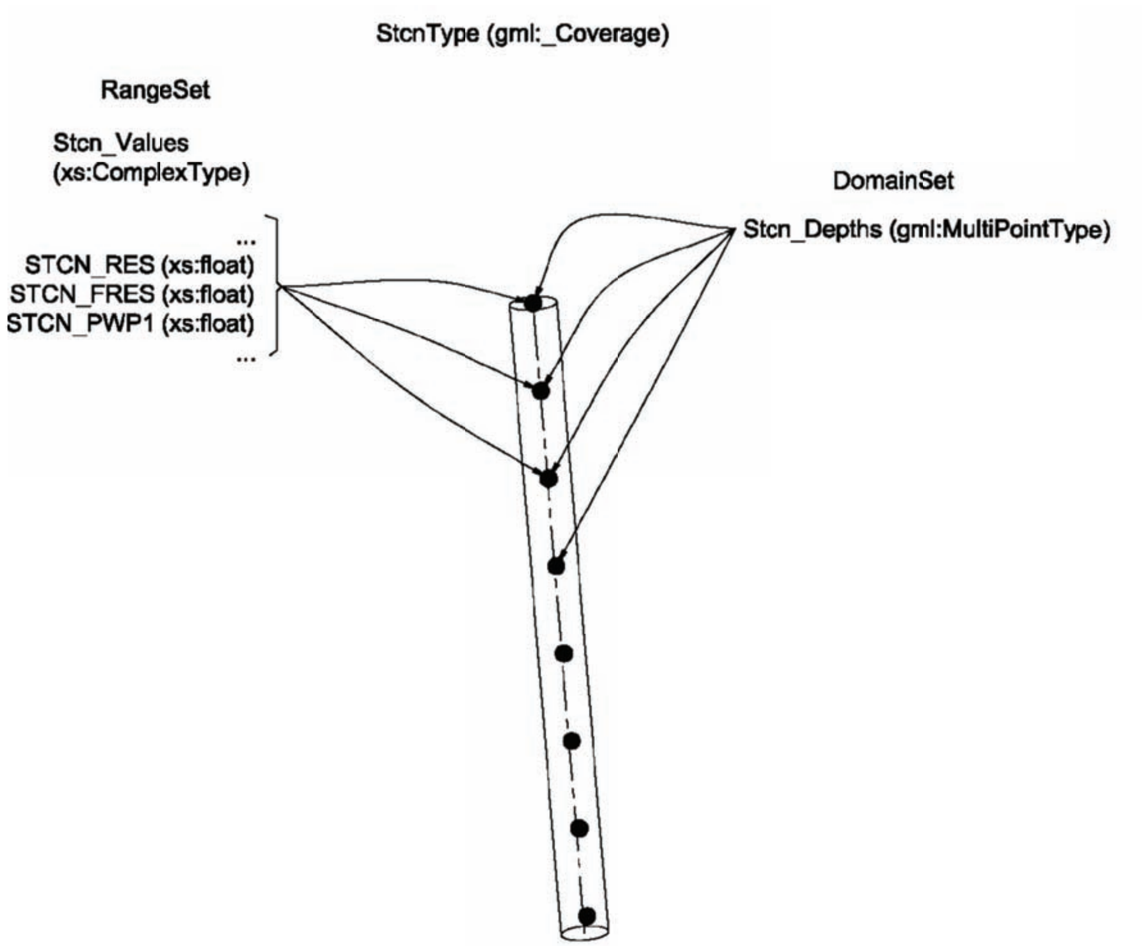


Figure 4-6. Geometry of STCN feature used for CPT data, encoded based on AGS STCN data group (Table 4-20).

Figure 4-7. Illustration of a sample GeotechGML file with two boreholes overlaid on Microsoft Virtual Earth satellite map in Gaia 3.0 GML viewer. There are a number of programs that support GML visualization but currently full GML support is limited, mainly due to flexibility of the data format and lack of rigid structure.

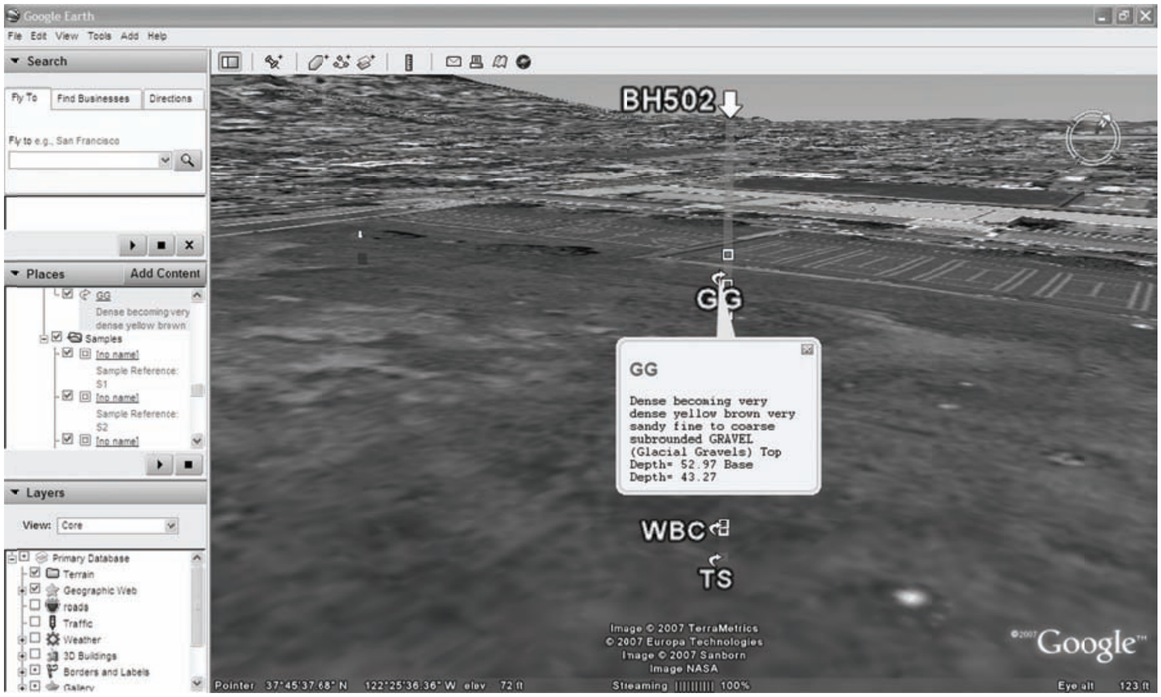


Figure 4-8. A sample GeotechGML file with two boreholes transformed to KML and visualized in Google Earth. KML is an XML based data format and transformation can be accomplished using existing tools such as XSLT or DOM (Data Object Model).

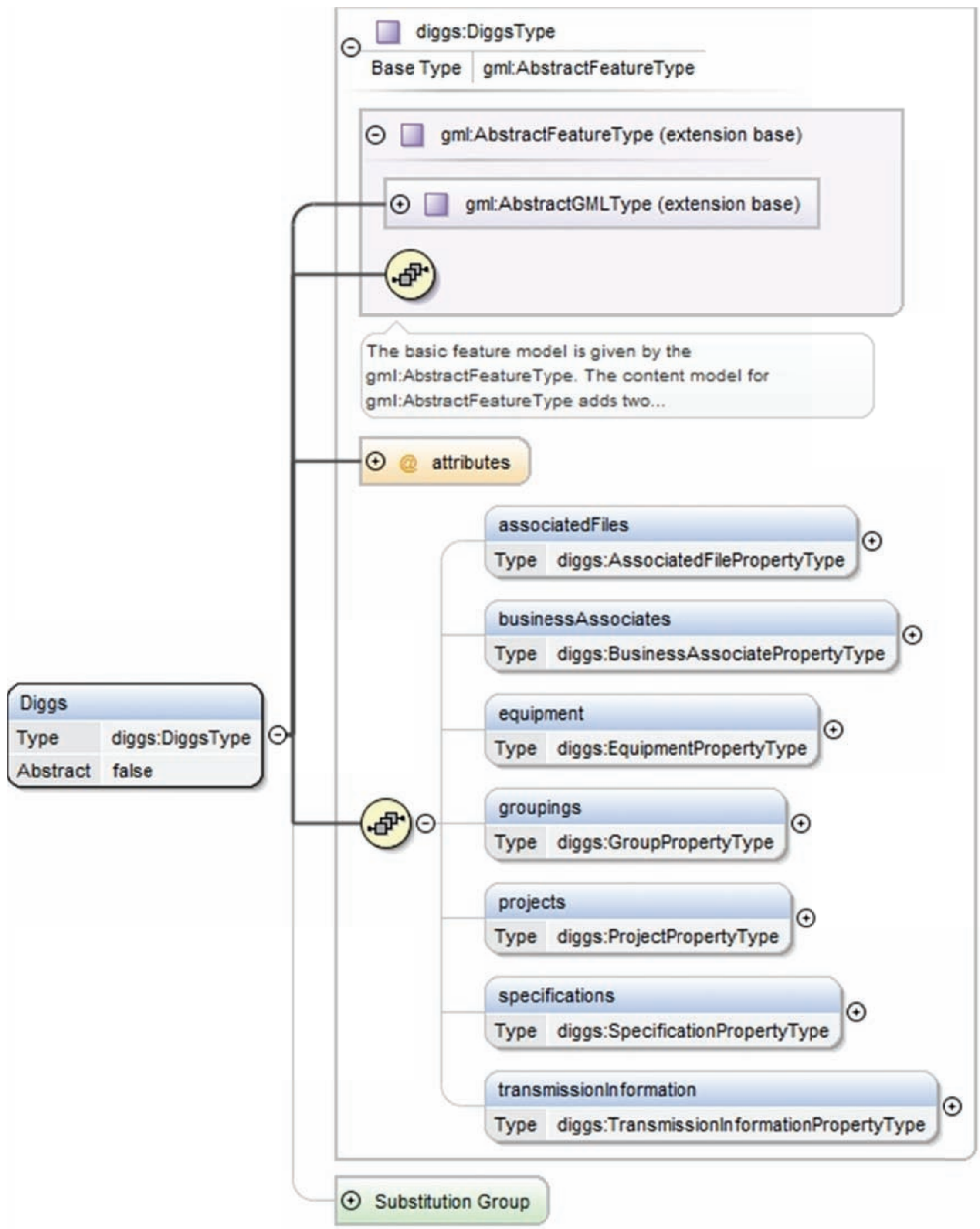


Figure 4-9. Diagram of DIGGS root element, "diggs:DiggsType", showing the base element and child elements. "Diggs" is an extension of "AbstractGMLType". This diagram is created using oXygen XML editor from the schemas developed by DIGGS (2010).

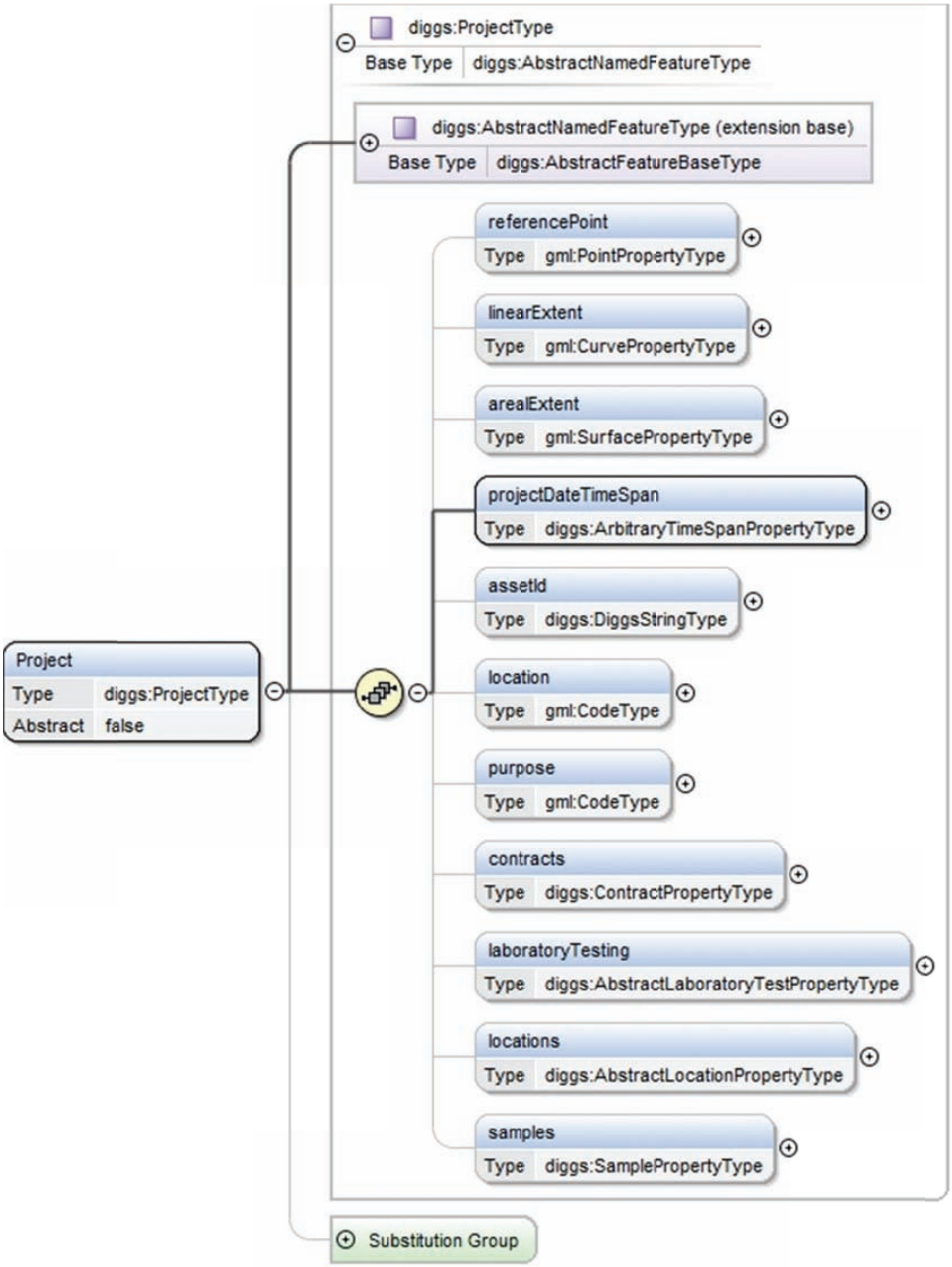


Figure 4-10. Diagram of "project" element (diggs:ProjectType). Boreholes are defined within "locations" element. The diagram also shows "laboratoryTesting" and "samples" elements which are in the same level as "locations". This diagram is generated using oXygen XML editor from the schemas developed by DIGGS (2010).

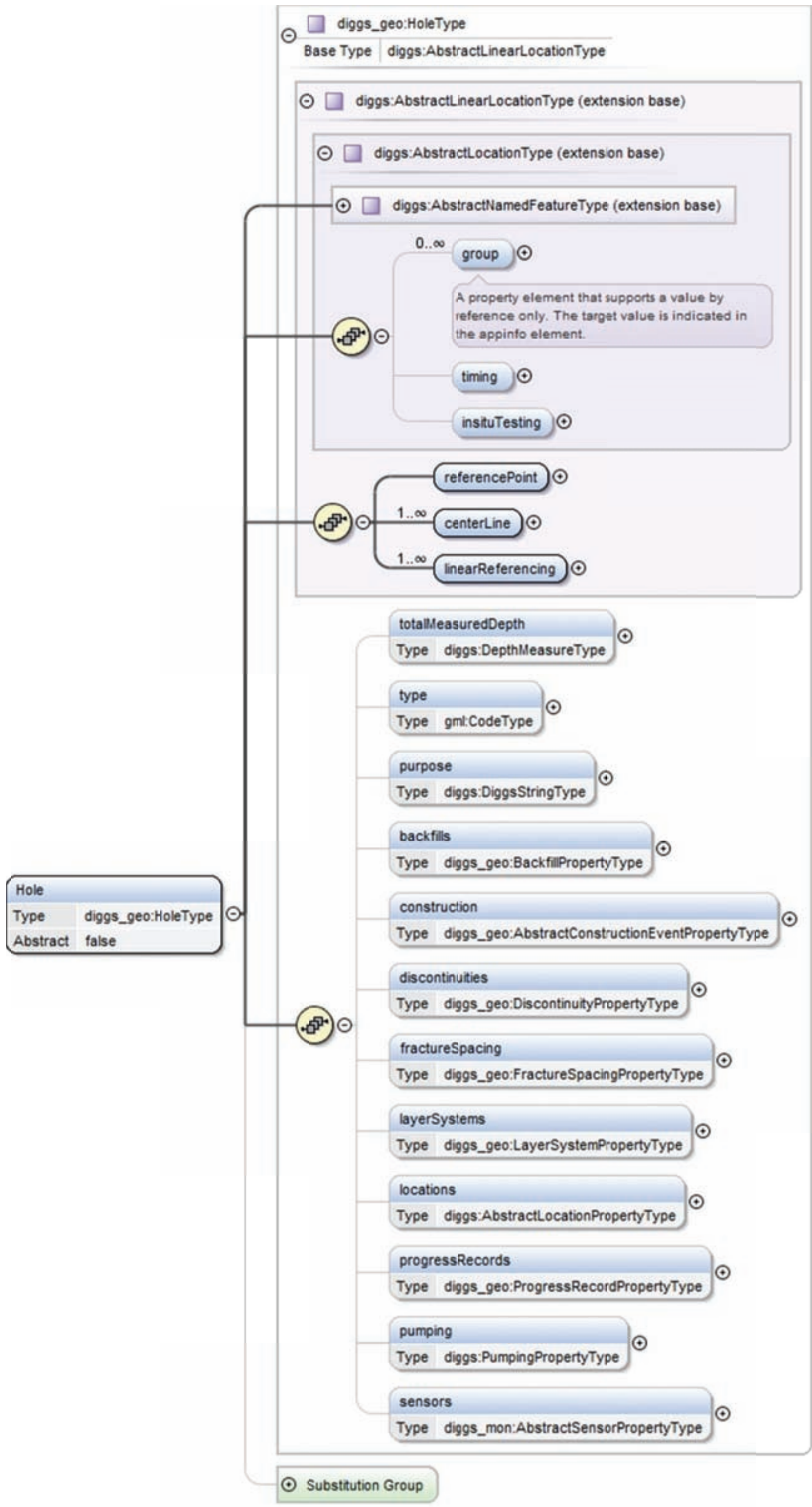


Figure 4-11. Diagram of Hole element schema (diggs_geo:HoleType), generated using oXygen XML editor from the schemas developed by DIGGS (2010).

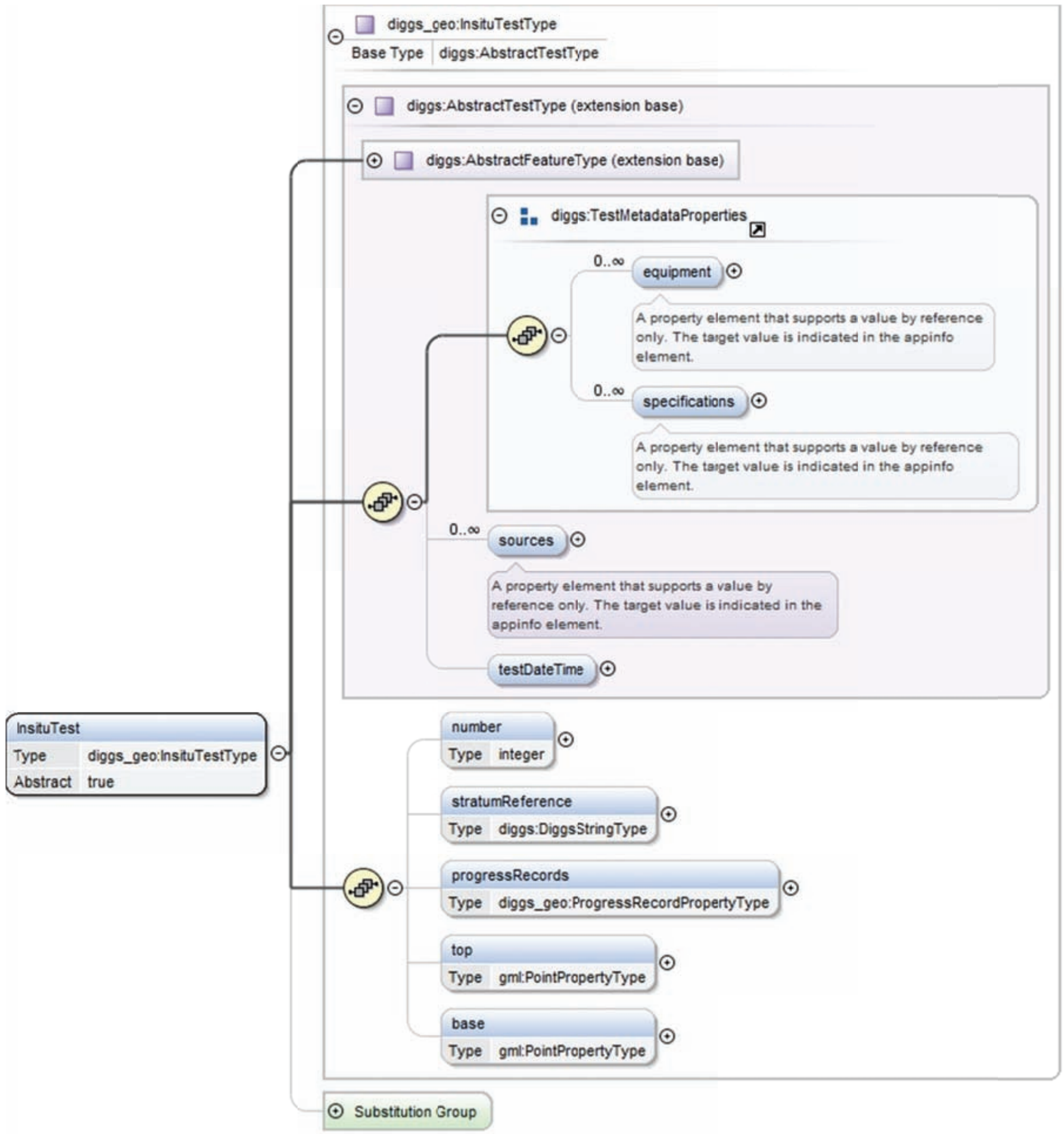


Figure 4-12. Diagram of "InsituTest" element schema (diggs_geo:InsituTestType), generated using oXygen XML editor from the schemas developed by DIGGS (2010).

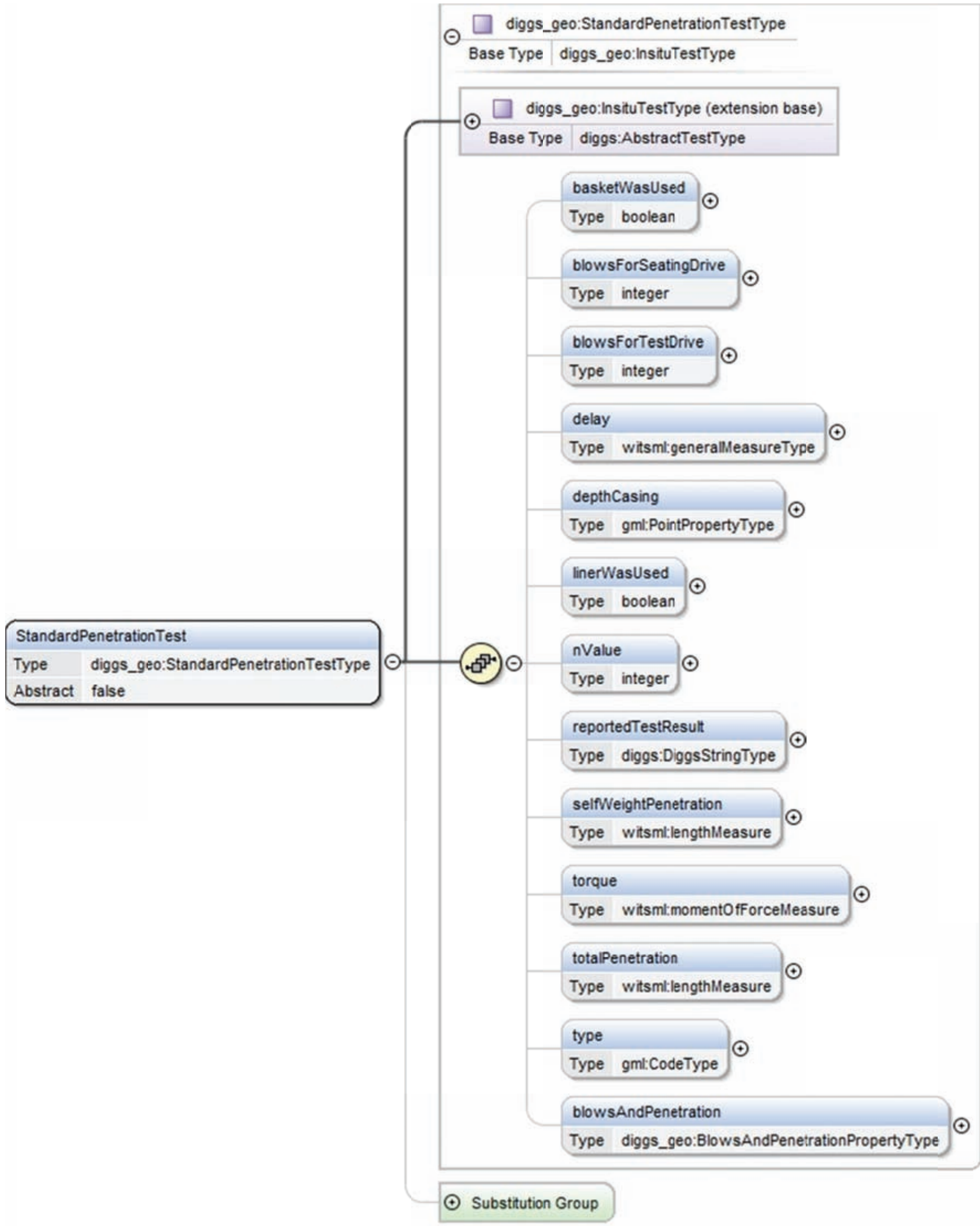


Figure 4-13. Diagram of "StandardPenetrationTest" element schema (diggs_geo: StandardPenetrationTestType, generated using oXygen XML editor from the schemas developed by DIGGS (2010).

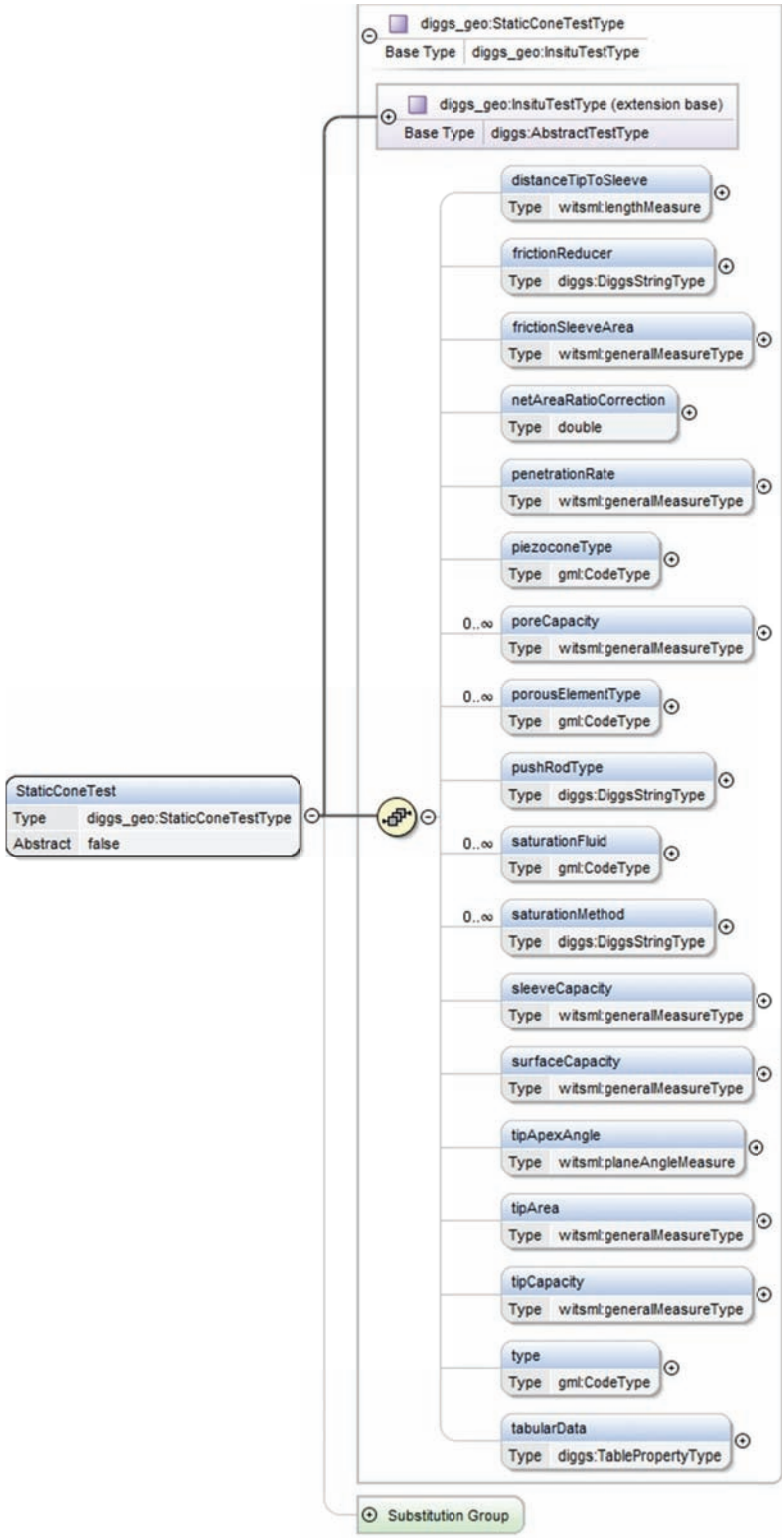


Figure 4-14. Diagram of "StaticConeTest" element schema (diggs_geo:StaticConeTestType), generated using oXygen XML editor from the schemas developed by DIGGS (2010).

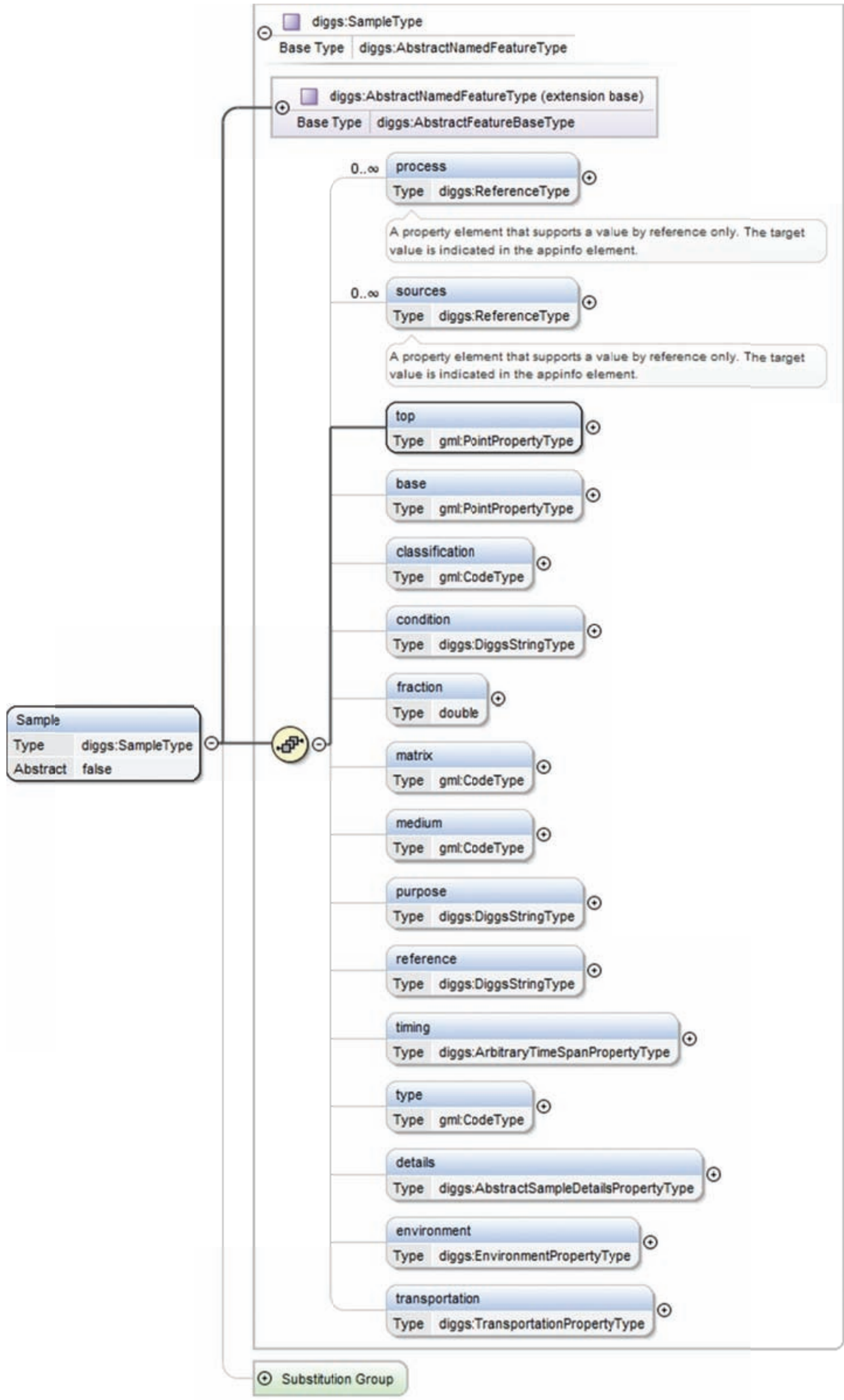


Figure 4-15. Diagram of "Sample" element schema (diggs:SampleType), generated using oXygen XML editor from the schemas developed by DIGGS (2010).

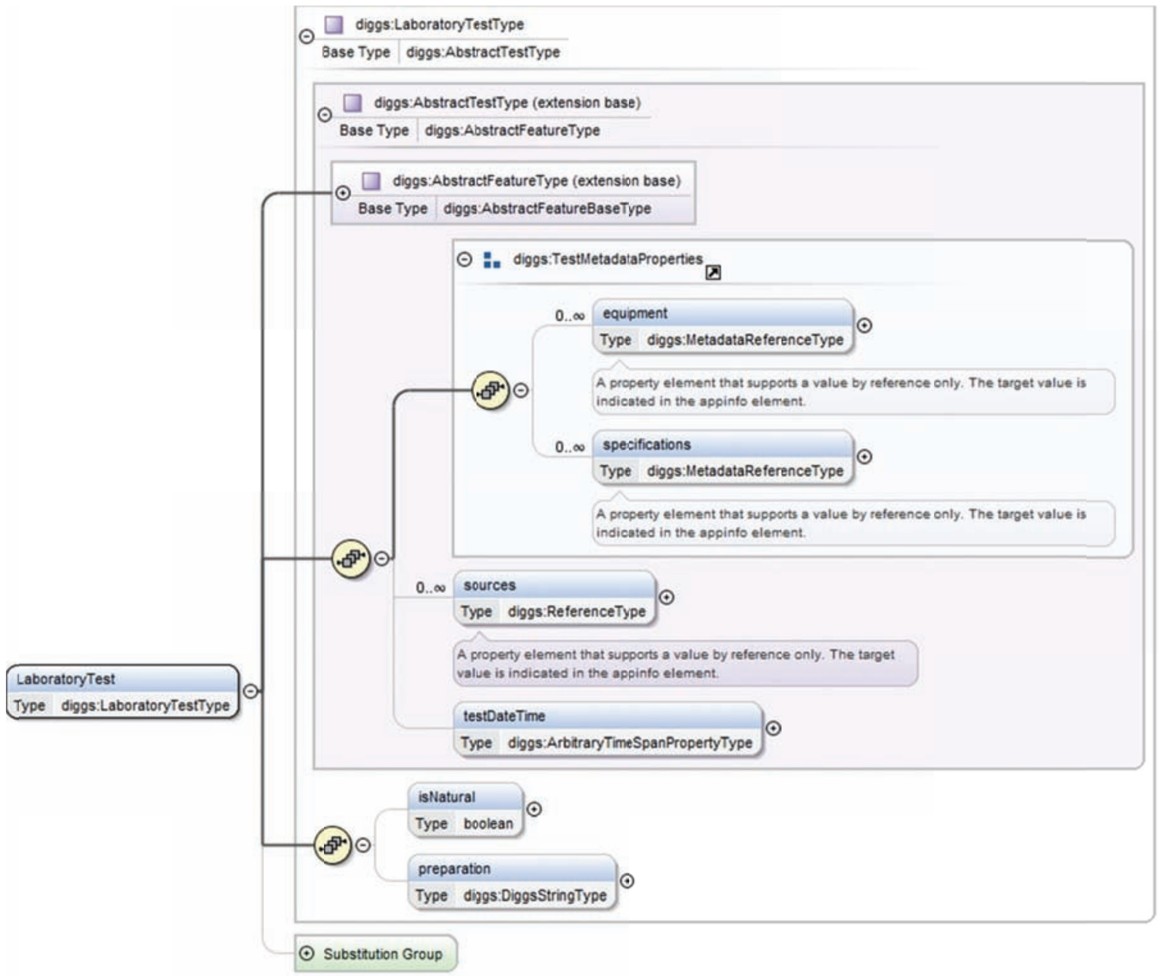


Figure 4-16. Diagram of "LaboratoryTest" element schema (diggs:LaboratoryTestType), generated using oXygen XML editor from the schemas developed by DIGGS (2010).

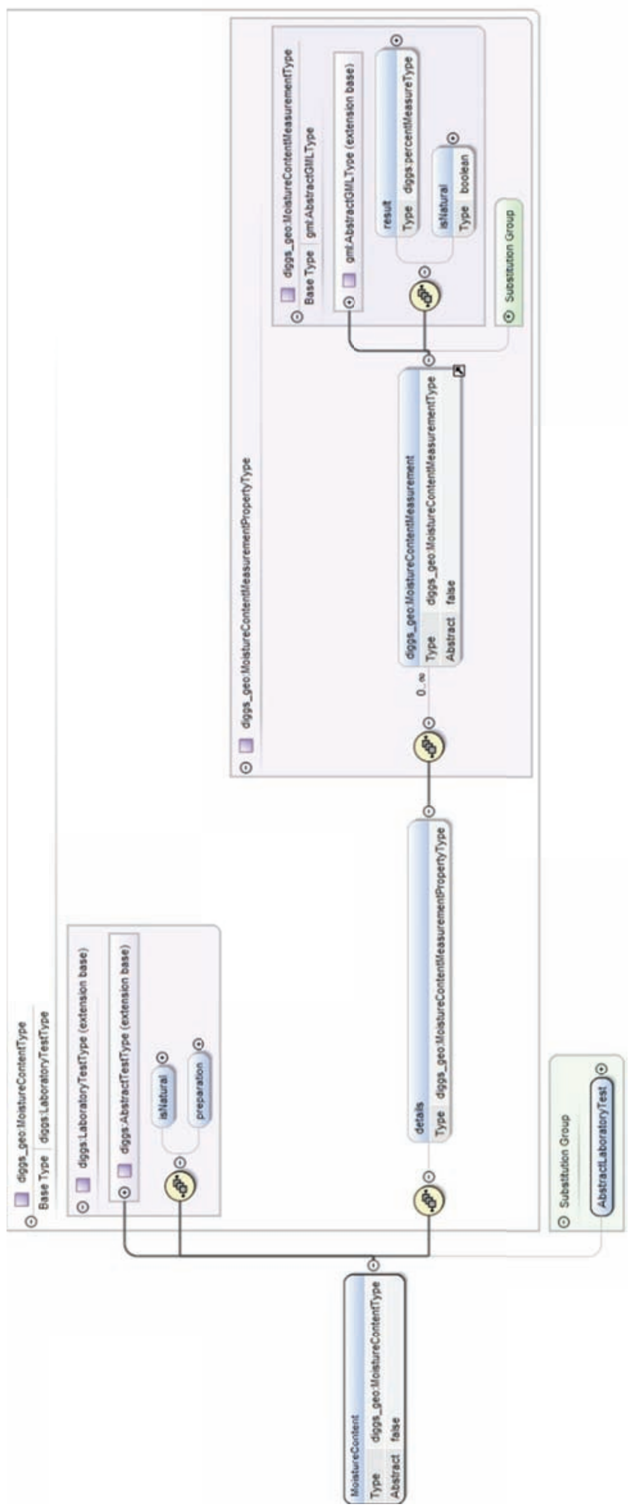


Figure 4-17. Diagram of "MoistureContent" element schema (diggs_geo:MoistureContentType), generated using oXygen XML editor from the schemas developed by DIGGS (2010).

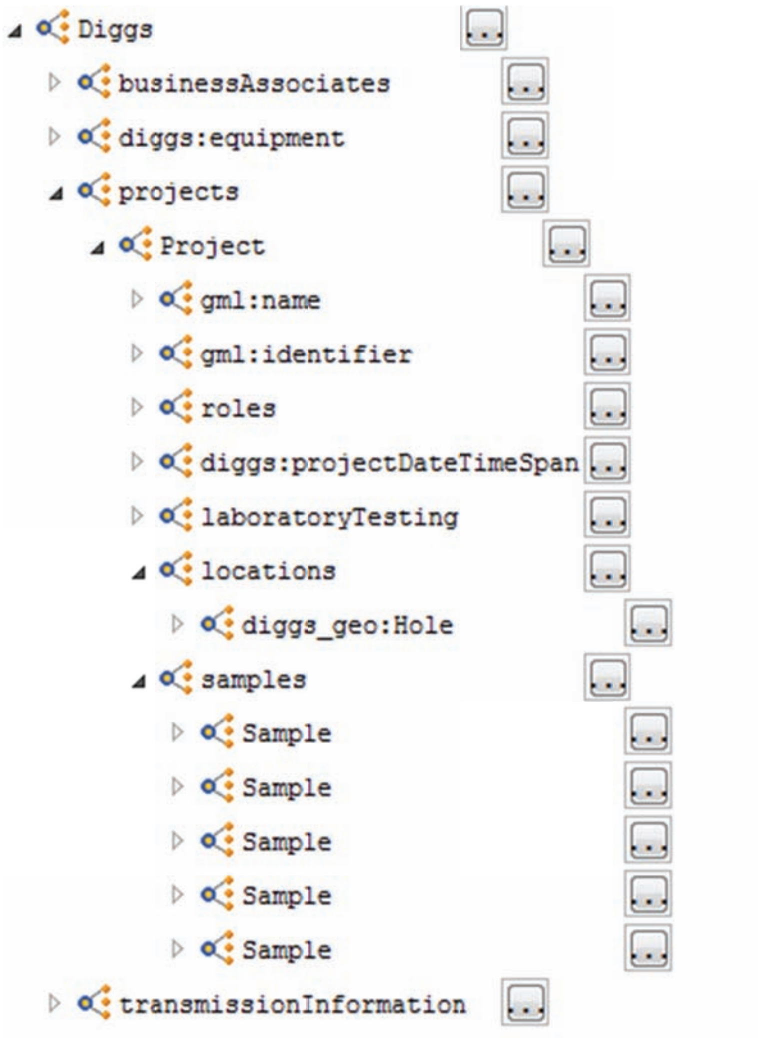


Figure 4-18. Schematic structure of a DIGGS Instance file. The diagram shows the high-level elements of the data format, Including "businessAssociates", "equipment", "projects" and "transmissionInformation".

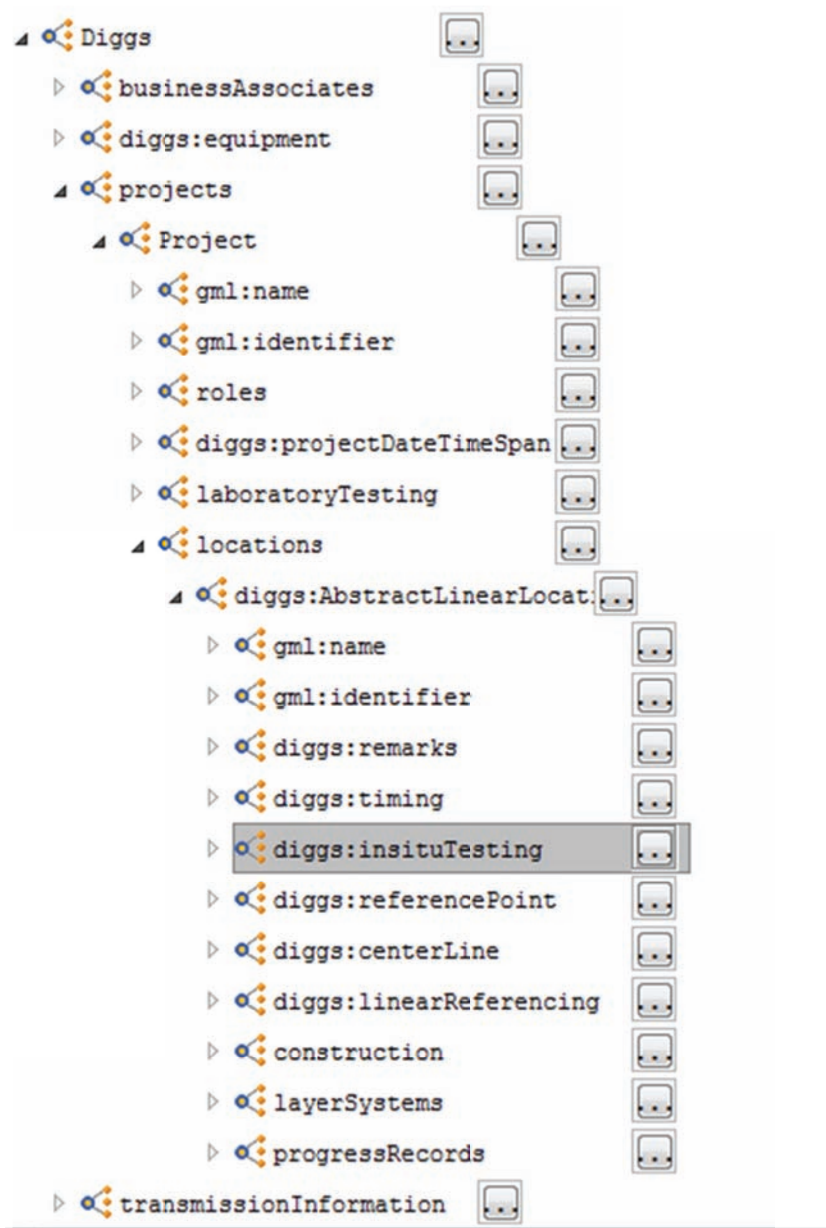


Figure 4-19. Schematic structure of DIGGS GML-compatible data format for geotechnical data, showing the businessAssociates, equipment, projects and transmissionInformation data. This example shows in-situ and laboratory tests.

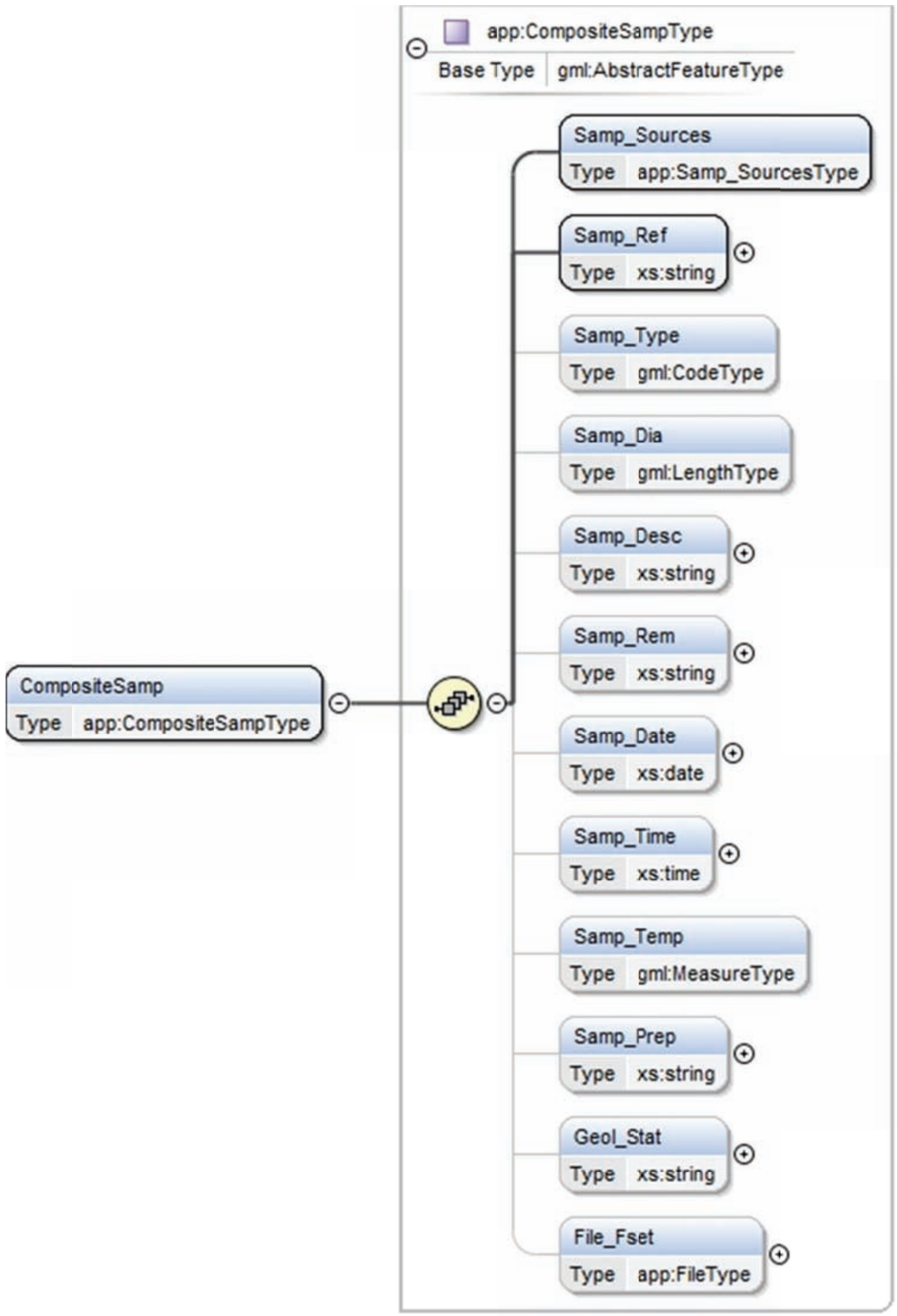


Figure 4-20. Design diagram of the *CompositeSamp* feature. This feature is used for composite samples that comprise of soils from multiple samples taken from one or several boreholes.

Chapter 5. Implementation and Case Study

In this chapter the implementation of the proposed GML model is demonstrated through several pilot applications that are developed based on the concepts discussed in the previous chapters. Furthermore, through a case study application of the proposed architecture for real-world infrastructure studies is demonstrated. The main objective of this chapter is to validate the concepts discussed throughout this research, rather than develop fully functional applications for practical purposes.

5.1 Implementation in a Geotechnical Data Management and Distribution Architecture

The GML data format is typically integrated in a Web-based data management and dissemination architecture. Beside the data format, a successful implementation depends on the supporting services that allow the users to find, explore and retrieve the data. Additional services are needed for administrators to insert the data to the data base and modify the existing data if needed.

Figure 5-1 shows the architecture of a Web-based distributed geotechnical database management and dissemination system. The diagram shows how several servers disseminate the data and multiple clients use the server for different purposes. Each server provides multiple Web services for uploading, downloading and visualizing data.

This architecture was used in the Geotechnical Information Management and Exchange (GIME) project (Zimmerman et al., 2006). GIME Web service was developed by the ITR group at the University of Southern California.

This architecture uses several archive servers for storing the data. Each archive server runs a database with support for XML and spatial data. Several Web services are shown in this architecture. The file storage web service (FSWS) is used to store the data in a temporary database before it is transferred to the main database. The Web feature service (WFS) is the standard OGC protocol for serving GML feature data through the Internet. In addition to WFS, other services can act on the same database. In this example, the query and exchange Web service (QEWS) is used for query and retrieval of data, and the visualization Web service (VWS) is used to visualize data by converting it to a scalable vector graphics (SVG) format.

Client applications use these Web services for different purposes. Some of these applications are shown in Figure 5-1 and include applications to get borehole data, liquefaction analysis, site response analysis and Web map service (WMS) applications.

All these client side applications use one of these services (WFS, QEWS, and WMS, etc.) for retrieving GML data containing geotechnical information. The rest of the analyses and visualization is performed within the client computer.

Hereafter, the pilot applications for site-response analysis and liquefaction analysis are discussed. Section 5.2 reviews the background and theory behind the site-response analysis and the methodology for performing such analysis using GML-encoded geotechnical data. Section 0 provides similar discussion for liquefaction hazard analysis using GML-encoded SPT and CPT data. Section 5.4 covers a case study involving REDARS analysis of earthquake hazard evaluation for a transportation network. Also discussed is how a site-response and liquefaction hazard analysis using GML-encoded geotechnical data can enhance the results of such analyses.

5.2 Site-Response Analysis

Earthquake waves travel from the fault rupture source to the site via the bedrock. The distance between the rupture source and the site could be several miles. If the site has a soil layer over the bedrock, the waves propagate from the bedrock to the ground surface through this soil layer. Although the thickness of soil layer is generally much less than the earthquake source distance from the site, the wave propagation through the soil has a significant impact on ground surface motions.

The influence of local soil conditions on ground surface motion has been recognized since the 1920s. It was observed that during earthquakes structures located on thick soil layers suffer more damage in comparison to similar structures built on rock. Later, with advent of seismographs seismologist were able to directly measure ground accelerations during earthquakes. Comparing soil site records with rock site records showed that soil sites do amplify ground motions in certain spectral frequencies.

Seismologists and geotechnical engineers have developed analytical methods to quantify this effect. This is achieved using an analysis called ground or site-response analysis. Site-response analysis is a procedure in which seismic excitation at a given location is propagated through soil strata to obtain ground notion at other locations.

Figure 5-2 shows a schematic diagram of the effect of site response on surface ground motions. Figure 5-3 shows a comparison between firm-ground and ground surface response spectra for several sites in the port of Long Beach. The firm-ground spectrum is the design spectrum developed using a probabilistic seismic hazard analysis (PSHA). This spectrum was then used to develop spectrum-compatible acceleration-time histories for the site. The spectrum-compatible strong motion was used in a one-dimensional site-response analysis for several sites (Zones A to D) and surface ground motions for each site was calculated. The site-effect adjusted spectrum is the envelope of surface ground motion acceleration response spectra from these four zones. The amplification of the acceleration response spectrum in longer spectral periods is evident from this figure.

Site-response analysis can be performed in 1-, 2- or 3-dimensional directions, however for most practical applications 1-dimensional analysis is performed. Both equivalent linear and non-linear analyses are used. In equivalent linear analysis, soil elastic properties are estimated in the beginning of each iteration, based on the strain levels calculated at the end of the previous iteration.

Site-response analysis is commonly performed not only to determine soil site ground motions, but also to determine stresses and strains in the soil layers for liquefaction evaluation and for soil-structure interaction analysis, as well as to determine earthquake-induced forces on slopes, embankments and retained earth structures.

5.2.1 Site-Response Analysis Using GML-Encoded Geotechnical Data

A one-dimensional site-response analysis is the most common type of site response analysis in engineering practice. This analysis is performed using an equivalent linear method using programs such as SHAKE (Schnabel et al., 1972) or EERA (Bardet et al., 2000), or non-linear method using programs like DESRA (Finn et al., 1977). The one-dimensional equivalent linear approach, which is more common for practical purposes, is discussed here.

Figure 5-4 shows a typical idealized soil profile for one-dimensional equivalent linear ground response analysis. In order to develop the model for equivalent site response analysis, the following parameters need to be determined for site soils:

- Soil stratigraphy
- Shear modulus and Poisson's ratio
- Density
- Shear strength
- Stress-strain relationship, e.g., shear modulus reduction and damping ratio curves

All these parameters can be evaluated using the data in a GML file. Although this can be achieved easily in an interactive application, the interactive approach is not suitable for large-scale studies that are the target users of this architecture. The goal of this framework is to perform the operation without user intervention for a large volume of data, e.g., site-response analysis for hundreds or thousands of sites. Therefore, all model parameters need to be determined automatically within the software. This process can be performed using SPT or CPT data, as discussed below.

5.2.1.1 Soil Stratigraphy

Soil stratigraphy can be established using visual soil classification from boring logs, SPT blowcounts, and soil classification based on laboratory test results. If SPT results are used for a site-response analysis model, each interval with SPT can be treated as a discrete layer, with soil properties estimated based on the blowcount and visual or laboratory soil classification from that test. The typical distance between SPTs is 5 to 10 feet, which results in layers 5- to 10-feet thick, appropriate for most site-response analyses.

CPT borings can also be used to determine soil stratigraphy. One of the main applications of CPT test is to determine soil profile and soil behavior type. The soil type can be established from correlations between side friction (f_s) and tip resistance (q_t) using a soil behavior chart. The charts were developed for non-normalized values (Robertson et al., 1986), or for normalized values (Robertson, 1990).

Figure 5-5 shows the non-normalized CPT-based soil behavior type classification system proposed by Robertson et al. (1986). The non-normalized charts can be used for depths up to about 60 feet. The chart identifies general trends in ground response, such as increasing relative density (D_r) for sandy soils, increasing stress history (OCR), soil sensitivity (S_t) and void ratio (e) for cohesive soils.

Figure 5-6 shows the normalized CPT soil behavior type classification (SBT_N) chart (Robertson, 1990). This chart is normalized for increase in f_s and q_t due to increase in effective stress at greater depths.

A major benefit of CPT-based soil profiling is that the boundary between different layers can be exactly determined due to continuous f_s and q_t data. If SPT is used for similar application, some kind of assumption needs to be made regarding the boundary between soil types.

If CPT is used to develop a site-response analysis model, the test results needs to be averaged over the soil layer thickness. For example, if a 5-feet layer thickness is adopted, CPT results should be averaged within that 5-feet interval to evaluate soil properties for analysis.

5.2.1.2 Soil Shear Wave Velocity

Shear wave velocity (V_s) is a key parameter in site-response analyses. For important projects it is often directly measured in the field using geophysical type tests (e.g., cross hole, down hole or suspension velocity methods). However, for less important projects or preliminary studies it may be estimated from SPT or CPT measurements. Several correlations, both SPT-based and CPT-based, are presented in this section.

SPT Correlation with Shear Wave Velocity and Modulus

There are several correlations between soil shear wave velocity (Vs) and SPT blowcount (SPT-N). Figure 5-7 shows two correlations: (a) for sandy soils and (b) for clayey soils, which are used in Japanese Highway Design Code (Towhata, 2008).

Other sources use different correlations. Caltrans (2008) recommends the following equations for correlating SPT-N to V_s:

For cohesionless soil (Sykora, 1987): $V_s = 100.5(N_{60})^{0.29}$

For cohesive soil (Ohta and Goto, 1978): $V_s = 86.9(N_{60})^{0.333}$

V_s is also related to soil shear modulus (G) and density (ρ) according to the following equation:

$$V_s = \sqrt{\frac{G}{\rho}}$$

CPT Correlations with Shear Wave Velocity and Modulus

CPT test results have been correlated to shear wave velocity and elastic or shear modules. Figure 5-8 shows a chart that correlates normalized shear wave velocity (V_{s1}) of soil to normalized CPT tip resistance and side friction (Robertson and Cabal, 2009). Using this chart, the shear wave velocity can be estimated from the following equation:

$V_s = [\alpha_{vs}(q_t - \sigma_v)/p_a]^{0.5}$ (m/s); where $\alpha_{vs} = 10^{(0.55I_c+1.68)}$

Caltrans (2008) recommends the following equations for correlating CPT results to V_s:

For cohesionless soil (Mayne, 2007): $V_s = 277(q_t)^{0.13}(\sigma'_{v0})^{0.27}$

For cohesive soil (Mayne and Rix, 1995): $V_s = 1.75(\sigma'_{v0})^{0.627}$

In above equations q_t is CPT tip resistance in kPa.

5.2.1.3 Poisson's Ratio

Poisson's ratio is generally correlated with soil type, as shown in Table 5-1. The soil classification from boring logs or soil behavior type from CPT provides the basis for estimating Poisson's ratio. This procedure is the norm for most practical cases and Poisson's ratio is rarely measured in practice.

5.2.1.4 Soil Density

There are no direct correlations between SPT and CPT and soil density, but both are correlated with relative density. For drilled boreholes with SPT samples soil density can be evaluated from direct laboratory measurements if available. Otherwise it can be correlated with SPT blowcounts and relative density, as shown in Table 5-2.

CPT results have also been correlated to soil density. Kulhawy and Maine (1990) recommend the following equation to estimate relative density from CPT:

$$D_r^2 = \frac{q_{c1}}{305 Q_c Q_{OCR} Q_A}$$

where:

q_{c1}: Normalized CPT tip resistance, $q_{c1} = (q_c/p_a)/(\sigma'_{v0}/p_a)^{0.5}$

p_a: Reference pressure of 1 tsf (100 kPa)

Q_c: Compressibility factor, ranges for 0.91 (low compressibility) to 1.09 (high compressibility)

Q_{OCR}: Overconsolidation factor, $Q_{OCR} = OCR^{0.18}$

Q_A: Aging factor, $Q_A = 1.2 + 0.05\log(t/100)$

5.2.1.5 Soil Shear Strength

Soil shear strength is not used in an equivalent linear analysis, but it is used in a non-linear analysis. For SPT-based analysis, the shear strength for sandy and clayey soils can be determined from Table 5-2 and Table 5-3, respectively.

For CPT-based analysis, the friction angle for sandy soils can be determined from the correlation in Figure 5-9, proposed by Robertson et al. (1983).

The shear strength of clayey soils can be determined from the tip resistance (q_t) and vertical stress (σ_v) using the following equation:

$$S_u = \frac{q_t - \sigma_v}{N_{kt}}$$

N_{kt} typically varies from 10 to 20, with 14 being the average for most clayey soils.

5.2.1.6 Shear Modulus Reduction and Damping Curves

The shear modulus degradation and damping curves are used to simulate the nonlinear stress-strain behavior of soil in an equivalent linear analysis, which is usually performed in a frequency domain. These curves are usually developed based on laboratory tests. For typical site-response analyses generic shear modulus reduction and damping curves are used. Site-specific curves are not common for the majority of projects involving site-response analysis.

Some of the common shear modulus reduction-damping relationships are listed as follows:

- Seed and Idriss (1970) for gravel and sand: these curves are functions of relative density (Dr) or void ratio (e_o)

- Vucetic and Dobry (1991) for cohesive soil: function of plasticity index (PI) and overconsolidation ratio (OCR)

The newer relationships listed below have confining pressure dependence:

- EPRI (1993), Hashash and Park (2001), Darendeli (2001)

Figure 5-10 shows a set of shear modulus reduction and damping curves for sandy soils (Darendeli, 2001). Figure 5-11 shows the curves for clayey soils developed by Vucetic and Dobry (1991).

5.2.2 Site-Response Analysis Application

The following example demonstrates the application of the geotechnical database and XML data format for SPT- and CPT-based site-response analysis. The example also shows the usage of XML features in Microsoft Excel 2003.

For this demonstration program, the analysis is conducted using a client application developed in Microsoft Excel 2003, based on the existing site-response analysis application EERA (Bardet et al., 2000). This implementation methodology is chosen to (1) show the potential of XML-based data formats for integration with existing commercial software, and (2) to eliminate the effort and time associated with developing a custom user interface from scratch.

Microsoft Excel provides a number of useful features for working with XML (and inherently GML) documents. Some of these features include:

- XML spreadsheet file format: XML Spreadsheet Schema (XML-SS) would be used to preserve the appearance and structure of the spreadsheet. XML-SS could also be used as a schema for XML data from other sources. XSL-style sheets could be used to change the format of XML data in the spreadsheet.
- Support for XML path language (XPath): XPath is an expression language used by XML-related languages to access or refer to different parts of an XML document. XPath expressions are now supported in the Excel 2003 object model.
- XML maps: XML maps are abstract objects that relate an XML schema (XSD) to a Workbook. In an XML map, different elements in XSD schema are mapped to certain locations in the spreadsheet. When an XML document conforming to the XSD file is loaded to the map, Excel places the mapped elements from the XML file into those specified cells in the spreadsheet. XML maps can be created interactively by the user, or can be defined programmatically using Visual Basic macros.
- Lists: A list is a continuous vertical sequence of cells in an XML map. Each column in a list is mapped through XPath expressions to a set of repeated elements in the XML schema. Excel automatically generates a list when a repeating element is added to the XML map. Upon insertion of an XML document to the XML map, the repeated elements of the XML document will be placed in the designated column of the list in successive order.

Excel 2003 adds a user interface to interact with XML documents. This interface is called XML Source Pane and can be used to show XML schemas and define new XML maps interactively (through drag-and-drop technique).

In this example the borehole data was encoded in XML format and stored in a remote database server. The data were served using the Web service GIME that was discussed in Section 5.1.

Figure 5-12 shows a screen shot of the site-response analysis application EERA. The XML Source Pane is used for defining a new XML map in Excel. An XML map links the elements in an XML schema to certain cells in an Excel spreadsheet. When the XML document is inserted into the spreadsheet, the mapped cell values are updated with corresponding values from the XML document. Figure 5-13 schematically demonstrates using XML documents and schemas in Excel spreadsheets. In Figure 5-12 the mapped cells are shown in bold font and blue border on the task pane and spreadsheet, respectively.

The Visual Basic code used to connect and query Web services is generated automatically using the Web Service References Tool. The Web Service References Tool is used in the Microsoft Office Visual Basic Editor to create a Visual Basic for Applications (VBA) proxy-class module from a selected Web Service Description Language (.wsdl) or Visual Studio .NET discovery (.vsdisco) file. The Web Service References Tool uses the information provided by the WSDL to the visual basic code for connecting to and executing Web services.

Similar procedures can be used in applications like REDARS or HAZUS-MH to perform SPT- and CPT-based liquefaction evaluation for large areas. This can be conducted by a module that interacts with Web services and retrieves geotechnical data with minimal user interaction. Such analyses will be time and cost prohibitive if user interaction is required.

5.3 Liquefaction Evaluation Using GML Data

Liquefaction is a phenomenon whereby saturated granular soils lose their inherent shear strength due to increased pore water pressures, which may be induced by cyclic loading such as that caused by an earthquake. Low-density granular soils, shallow groundwater and long-duration/high-acceleration seismic shaking are some of the factors favorable to cause liquefaction. Liquefaction is generally considered possible when the depth to groundwater is less than about 50 feet below the ground surface.

There are several methods for evaluating liquefaction potential due to earthquakes. The analytical method is based on evaluating cyclic stress ratio (CSR) due to earthquakes and comparing it with cyclic resistance ratio (CRR) of the soil. The CSR is evaluated based on earthquake PGA and depth. The cyclic shear resistance can be evaluated from several tests, e.g., SPT, CPT, V_s, and Becker penetration test (BPT). Hereafter, several liquefaction analysis procedures are discussed, and the soil parameters used in each model are reviewed.

5.3.1 SPT-Based Liquefaction Evaluation

The most common procedure is the SPT method, in which the CRR is evaluated using corrected SPT blowcounts. The procedure is explained in detail in Youd et al. (2001). A summary of procedures highlighting the use of geotechnical data is presented here.

The first step in evaluation of liquefaction is evaluating the CSR, which is calculated from the following equation:

$$CSR = (\tau_{av}/\sigma'_{v0}) = 0.65(a_{max}/g)(\sigma_{v0}/\sigma'_{v0})r_d$$

where a_{max} is peak earthquake horizontal acceleration at the ground surface; g is acceleration of gravity; σ_{vo} and σ'_{vo} are total and effective vertical overburden stresses; and r_d is a stress reduction coefficient for depth (z) that accounts for flexibility of the soil profile. It can be evaluated from the following equation or other similar equations:

$$r_d = \begin{cases} 1.0 - 0.00765z & for\ z \leq 9.15\ m \\ 1.174 - 0.0267z & for\ 9.15\ m < z \leq 23\ m \end{cases}$$

The CRR for clean granular sands is evaluated from the following equation from corrected blowcount $(N_1)_{60}$:

$$CRR_{7.5} = \frac{1}{34 - (N_1)_{60}} + \frac{(N_1)_{60}}{135} + \frac{50}{[10.(N_1)_{60} + 45]^2} - \frac{1}{200}$$

The above equation is valid for $(N_1)_{60}$ <30. If $(N_1)_{60} \geq 30$ the soil is non-liquefiable.

The blowcount needs to be corrected for hammer energy efficiency ratio, overburden pressure, fines content, rod length, borehole diameter and sampling method, as explained in Youd et al. (2001).

The factor of safety against liquefaction is evaluated from the following equation:

$$FS = (CRR_{7.5}/CSR)MSF$$

where MSF is earthquake magnitude scaling factor for earthquakes with magnitude other than 7.5. MSF is a function of earthquake magnitude (independent of soil properties) and can be estimated using several methods, as explained in Youd et al. (2001).

5.3.2 CPT-Based Liquefaction Evaluation

In recent years CPT-based methods have become more popular. This is due to both widespread use of CPT in geotechnical investigations, as well as availability of commercial software for this analysis. In these methods the cyclic shear resistance is calculated using CPT measurements. The following equations are used to evaluate CRR from clean-sand normalized cone penetration resistance $(q_{c1N})_{cs}$:

$$CRR_{7.5} = \begin{cases} 0.833[(q_{c1N})_{cs}/1{,}000] + 0.05 & for\ \ (q_{c1N})_{cs} < 50 \\ 93[(q_{c1N})_{cs}/1{,}000]^3 + 0.08 & for\ 50 \leq (q_{c1N})_{cs} < 160 \end{cases}$$

Cone penetration resistance is normalized to 100kPa overburden pressure for use in the above equation. Details of this normalization can be found in Youd et al. (2001). The rest of the liquefaction evaluation procedure is similar to the SPT-based method. Figure 5-14 shows the CPT-based liquefaction evaluation procedure for sandy soils.

5.3.3 Other Empirical Models

The methods discussed above require information on soil profile and field test results for the soil strata. In a large number of practical problems these data are not available for the site. For these sites there are empirical methods that provide an estimation of liquefaction potential and liquefaction-induced lateral displacement. Bardet et al. (2002) proposed FFGS4, a model for prediction of liquefaction-induced lateral displacement. This method is currently used in REDARS to evaluate liquefaction potential.

The FFGS4 model uses the following equation to evaluate the liquefaction-induced lateral displacement (D):

$$\log(D+0.01) = b_0 + b_{off} + b_1 M + b_2 \log(R) + b_3 R + b_4 \log(W) + b_5 \log(S) + b_6 \log(T_{15})$$

where D is in meters; M the moment magnitude; R the epicentral distance (km); S the slope (%) of the ground surface; W the free-face ratio (%); T_{15} the thickness of saturated cohesionless soils with $N1_{60}$ < 15 in meter (excluding depth > 20 meter and >15% clay content). The coefficients b_0 to b_6 are given in Table 5-4.

5.3.4 Liquefaction Analysis Application

The following example demonstrates the application of the geotechnical database and XML data format for SPT- and CPT-based liquefaction evaluation. This program is also developed in Microsoft Excel 2003 and uses GIME Web service to retrieve the geotechnical data.

Figure 5-15 shows a screen shot of the developed application. The XML Source Pane was used for defining a new XML map in Excel based on the GML schema. The mapped cells are shown in bold font and blue border in task pane and spreadsheet, respectively.

The same Visual Basic subroutines that used to connect and query the Web service for the site-response analysis are used here.

Similar procedure can be used in applications like REDARS to perform SPT- and CPT-based liquefaction evaluation for large areas. This type of analysis can be performed by a client software module that retrieves the geotechnical data from the database through the Web service and feeds it to the REDARS analysis module. This scenario is discussed further in the case study presented in the next section.

Using a standardized GML data model and Web services, this operation can be performed with minimal user interaction. While in lieu of a standard data format and Web services, this operation will be time and cost prohibitive if user interaction is required.

5.4 Case Study: REDARS Analysis

In this section a case study is presented to demonstrate how the proposed geotechnical data management system and the developed data format may improve current studies in risk analysis and infrastructure planning. The case presented here is a REDARS analysis for the Port of Los Angeles (POLA) and the Port of Long Beach (POLB) road transportation network, which transits a major part of the cargo from POLA and POLB to the continental United States.

The original study report is presented in Appendix. That report includes an introduction to REDARS analysis methodology, details of seismic hazard analysis and a brief summary of the analysis results. In the remainder of this section potential improvements in the project results using a geotechnical database are discussed.

5.4.1 Case Study Background

The study network is shown in Figure 5-16. The entire network included more than 6,000 bridges. Three rupture scenarios were considered in this study. These scenarios included the Newport-Inglewood Fault, the Palos Verdes Fault,and the South San Andreas Fault. The first two are the dominant faults in the POLA and POLB area, while the last one is the fault capable of generating the highest magnitude earthquake in the Southern California.

Figure 5-17 shows a typical PGA contour map due to rupture on the Newport-Inglewood Fault. These values are estimated for a Soil Profile Type B/C (Vs = 760 m/sec), which represents a firm-ground condition.

Figure 5-18 shows the spectral accelerations at 1.0 second for all bridges in the network, due to the same rupture scenario on the Newport-Inglewood Fault. These values include the local site effects, based on NEHRP generalized soil profile types (Table 5-5). The spectral acceleration value at 1.0 second was used to evaluate the damage state and economic loss due to damage to each bridge in the analysis.

Figure 5-19 shows the estimated damage state for each bridge in the network, due to ground motions induced by a rupture scenario on the San Andreas Fault. A damage state of 1 indicates no damage, while a damage state of 5 indicates complete failure of the structure. This index was used to evaluate the economic loss due to damage to the bridges resulting from considered earthquake scenarios.

More details about this REDARS analysis study are included in Appendix.

5.4.2 Improving REDARS Analysis by Including Local Site Effects

REDARS is a modular program, a complete deterministic or probabilistic seismic and liquefaction hazard analysis involving 4 modules: 1) system module, 2) hazards module, 3) component module, and 4) economic module. The project discussed here is mainly focused on the hazards module, where the seismic hazard due to ground shaking, liquefaction and fault rupture are evaluated for the subject system.

One of the input parameters in the hazards module is local soil conditions. The local soil condition data are used in ground -shaking estimation and liquefaction assessment. Figure 5-20 shows a flowchart summarizing current Caltrans seismic hazard analysis methodology for highway bridges. In most REDARS analyses the soil data is obtained from generalized NEHRP soil classes. In this method soils are classified in five categories, as shown in Table 2-1. This classification provides a generalized estimation of soil conditions at the site. However, as shown on Figure 5-20, this method cannot be used for Soil Profile Type F, and in final design for Soil Profile Type E. Even for other site soil conditions, both seismic-hazard and liquefaction evaluations may be improved by utilizing site-specific geotechnical information using borehole data, rather than the generalized NEHRP soil classification.

The effect of soil conditions on the spectral acceleration in 1.0 second, which was used to estimate the damage to each bridge, can be significant. The importance of this parameter can be seen in Figure 5-3, which compares the ground motion response spectrum for firm-ground conditions, and several local site conditions in the POLA.

The results of REDARS study in Appendix can be improved if a site-specific site-response analysis could be performed for each bridge site. As discussed in the case study report, the main obstacle in using a site-specific ground response analysis in a REDARS analysis is unavailability of geotechnical data for bridge sites. Theoretically the data are available because Caltrans has multiple boreholes at each bridge site, and the LOTBs are included in the bridge plans as hardcopies and electronic PDF files. An engineer can analyze the data in these LOTBs and perform a site-response analysis for each bridge. However these PDF files are not usable by applications, therefore this process cannot be automated in an application. Therefore, performing this task for the example project with more than 6,000 bridge sites is time and cost prohibitive. On the other hand if the data were available in GML format, the process could be automated, as explained in Section 5.2.

Similarly, a liquefaction hazard analysis could be performed using one of the methods discussed in Section 5.3.

Table 5-1. Typical ranges of elastic properties of soils and rocks.

Soil / Rock Type	Young's Modulus, Es (tsf)	Poisson's Ratio, ν
Soils		
Clay:		
Soft sensitive	25-150	
Firm to stiff	150-500	0.4-0.5
Very stiff	500-1000	(undrained)
Loess	150-600	0.1-0.3
Silt	20-200	0.3-0.35
Fine sand:		
Loose	80-120	
Medium dense	120-200	0.25
Dense	200-300	
Sand		
Loose	100-300	0.2-0.35
Medium dense	300-500	
Dense	500-800	0.3-0.4
Gravel		
Loose	300-800	
Medium dense	800-1000	
Dense	1000-2000	
Rocks		
Sound, intact igneous and metamorphic	6-10x10^5	0.25-0.33
Sound, intact sandstone and limestone	4-8x10^5	0.25-0.33
Sound, intact shale	1-4x10^5	0.25-0.30
Coal	1-2x10^5	

(after Hunt, 2005)

Table 5-2. SPT soil correlations for sandy soils recommended by FHWA (1986).

Type of Soil	Resistance N (blows/ft)	Relative Density D_r	Angle of Internal Friction ϕ (Deg)	
			Peck et al. (1974)	Meyerhof (1956)
Very loose sand	< 4	< 0.2	< 29	< 30
Loose sand	4 – 10	0.2 – 0.4	29 – 30	30 – 35
Medium sand	10 – 30	0.4 – 0.6	30 – 36	35 – 40
Dense sand	– 50	0.6 – 0.8	36 – 41	40 – 45
Very dense sand	> 50	> 0.8	> 41	> 45

Table 5-3. SPT soil correlations for fine-grained soils recommended by FHWA (1986).

Penetration Resistance N (blows/ft)	Undrained Shear Strength c (ksf)	Consistency
< 2	< 0.25	Very soft
2 – 4	0.25 – 0.50	Soft
4 – 8	0.50 – 1.00	Medium
8 – 15	1.00 – 2.00	Stiff
– 30	2.00 – 4.00	Very stiff
> 30	> 4.00	Hard

Table 5-4. FFGS4 liquefaction prediction model parameters. This is an empirical model for liquefaction and lateral displacement evaluation.

Coefficients	Data Set A	Data Set B
b_0	-6.815	-6.747
b_{off}	-0.465	-0.162
b_1	1.017	1.001
b_2	-0.278	-0.289
b_3	-0.026	-0.021
b_4	0.497	0.090
b_5	0.454	0.203
b_6	0.558	0.289
R^2 adjusted	64.25%	64.27%
Number of data	467	213

Set A: complete data for all ranges of displacement amplitude. Set B: data limited to displacement amplitudes smaller than 2 meters.

(Bardet et al., 2002)

Table 5-5. Soil profile types as defined by Applied Technology Council-32-1 (1996).

Soil Profile Type	Soil Profile Description
A	Hard rock with measured shear wave velocity v_{S30} > 5000 ft/s (1,500 m/s)
B	Rock with shear wave velocity 2,500 < v_{S30} < 5000 ft/s (760m/s < v_{S30} < 1,500 m/s)
C	Very dense soil and soft rock with shear wave velocity 1,200 < v_{S30} < 2,500 ft/s (360m/s < v_{S30} < 760 m/s) or with either standard penetration resistance N > 50 or undrained shear strength s_u ≥ 2,000 psf (100 kPa)
D	Stiff soil with shear wave velocity 600 < v_{S30} < 1,200 ft/s (180 m/s < v_{S30} < 360 m/s) or with either standard penetration resistance 15 ≤ N ≤ 50 or undrained shear strength 1,000 < s_u < 2,000 psf (50 < s_u < 100 kPa)
E	A soil profile with shear wave velocity v_{S30} < 600 ft/s (180 m/s) or any profile with more than 10 ft (3 m) of soft clay, defined as soil with plasticity index *PI* > 20, water content *w* ≥ 40 percent, and undrained shear strength s_u < 500 psf (25 kPa)
F	Soil requiring site-specific evaluation: Soils vulnerable to potential failure or collapse under seismic loading, i.e., liquefiable soils, quick and highly sensitive clays, collapsible weakly-cemented soils Peat and/or highly organic clay layers more than 10 ft (3 m) thick Very high-plasticity clay (*PI* > 75) layers more than 25 ft (8 m) thick Soft-to-medium clay layers more than 120 ft (36 m) thick

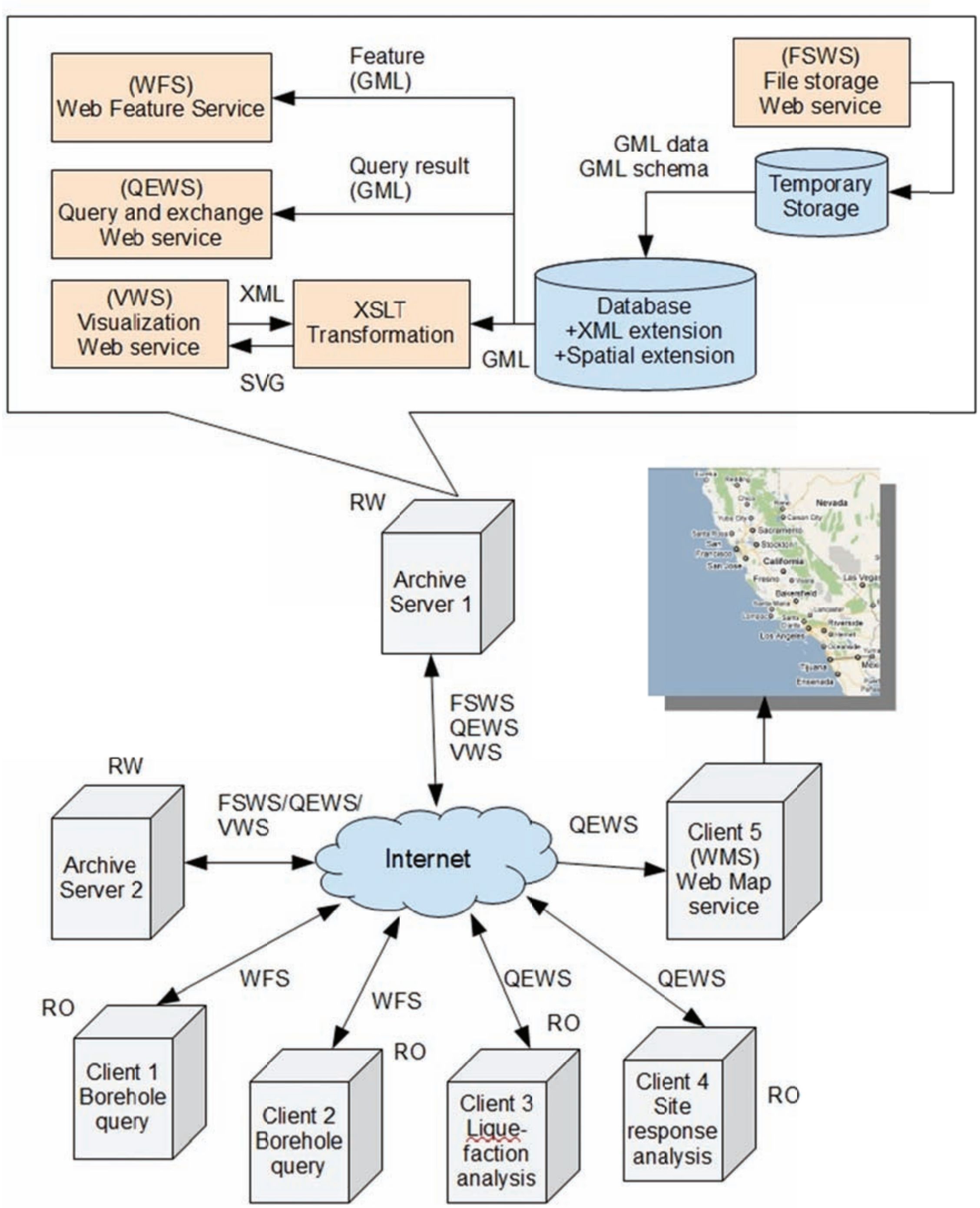


Figure 5-1. Schematic diagram of a Web-based distributed geotechnical database management and dissemination system.

(after Zimmerman et al., 2006)

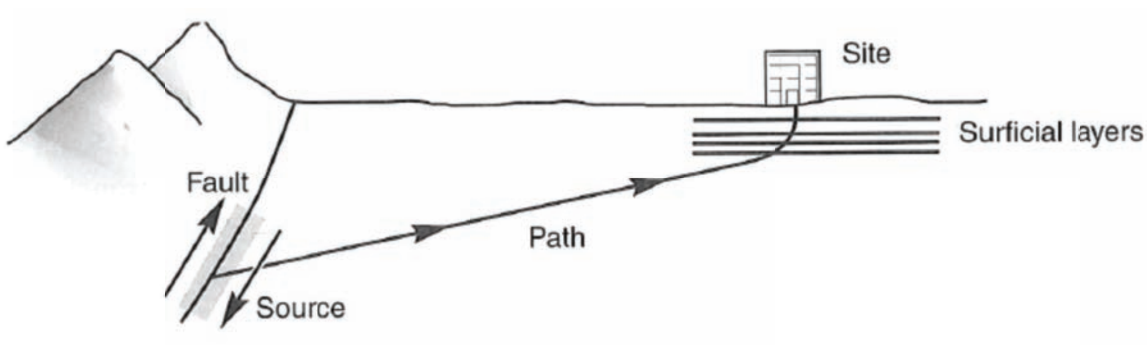


Figure 5-2. Schematic diagram showing the mechanism of the near vertical wave propagation through local site soils during an earthquake. This mechanism is simulated in a site soil response analysis.

(original figure from Kramer, 1996)

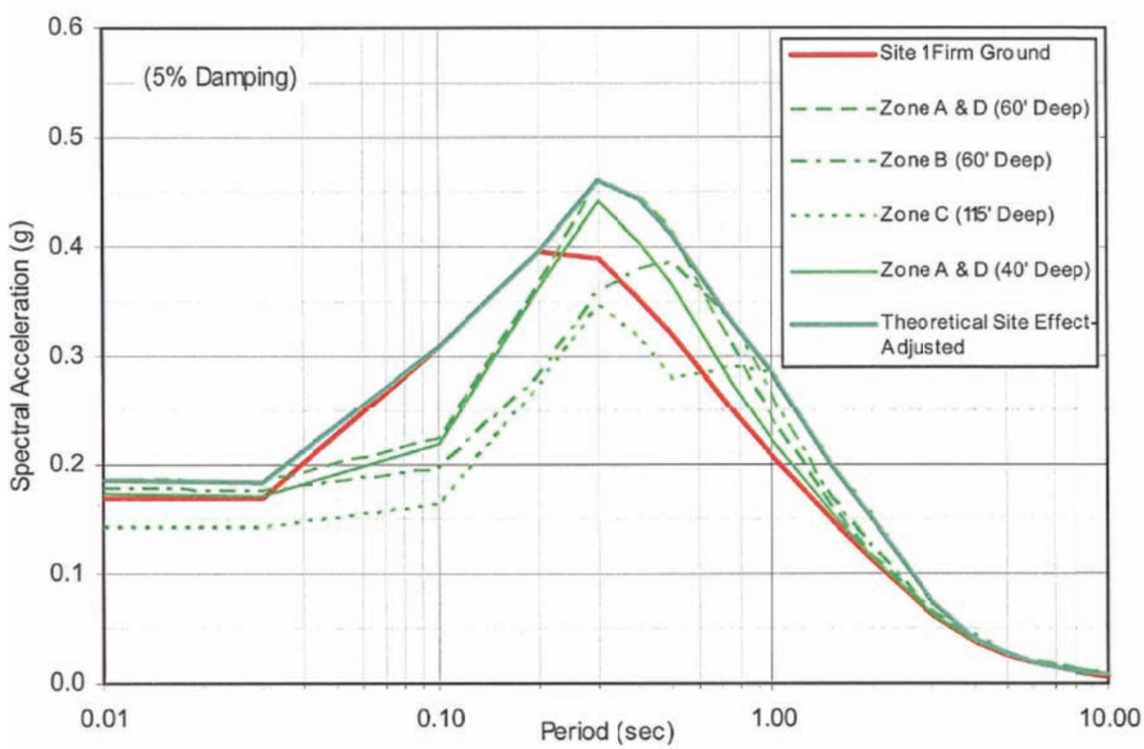


Figure 5-3. Effect of site soil response on acceleration response spectra for several sites in the Port of Los Angeles area.

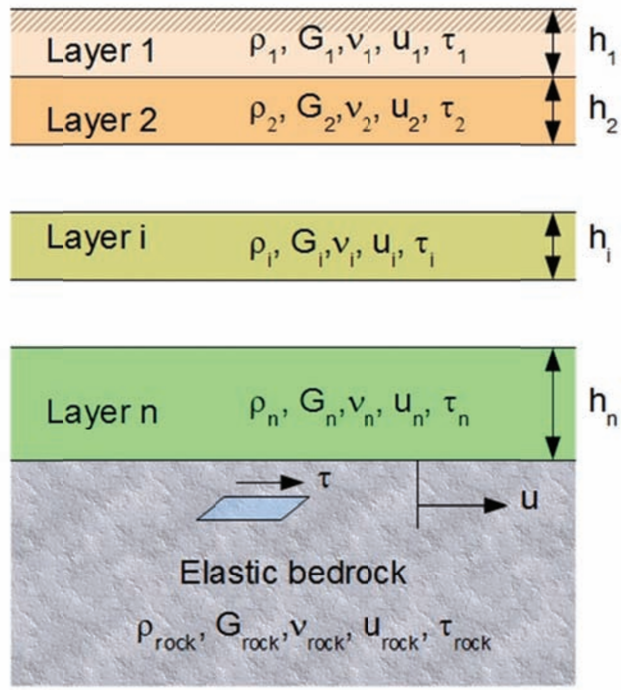
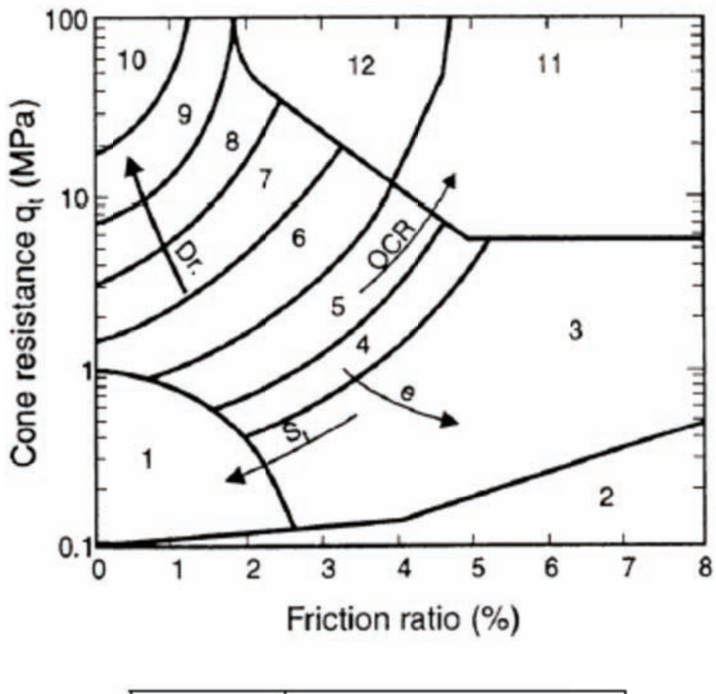


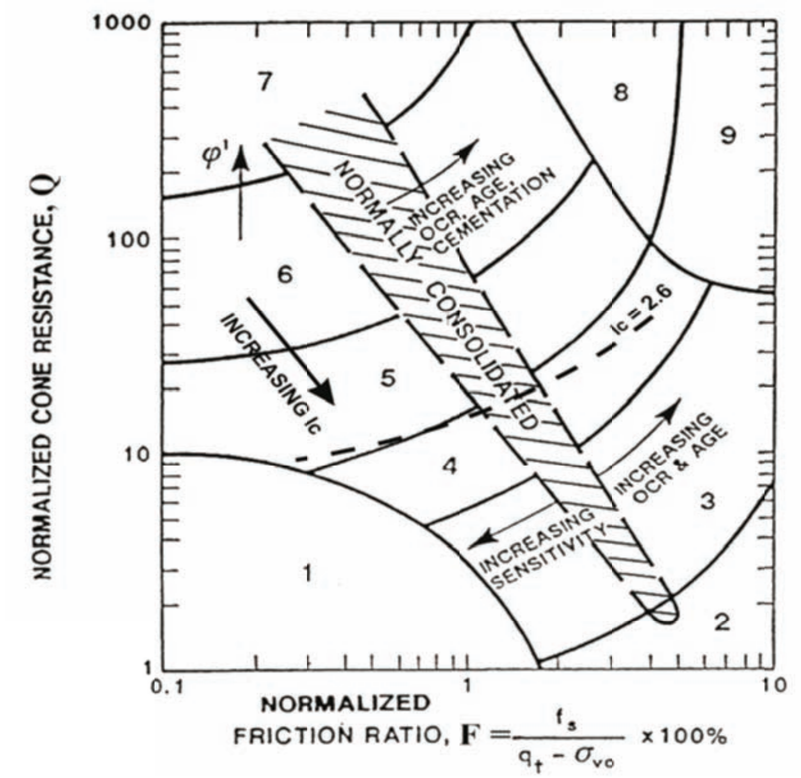
Figure 5-4. Idealized ground model for one-dimensional equivalent linear site response analysis. For each soil layer the following parameters are needed: Thickness, Density, Shear Modulus, Poisson's Ratio. The analysis determines the Shear Stress and Displacement for each layer based on the bedrock motion.



Zone	Soil Behavior Type
1	*Sensitive fine grained*
2	*Organic material*
3	*Clay*
4	*Silty Clay to clay*
5	*Clayey silt to silty clay*
6	*Sandy silt to clayey silt*
7	*Silty sand to sandy silt*
8	*Sand to silty sand*
9	*Sand*
10	*Gravelly sand to sand*
11	*Very stiff fine grained**
12	*Sand to clayey sand**

** Overconsolidated or cemented*

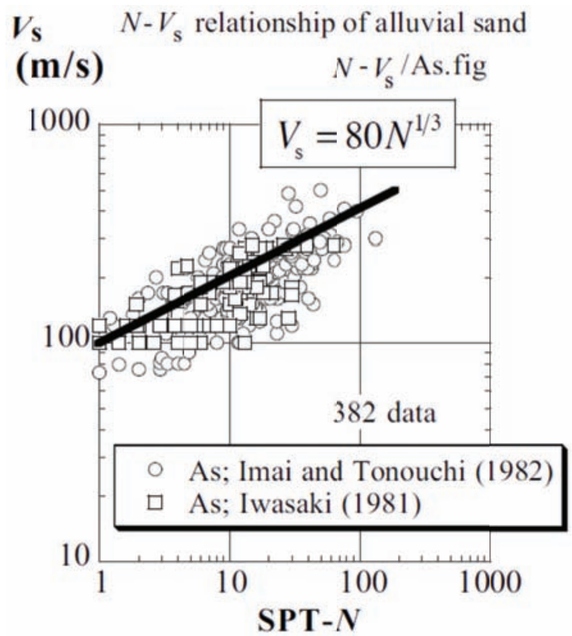
Figure 5-5. Non-normalized CPT Soil Behavior Type (SBT) chart by Robertson et al. (1986). This chart can be used for depths up to about 60 feet.



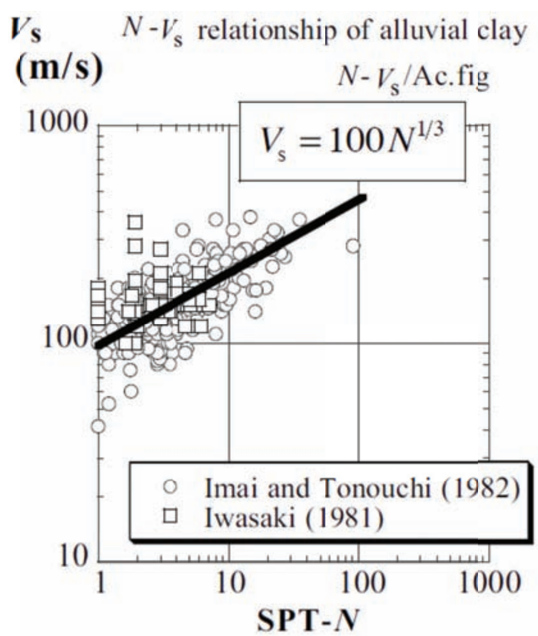
Zone	Soil Behavior Type	I_c
1	Sensitive, fine grained	N/A
2	Organic soils, peats	> 3.6
3	Clays – silty clay to clay	2.95 – 3.6
4	Silt mixtures – clayey silt to silty clay	2.60 – 2.95
5	Sand mixtures – silty sand to sandy silt	2.05 – 2.6
6	Sands – clean sand to silty sand	1.31 – 2.05
7	Gravelly sand to dense sand	< 1.31
8	Very stiff sand to clayey sand*	N/A
9	Very stiff, fine grained*	N/A

* Heavily overconsolidated or cemented

Figure 5-6. Normalized CPT Soil Behavior Type (SBTN) chart, Qt – F (Robertson, 1990).



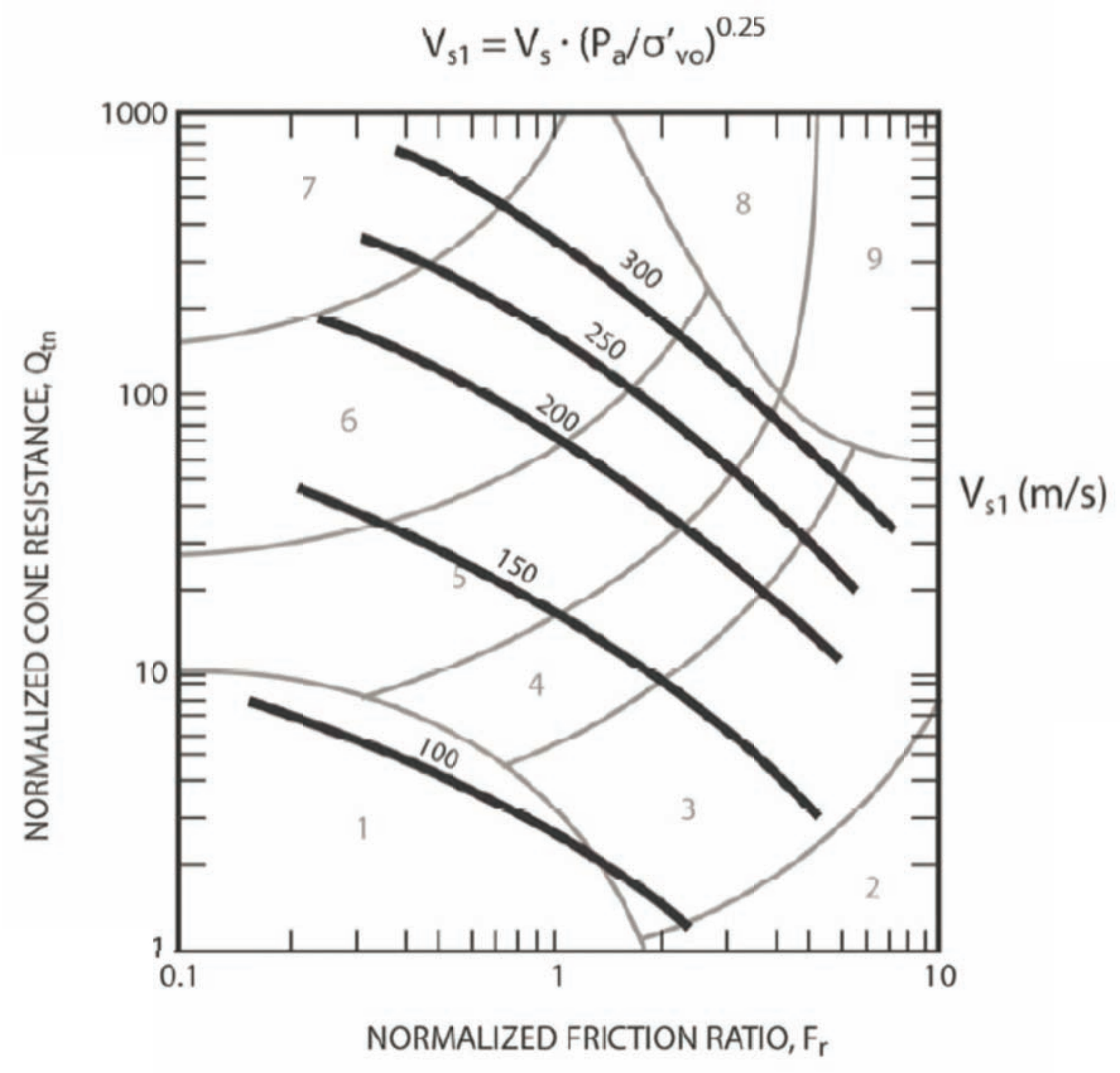
(a)



(b)

Figure 5-7. Correlation between SPT-N value and soil shear wave velocity for (a) sand and (b) clay.

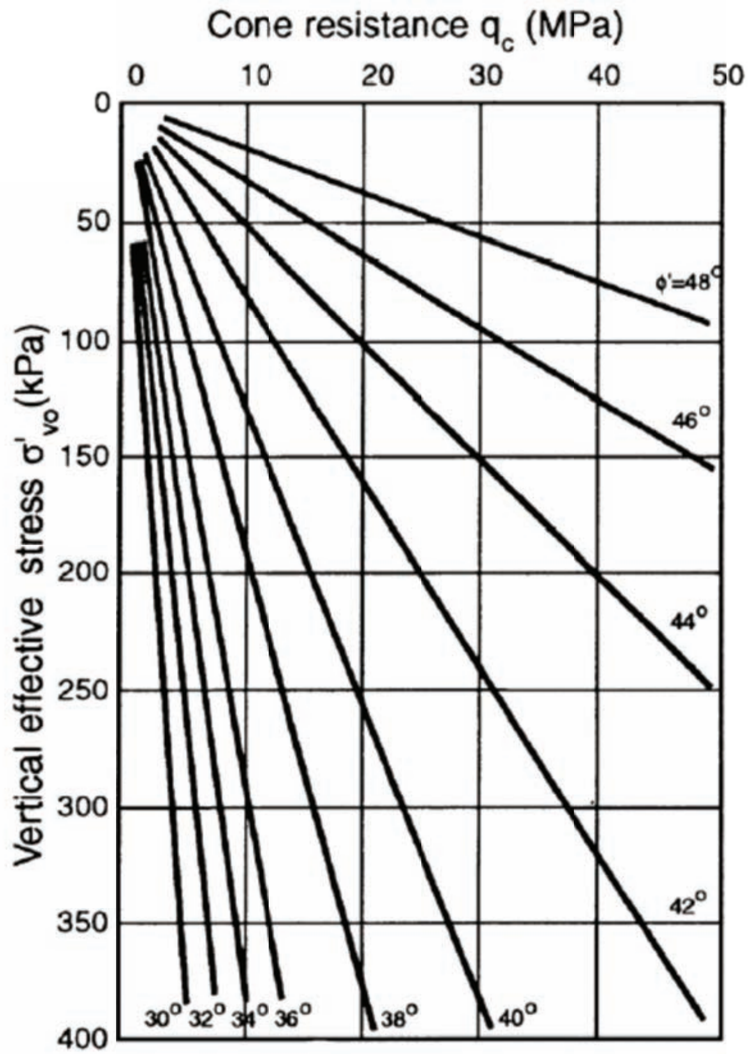
(figure from Towhata, 2008)



$V_s = [\alpha_{vs}(q_t - \sigma_v)/p_a]^{0.5}$ (m/s); where: $\alpha_{vs} = 10^{(0.55 I_c + 1.68)}$

Figure 5-8. Correlation between normalized CPT tip resistance and normalized side friction and normalized soil shear wave velocity.

(after Robertson and Cabal, 2009)



Note: 0.1MPa = 100 kPa = 1 bar ≈ 1 tsf ≈ 1 kg/cm²

$$\tan \phi' = \frac{1}{2.68}\left[\log\left(\frac{q_c}{\sigma'_{vo}}\right) + 0.29\right]$$

Figure 5-9. Correlation between CPT tip resistance and friction angle for sandy soils.

(from Robertson et al., 1983)

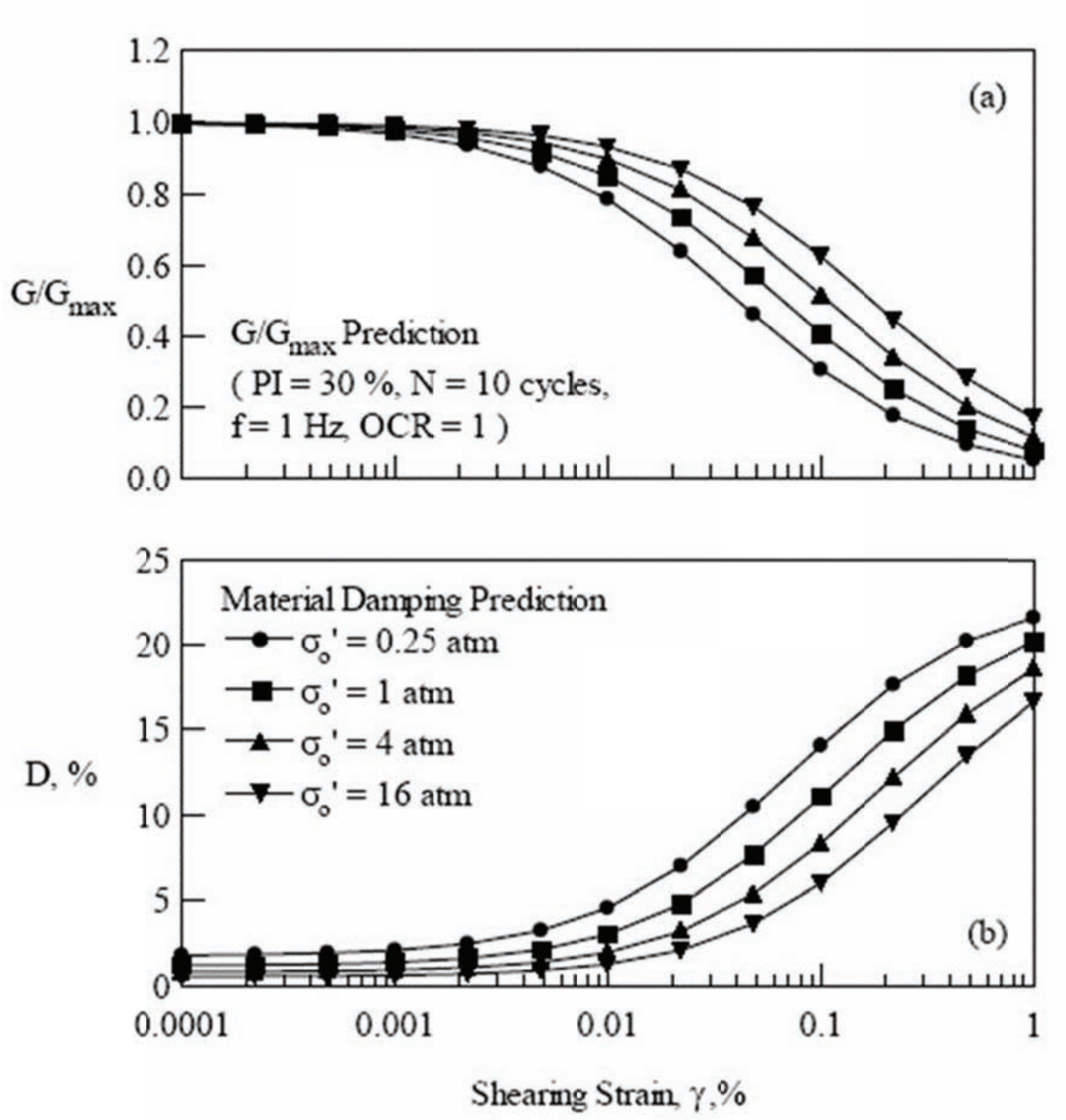


Figure 5-10. Normalized shear modulus reduction and damping curves for sandy soils. These curves are typically used in equivalent site response analyses to model non-linear stress-strain response on soil.

(original firgure from Darendeli, 2001)

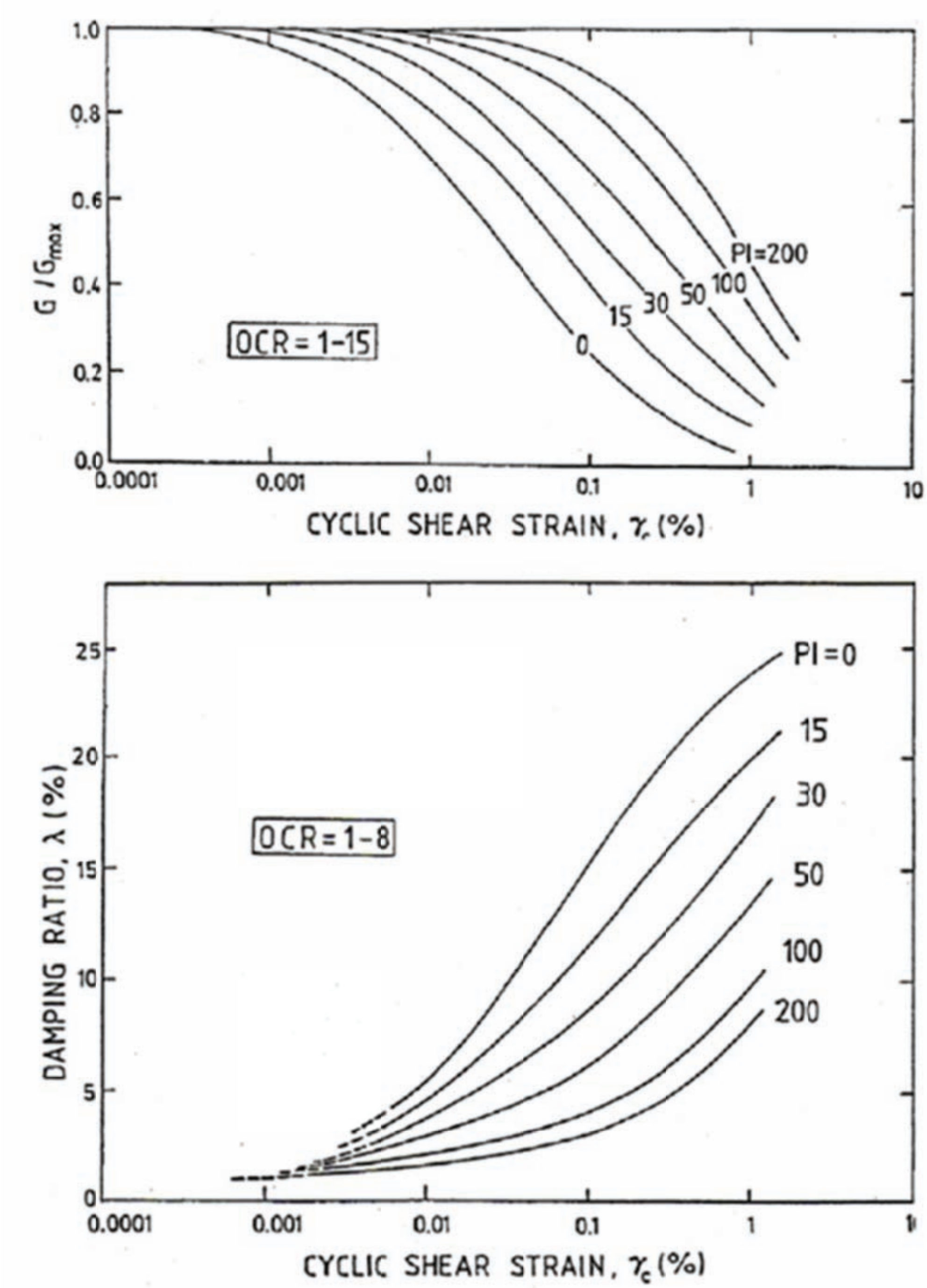


Figure 5-11. Normalized shear modulus reduction and damping curves for clayey soils recommended by Vucetic and Dobry (1991). These curves are typically used in equivalent site response analyses to model non-linear stress-strain response on soil.

Example -- 150-ft layer; input:Diam @ .1g

Fundamental period (s) = 0.48

Average shear wave velocity (m/sec) = 382.02

Total number of sublayers = 17

	Layer Number	Soil Material Type	Number of sublayers in layer	Thickness of layer (m)	Maximum shear modulus G_{max} (MPa)	Initial critical damping ratio (%)	Total unit weight (kN/m^3)	Shear wave velocity (m/sec)	Location and type of earthquake input motion	Location of water table	Depth at middle of layer (m)	Vertical effective stress (kPa)
Surface	1	2		1.5	186.19		19.66	*304.8*			0.8	14.98
	2	2		1.5	150.81		19.66	*274.32*			2.3	44.94
	3	2		3.0	150.81		19.66	*274.32*			4.6	89.89
	4	2		3.0	168.03		19.66	*289.56*			7.6	149.81
	5	1		3.0	186.19		19.66	*304.8*			10.7	209.74
	6	1		3.0	186.19		19.66	*304.8*			13.7	269.66
	7	1		3.0	225.29		19.66	*335.28*			16.8	329.59
	8	1		3.0	225.29		19.66	*335.28*			19.8	389.51
	9	2		3.0	327.24		20.45	*396.24*			22.9	450.63
	10	2		3.0	327.24		20.45	*396.24*			25.9	512.95
	11	2		3.0	379.52		20.45	*426.72*			29.0	575.28
	12	2		3.0	379.52		20.45	*426.72*			32.0	637.60
	13	2		3.0	435.68		20.45	*457.2*			35.1	699.92
	14	2		3.0	435.68		20.45	*457.2*			38.1	762.24
	15	2		3.0	495.71		20.45	*487.68*			41.1	824.56
	16	2		3.0	627.38		20.45	*548.64*			44.2	886.88
Bedrock	17	0			3336.48	1	22.02	*1219.2*	Outcrop		45.7	918.04

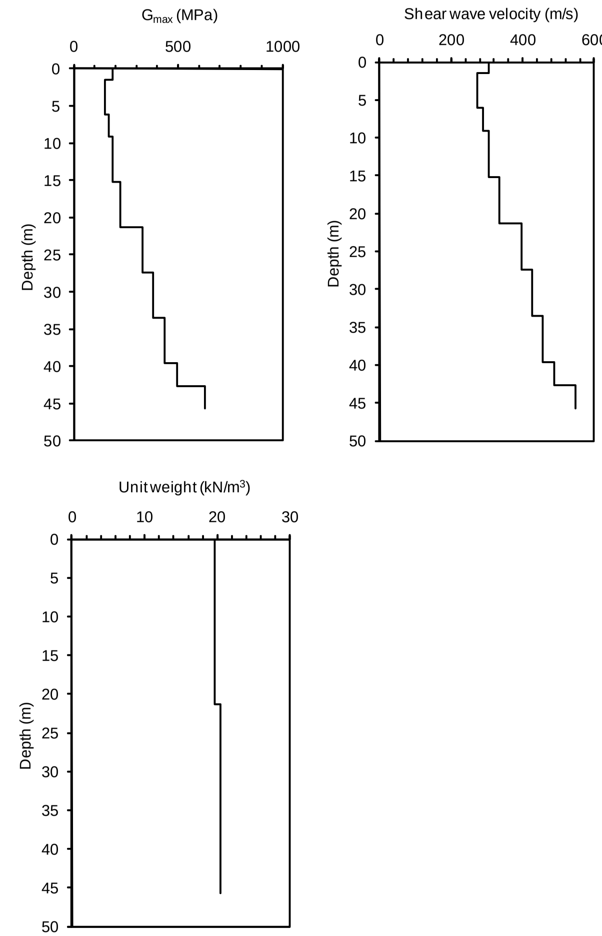


Figure 5-12. Site Response Analysis with EERA using Geotechnical Data from GIME Web Service.

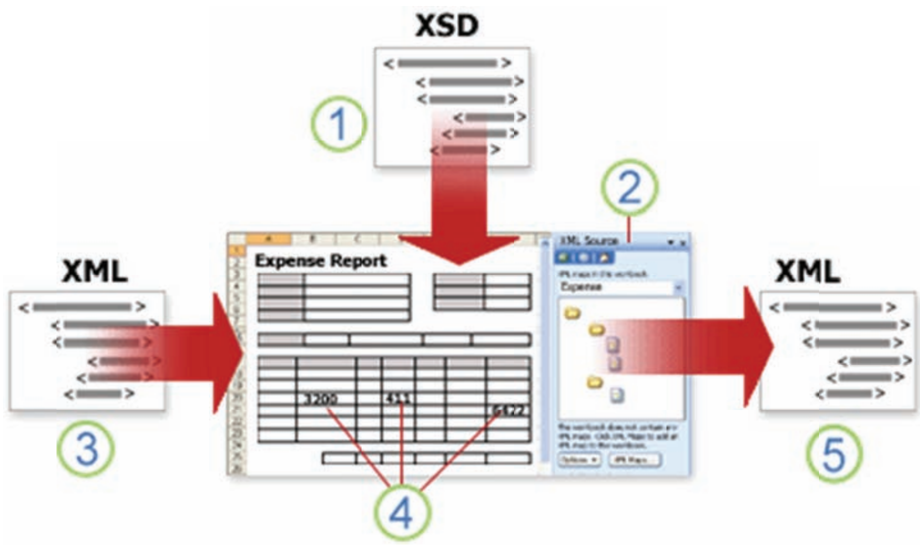


Figure 5-13. XML documents and schemas (XSD) in Microsoft Excel. An XML schema can be associated with a workbook using XML Source Pane. The data in the XML document are linked to corresponding cells in the spreadsheet.

(from Microsoft, 2007)

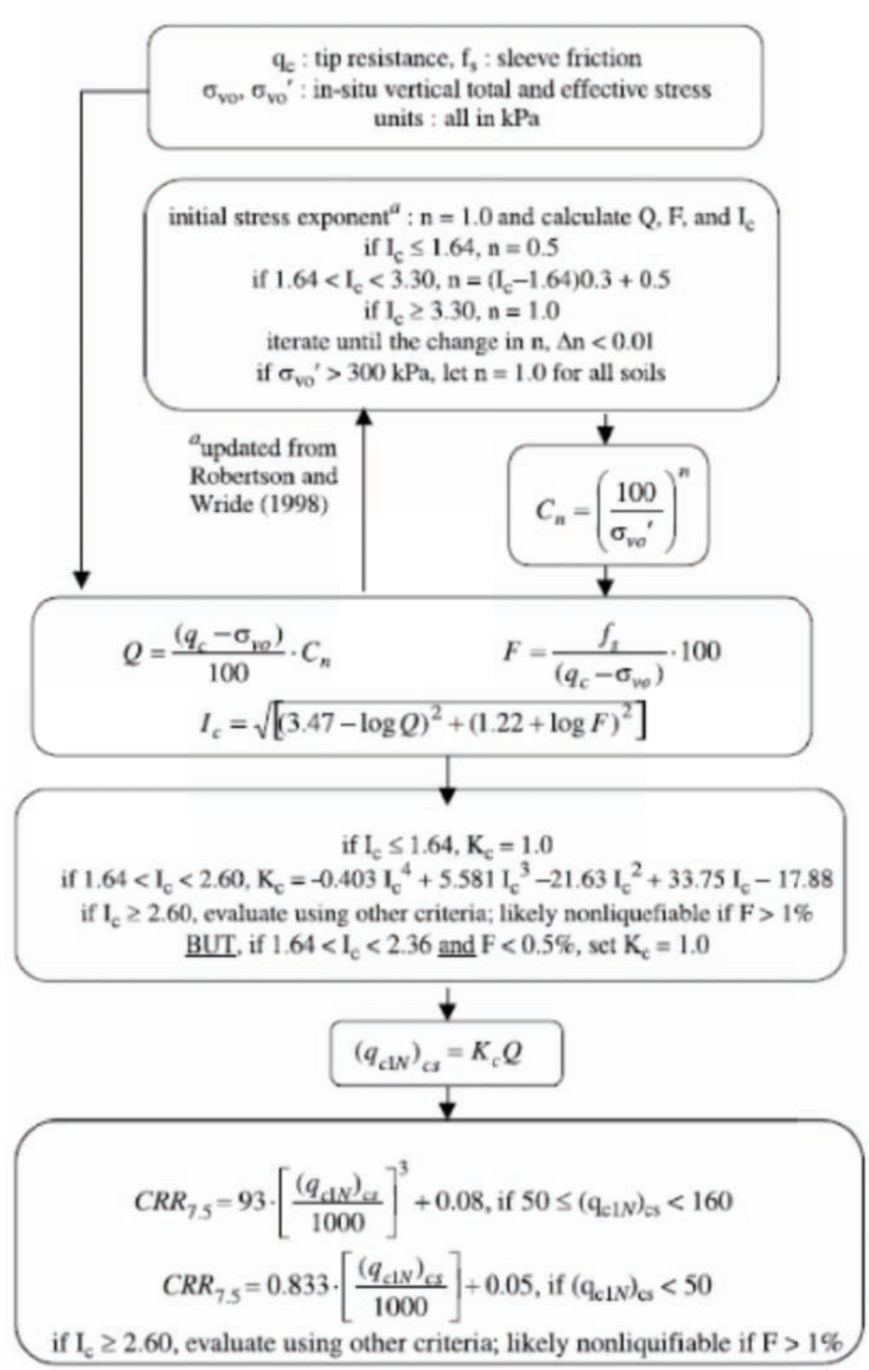


Figure 5-14. CPT-based liquefaction evaluation for sandy soils.

(original diagram by Zhang et al., 2002)

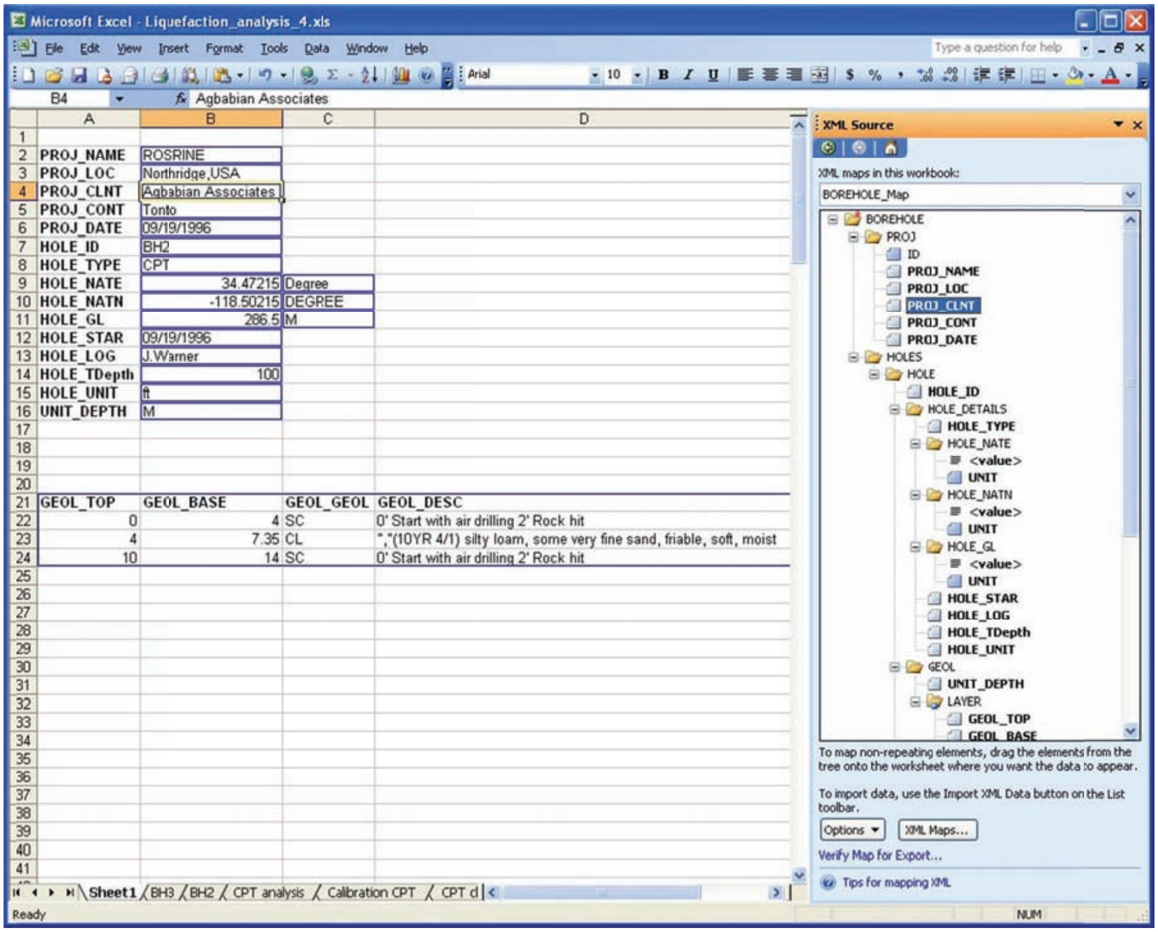


Figure 5-15. SPT- and CPT-based liquefaction evaluation using Web services and XML maps in Microsoft Excel. Borehole data were imported to Excel by querying the GIME Web service in a Visual Basic macro. The data in XML were mapped to the corresponding spreadsheet cells using the XML map feature in Excel.

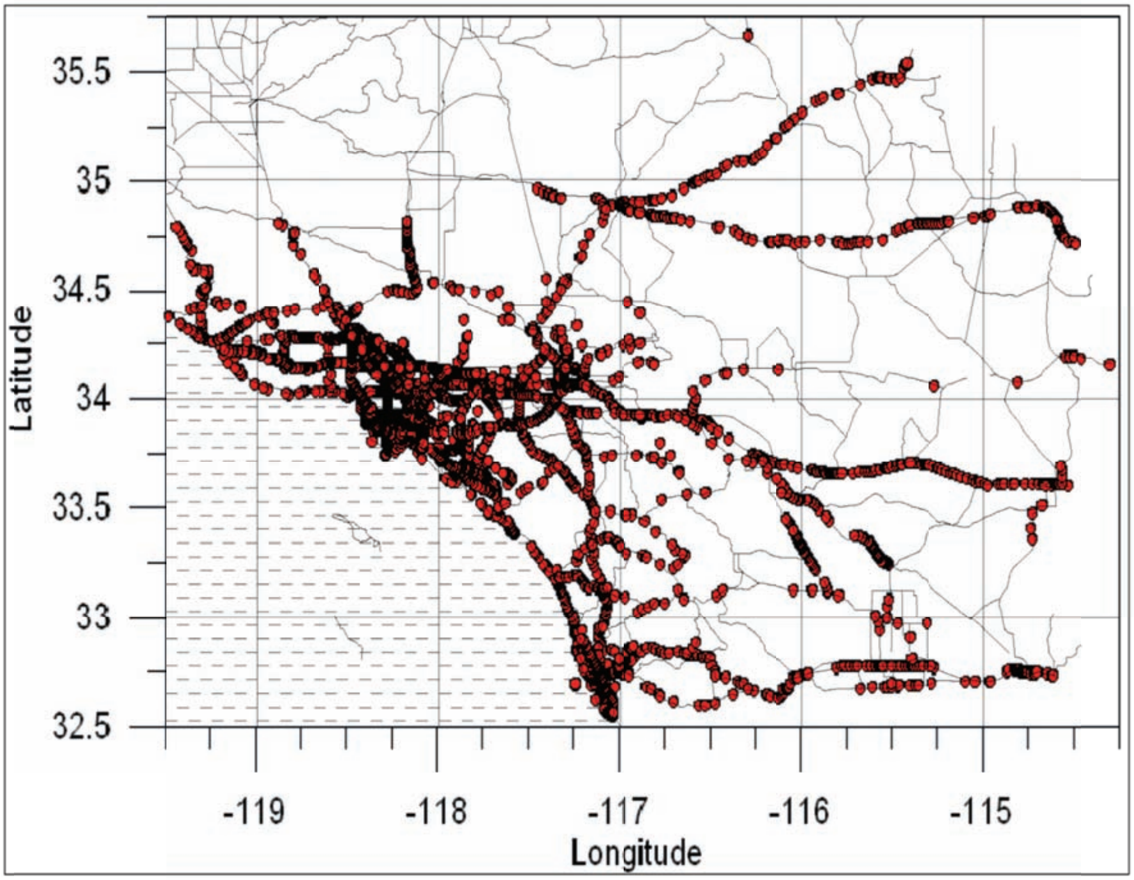


Figure 5-16. The road transportation network evaluated in the REDARS analysis is shown. The red dots show the location of the studied bridges. Total number of bridges was more than 6,000.

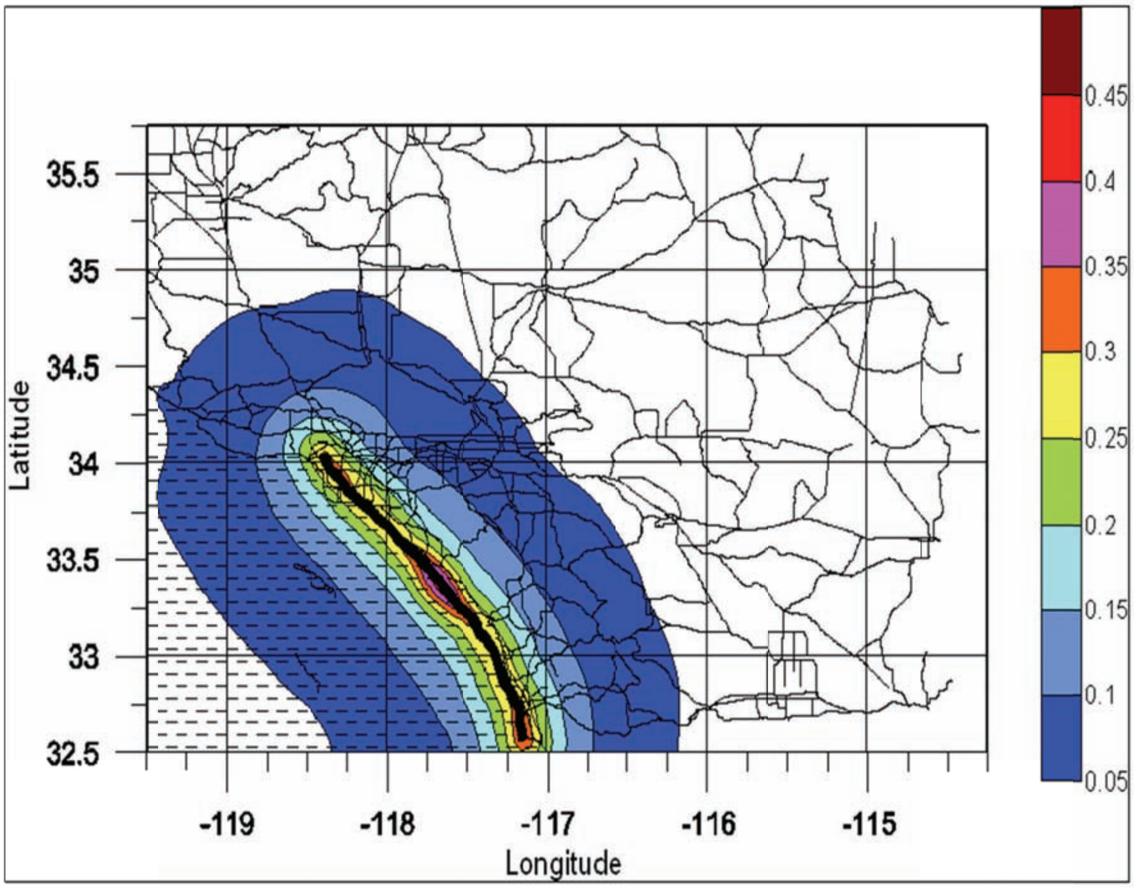


Figure 5-17. Typical deterministic seismic hazard analysis result for the studied network is shown. This figure shows the PGA values for the Newport-Inglewood Fault rupture scenario for Soil Profile Type B/C (V_s=760 m/sec). Similar results were obtained for other spectral periods, rupture scenarios and soil conditions.

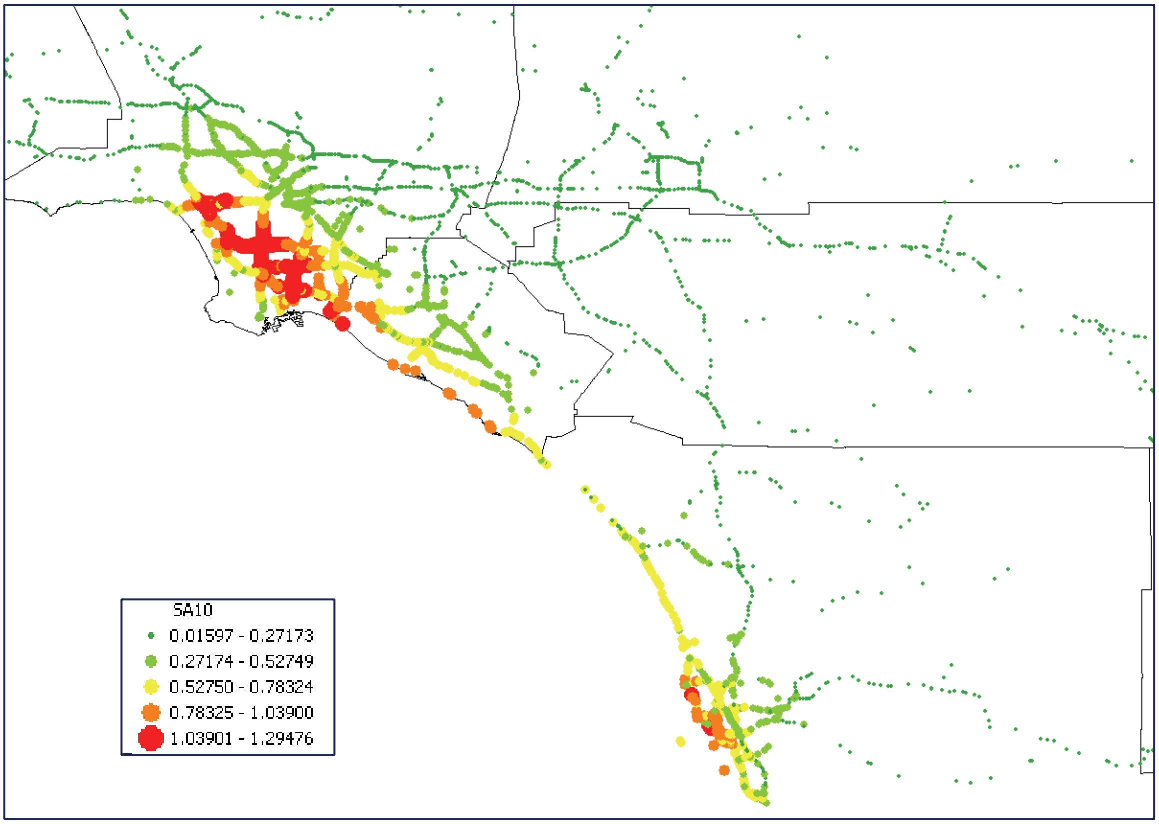


Figure 5-18. The figure shows the spectral acceleration values at 1.0 second for the Newport-Inglewood Fault rupture scenario for the studied bridges. This spectral acceleration was used to assess the damage to the bridges due to earthquake.

(Figure from Cho, 2010)

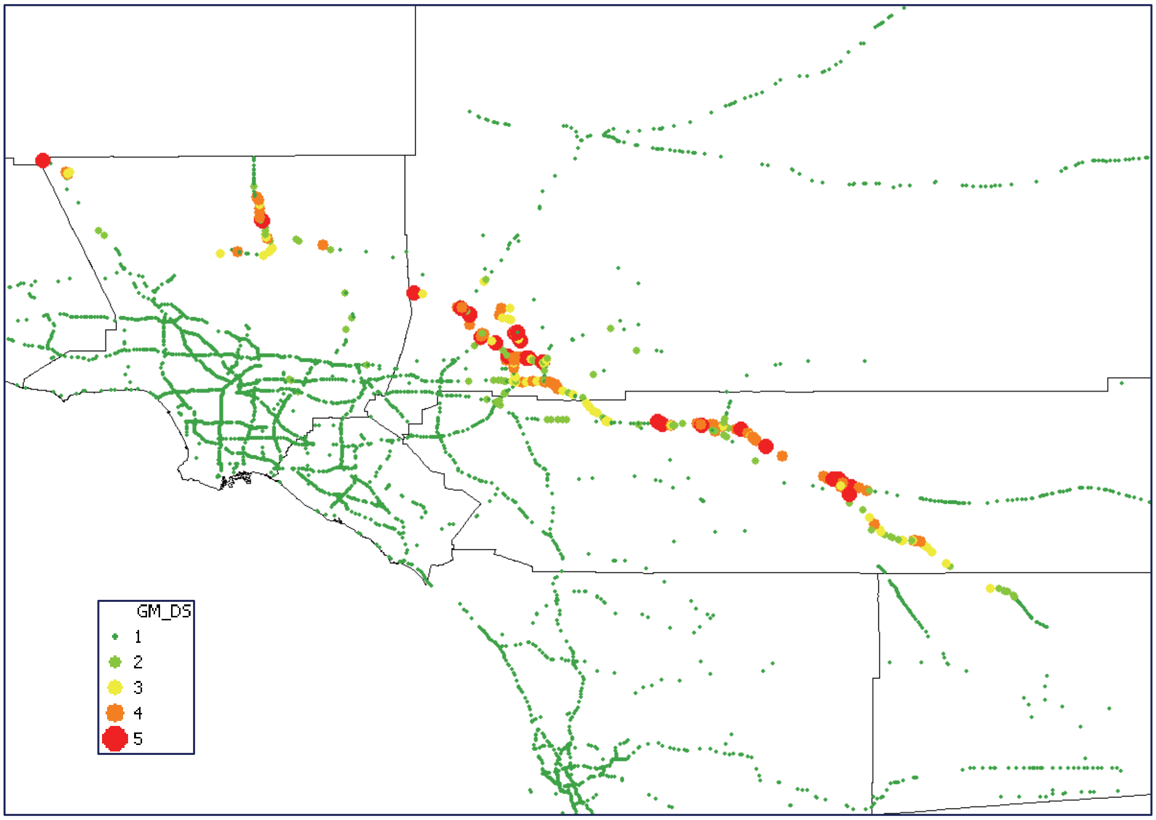


Figure 5-19. The figure shows the damage state for each bridge due to ground motion from the San Andreas Fault Rupture scenario. A damage state of 1 means no damage, while 5 means complete failure. This parameter was used to assess the economic loss due to damage to each bridge structure.

(Figure from Cho, 2010)

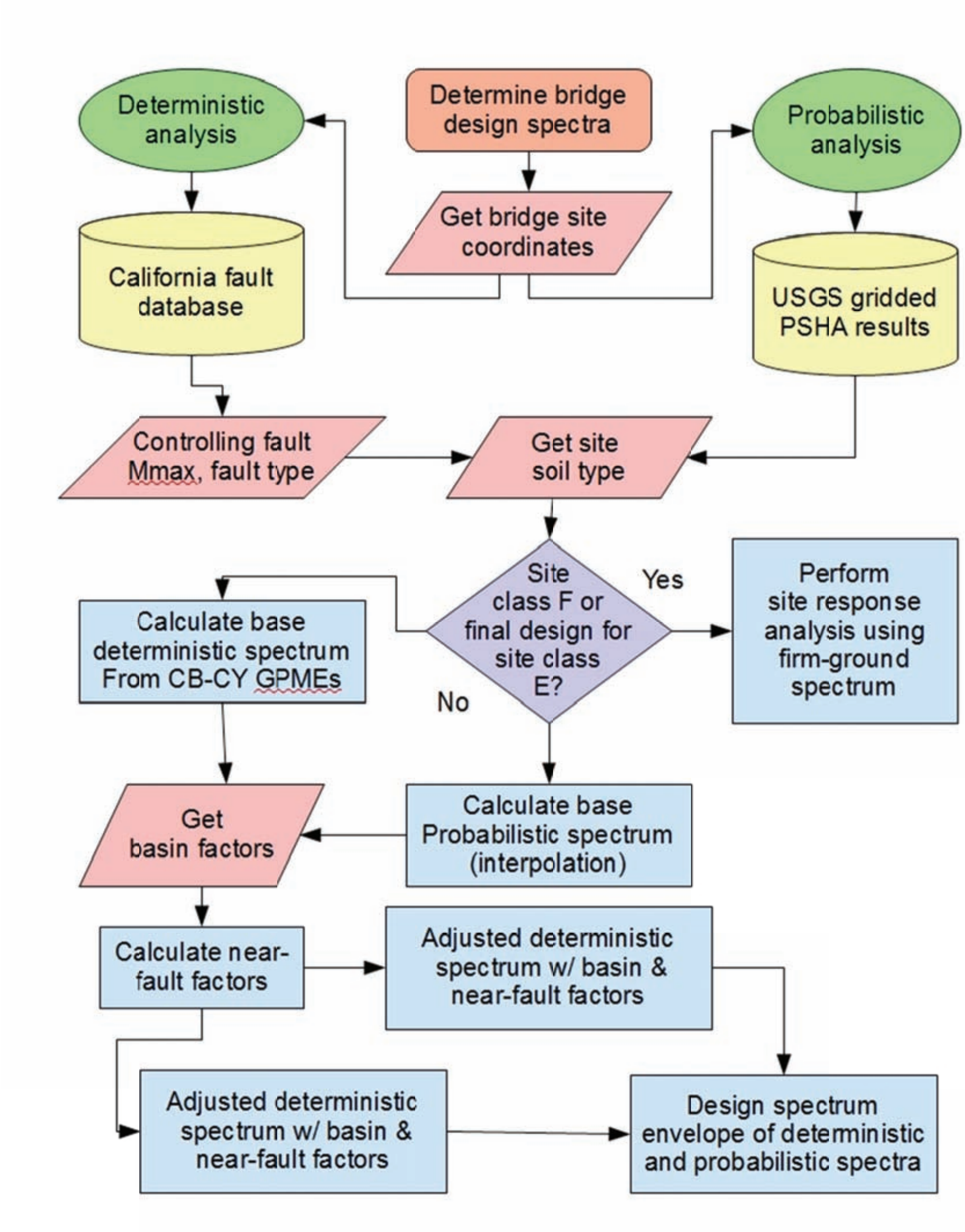


Figure 5-20. Flowchart of Caltrans seismic design criteria (Caltrans, 2010) for highway bridges.

Chapter 6. Summary and Conclusions

In this chapter a summary of the results of the dissertation is presented. The lessons learned and difficulties encountered from each research topic are discussed. The most significant achievements of the research for practical geotechnical engineering applications are highlighted. At the end, suggestions have been made to improve the current results and continue this research and related research lines that have potential to gain significance in the future.

6.1 Summary and Recap

6.1.1 Objectives

The main objective of this research was to explore methods to improve access to geotechnical data for major infrastructure studies that rely on spatial data. In Chapter 1 through several examples it was shown that some of these studies benefit from access to geotechnical data. It was argued that since these types of studies use SDIs for spatial data, the geotechnical data need to be accessible through SDIs. Over the next chapters various technologies needed to integrate the geotechnical data in SDIs were explored.

6.1.2 Review of the Current State of Geotechnical Data Exchange

In Chapter 2 the geotechnical data acquisition, processing and exchange methods were reviewed. The current state of practice for projects involving geotechnical data was discussed. This review included various methods used for data processing and exchange, e.g., proprietary software, data formats and emerging technologies such as Web-based data dissemination. The review included geotechnical software gINT and the COSMOS geotechnical database.

6.1.3 Spatial Databases

In Chapter 3 spatial databases for storage of spatial data types were discussed. Various types of spatial databases and their advantages in handling spatial data were reviewed. It was shown that geotechnical data are inherently spatial, therefore, using spatial databases will improve the speed and efficiency of the operations involving large geotechnical datasets. In addition to improvement in speed and efficiency, spatial databases provide optimized operations on spatial data that can be used by Web services and client applications to simplify their development.

6.1.4 Spatial Data Infrastructures

The concept of SDIs and various standards used by SDI developers were also covered in Chapter 3. These standards included various OGC standards including GML and WFS. The role of these standards in SDIs was discussed. It was demonstrated that a GML-based data model for geotechnical data can be easily integrated with SDIs similar to other types of spatial data.

6.1.5 GML-Based Geotechnical Data Format

In Chapter 4 the advantages of a standard data-transfer format for geotechnical information were discussed. A data format was proposed that supports spatial data types and is used in similar applications in other disciplines. The proposed data format is compatible with GML (Geography Markup Language), which is an XML-based data format. GML and XML compatibility allows the data format to be used with a wide range of standard applications developed for these markup languages. This compatibility reduces the time and resources needed to develop server and client side applications for database systems.

A comparison was made between the proposed data format, and similar initiatives from other researchers and pros and cons of different approaches to encoding geotechnical data in GML were discussed.

6.1.6 Implementation and Case Study

The architecture of a system for distribution and management of geotechnical data was presented. The use of Web services for interaction between database and client-side applications was discussed. It was shown that Web services provide an efficient method for storage and retrieval of data from the database. In addition, Web services can be used to provide various types of data for different client applications. Using Web services for these applications ensures that data can be quickly accessed and retrieved.

As an example, a geotechnical data management system, GIME, was presented. Using this system, two client applications for site-response analysis and liquefaction analysis were explored. A case study was presented in order to illustrate the benefits of this type of geotechnical data management system in infrastructure studies involving spatial data.

6.2 Suggestions for Improvement and Further Research

Further work is needed to complete and improve on this research. Based on the research conducted in this dissertation, a number of topics have been proposed for further research in the following section.

6.2.1 Algorithms for Developing Soil Stratigraphy Models from Borehole Data

In this research correlations were presented to develop soil stratigraphy from in-situ and laboratory soil data associated to a borehole. One-dimensional stratigraphy models were developed using data from a single boring. This assumption simplifies the modeling process, but it results in some limitations for general soil conditions. By using a single borehole for modeling, it is assumed that the soil profile is uniform and the borehole is representative of the soil stratigraphy. While this assumption is valid for some sites, there are cases where the data from several borings should be combined to develop the stratigraphy model. The following scenarios may cause this situation to happen:

- A single boring does not have enough data to develop the soil profile and the data from other boreholes should be included, or additional assumptions needs to be made.
- The data from several boreholes are combined to decrease the spatial variation of soil layers and variability of soil test measurements, or to resolve conflicts between several borings.
- The soil profile is not uniform, and soil profile should be interpolated from several boreholes.
- A uniform soil profile is not applicable, and a two- or three-dimensional soil profile is needed.

Each of these scenarios provides challenging situations that could not be dealt with within the scope of work and limitations of this dissertation and require additional research. Beside knowledge of soil mechanics and geotechnical engineering and computer programming, additional skills in the field of geostatistics and intelligent algorithms are needed.

The author believes that this subject has a lot of potential for research, and the outcomes can be used for a wide range of practical and experimental problems.

6.2.2 Developing a More Comprehensive Earthquake Engineering Database

A similar GML schema can be developed for seismic data. The GML data model may include seismic data including fault geometries and parameters, ground motion attenuation models and strong motion records. The data can be integrated with geotechnical data in seismic hazard analysis and liquefaction analysis applications.

This database can be used to provide Web-based services for earthquake engineering applications. Similar to geotechnical data, earthquake engineering databases are inherently dynamic and need to be frequently updated due to additional data from recent events and advancements in seismology.

6.2.3 Improving the GML-Based Geotechnical Data Format

Although a lot of effort was made to develop a comprehensive GML data model for geotechnical data, the author believes that this format can be improved with additional research. Moreover, additional effort is required to convert the available data from existing borings to the proposed format. Existing boring data are stored in literally hundreds of proprietary digital data formats, in addition to hard copies and raster image formats, which cannot be easily converted to a new digital format. A separate focused research project is required to identify the data sources with the most valuable geotechnical data and to develop methods to convert those to the proposed GML-based format efficiently. The results, however, can be rewarding in increasing the usefulness of the proposed geotechnical database and preserving the investment made in drilling and testing made in the existing boreholes.

6.2.4 Developing Client Software for Other Applications

The pilot applications discussed in this dissertation were mainly aimed at illustrating the concepts, rather than being used in actual analyses. Additional work is needed to develop fully functional applications for more sophisticated analyses and to validate this software.

Bibliography

Aggour, M. S. and Radding, W. R. "Standard Penetration Test (SPT) Correction." Report submitted to Maryland State Highway Administration, Office of Policy and Research, Contract No: SP007B48, 2001.

American Society for Testing and Materials (ASTM). "ASTM Book of Standards, Volume 04.08: Construction: Soil and Rock (I): D420 - D5611," 2007a.

ASTM. "ASTM Book of Standards, Volume 04.09: Construction: Soil and Rock (II): D5714," 2007b.

ASTM."ASTM D1586 - 08a, Standard Test Method for Standard Penetration Test (SPT) and Split-Barrel Sampling of Soils," 2008.

Association of Geotechnical and Geoenvironmental Specialists (AGS). "Electronic Transfer of Geotechnical and Geoenvironmental Data" (Edition 3.1), including addendum, 2005. http://www.ags.org.uk/datatransfer/download.cfm (accessed August 15, 2008).

AGS. "AGS data transfer format example file", 2008. http://www.ags.org.uk/site/datatransfer/example.cfm (accessed August 15, 2008).

Applied Technology Council (ATC). "ATC-32, Improved Seismic Design Criteria for California Bridges: Provisional Recommendations." Report ATC-32, Redwood City, California, 1996.

Bardet, J. P. *Experimental Soil Mech*anics. Englewood-Cliffs, NJ: Prentice Hall, 1997.

Bardet, J. P., Ichii, K. and Lin, C. H. "EERA – A Computer Program for Equivalent-Linear Earthquake Site Response Analyses of Layered Soil Deposits." University of Southern California, Dept. of Civil Engineering, Los Angeles, CA, August 2000.

Bardet, J. P., Tobita, T., Mace, N. and Hu, J. "Regional Modeling of Liquefaction-Induced Ground Deformation." *Earthquake Spectra* 18, no. 1 (2002): 19–46.

Bardet, J. P. and Zand, A. "Spatial Modeling of Geotechnical Information Using GML." *Transactions in GIS* 13, no. 1 (2009): 125–165.

Blasby, D. "Building a Spatial Database in PostgreSQL." PostGIS presentation at the Open Source Database Summit, 2001. http://postgis.refractions.net/files/OSDB2_PostGIS_Presentation.pdf (accessed August 15, 2008).

Bobbitt, J. I. "Unit of Measure Recommendations, Petrotechnical Open Standards Consortium (POSC)", 2008. http://www.posc.org/ebiz/Guidelines/UOMFinalRecs.doc (accessed August 15, 2008).

Bobbitt, J. I. “Examples of XML Dictionary Usage. Petrotechnical Open Standards Consortium (POSC),” 2004. http://www.posc.org/ebiz/Guidelines/XMLDictionaryUsage.html (accessed August 15, 2008).

Bowman, D. “Civil Engineering Data Meets GIS.” *Journal of Computing in Civil Engineering* 12, no. 1 (1998): 5–7.

Burggraf, D. “Standards, GML and AIXM.” PowerPoint presentation, Galdos Systems, Inc., 2010.

Burk, B. G. Handbook of Geotechnical Investigation and Design Tables. New York: Taylor & Francis, 2007.

California Geological Survey (CGS). “Recommended Criteria for Delineating Seismic Hazard Zones in California.” CGS Special Publication 118, May 1992, revised April 2004.

CGS.“Guidelines for Evaluating and Mitigating Seismic Hazards in California.” CGS Special Publication 117A, 2008.

CGS. “CGS Seismic Hazard Zonation Program (SHZP),” 2010. http://www.conservation.ca.gov/cgs/shzp/Pages/Index.aspx, (accessed September 14, 2010).

Caltrans. “Geotechnical Services Design Manual, Version 1.0.” August 2009.

Caltrans. “Caltrans Seismic Design Criteria, Version 1.6.” November 2010.

Canadian Well Logging Society (2008), “LAS - Log ASCII Standard Information.” 2008. http://www.cwls.org/las_info.php (accessed August 15, 2008).

Cho, S. Internal Web Presentation for METRANS/Megacities Project, 2010.

COSMOS/PEER-LL.“Archiving and Web Dissemination of Geotechnical Data: Development of a Pilot Geotechnical Virtual Data Center.” PEER Lifelines Project 2L02 Draft Final Report, December 2004.

COSMOS.“COSMOS Virtual Data Center,” 2010. www.cosmos-eq.org, (accessed January 22, 2011)

Craglia, M. and Annoni, A., “INSPIRE: An Innovative Approach to the Development of Spatial Data Infrastructures in Europe.” In *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ed. Onsrud, H., 93–105. Redlands, CA: ESRI Press, 2007.

Daniel, C. R., Howie, J. A. and Sy, A. (2003), “A Method for Correlating Large Penetration Test (LPT) to Standard Penetration Test (SPT) Blow Counts.” *Canadian Geotechnical Journal* 40, no. 1 (2003): 66–77.

Data Interchange for Geotechnical and GeoEnvironmental Specialists (DIGGS) (2010), DIGGS Homepage, www.diggsml.com (accessed 10/22/2010).

Ellis, R. A. “Jobs, 1996–2006: Strong Growth.” *Engineers* 4, no. 1 (1998): 2–13.

Electric Power Research Institute (EPRI). “Guidelines for Determining Design Basis Ground Motions.” Electric Power Research Institute, Palo Alto, California, 1993.

Federal Emergency Management Agency (FEMA).“HAZUS: FEMA’s Methodology for Estimating Potential Losses from Disasters,” 2010.

http://www.fema.gov/plan/prevent/hazus (accessed September 12, 2010).

Federal Emergency Management Agency (FEMA). “HAZUS-MH MR4 Earthquake Model User Manual,” 2010b.

http://www.fema.gov/plan/prevent/hazus (accessed September 12, 2010).

Federal Emergency Management Agency (FEMA). “HAZUS-MH MR4 Flood Model User Manual,” 2010c.

http://www.fema.gov/plan/prevent/hazus (accessed September 12, 2010).

Federal Highway Administration (FHWA). “Seismic Design of Highway Bridge Foundations.” Report No. FHWA/RD-86/102, June 1986.

Finn, W. D. Liam, Lee, K. W. and Martin, G. R. “An Effective Stress Model for Liquefaction.” *Journal of Geotechnical Engineering Division* 103, no. GT6 (1977): 517–533.

Galdos Systems, Inc. “Developing and Managing GML Application Schemas: A Best Practices Guide prepared by Galdos Systems Inc.” The Ministry of Sustainable Resource Management and GeoConnections, TR2003-232-01, 2003.

GeoServer. “GeoServer 2.1 User Manual,” 2010. http://docs.geoserver.org/stable/en/user/index.html (accessed September 12, 2010).

Geotechnical and Geoenvironmental Software Directory (GGSD).

www.ggsd.com (accessed January 1, 2011).

gINT Softwar. “The gINT Geotechnical and Geoenvironmental Software System,” 2008.

http://www.gintsoftware.com.

Global Spatial Data Infrastructure (GSDI).*GSDI Newsletter,* April 2006.

GSDI. “SDI Cookbook,” 2010. http://www.gsdidocs.org/GSDIWiki/index.php/Main_Page, (accessed September 12, 2010).

Goodman, R. E. “Karl Terzaghi’s Legacy in Geotechnical Engineering.” Geo-Strata, October 2002.

Google Inc. Google Earth Website, 2008a. http://earth.google.com

Google Inc. “KML Documentation Introduction,” 2008b. http://code.google.com/apis/kml/documentation (accessed August 15, 2008).

Grigg, S. N. “Demographics and Industry Employment of Civil Engineering Workforce.” *Journal of Professional Issues in Engineering Education and Practice* 126, no. 3 (2000): pp. 116-124.

Heslop, K., Karst, J., Prensky, S. and Schmitt, D. “LAS 3.0, Log ASCII Standard Document #1 - File Structures.” Canadian Well Logging Society, 2003.

Hunt, R. E. Geotechnical Engineering Investigation Handbook, Second Edition. New York: Taylor and Francis Group, 2005.

Kobe City.“Geotechnical Database KOBE JIBANKUN.” http://www.kobe-toshi-seibior.jp/matisen/jibankun/index.htm (accessed September 12, 2010).

Kramer, S. L. *Geotechnical Earthquake Engineering*. Englewwood-Cliffs, NJ: Prentice Hall, 1997.

Kuhn, W. “Introduction to Spatial Data Infrastructures.” Presentation held on March 14, 2005.

Kulhawy, F. H. and Mayne, P. W. “Manual on Estimating Soil Properties for Foundation Design.” Report No. EL-6800, Electric Power Research Institute, Palo Alto, California, August 1990.

Kunapo, J., Dasari, G. R., Phoon, K. K. and Tan, T. S. “Development of a Web GIS-Based Geotechnical Information System.” *Journal of Computing in Civil Engineering* 19, no. 3 (2005): 323–327.

Lamb, T. W. and Whitman, R. V. *Soil Mechanics*. New York: John Wiley and Sons, 1969.

Mayne P. W., Christopher, B. R. and DeJong, J. “Manual on Subsurface Investigations.” National Highway Institute, NHI-01-031, 2001.

Mayne, P. W. and Rix, G. J. “Correlations between Shear Wave Velocity and Cone Tip Resistance in Natural Clays.” *Soils and Foundations* 35, no. 2 (1995): 107–110.

MCEER. “Risk Due to Earthquake Damage to Highway Systems (REDARS),” 2010. http://mceer.buffalo.edu/research/redars/ (accessed September 12, 2010).

Meyerhof, G. G. “Penetration Tests and Bearing Capacity of Cohesionless Soils.” *Journal of the Soil Mechanics and Foundation Engineering* 82, no. 1 (1956).

Microsoft. Microsoft Office. http://office.microsoft.com.

Mokarram, N. “Data Exchange in Geotechnical Engineering.” Ph.D. Dissertation, Dept. of Civil Engineering, University of Southern California, 2010.

National Geospatial Intelligence Agency "NGA announces Requirements for OGC and Complementary Standards," 2005. http://www.nga.mil/NGASiteContent/StaticFiles/OCR/nga0518.pdf (accessed September 12, 2010).

Nationl Spatial Data Infrastructure. NSDI Website. www.fgdc.gov/nsdi/nsdi.html, (accessed March 22, 2011).

Nebert, D., Reed, C. and Wagner, R. M. "Proposal for a Spatial Data Infrastructure Standards Suite: SDI 1.0." In *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ed. Onsrud, H. Redlands, CA: ESRI Press, 2007.

Ohta, Y. and Goto, N. "Empirical Shear Wave Velocity Equations in Terms of Characteristic Soil Indexes." *Earthquake Engineering and Structural Dynamics* 6 (1978): 167–187.

Onsrud, H. Research and Theory in Advancing Spatial Data Infrastructure Concepts. Redlands, CA: ESRI Press, 2007.

Open Geospatial Consortium (OGC). "Simple Features Specification for SQL." Open Geospatial Consortium (formerly called the Open GIS Consortium), Wayland, MA, 1999.

OGC. "Geography Markup Language, Encoding Standards," 2008. http://www.opengeospatial.org/standards/gml (accessed August 15, 2008).

OGC. "GML 2.1.2 Validation," 2008. http://cite.opengeospatial.org/test_engine/gml/2.1.2 (accessed August 15, 2008).

Parsons, R. L. and Frost, J. D. "Interactive Analysis of Spatial Subsurface Data Using GIS-Based Tool." *Journal of Computing in Civil Engineering* 14, no. 4 (2000): 215–222.

Peck, R. B., Hanson, W. E. and Thornburn, T. H. *Foundation Engineering*. New York: John Wiley, 1974.

Petrotechnical Open Standards Consortium (POSC)."POSC Web Site Specifications," 2008. http://www.posc.org (accessed August 15, 2008).

Ponti, D., Hoit, M., Mcvay, M., Styler, M., Benoit, J., Bobbitt, J., Caronna, S., Chandler, R. and Quinn, P. "DIGGS: An XML-Based Interchange Standard for Geotechnical and Geoenvironmental Data." Workshop for Site Selection, Installation and Operation of Geotechnical Strong-Motion Arrays: Workshop 2, Guidelines for Installation, Operation, and Data Archiving and Dissemination, San Diego Supercomputer Center, University of California San Diego, San Diego, CA, 2006.

Robertson, P. K. "Soil Classification Using the Cone Penetration Test." *Canadian Geotechnical Journal* 27, no. 1 (1990): 151–158.

Robertson, P. K. and Cabal, K. L. (2009), “Guide to Cone Penetration Testing.” Gregg Drilling and Testing Inc., 3rd Edition, January.

Robertson, P. K., Campanella, R. G. and Wightman, A. “SPT-CPT Correlations.” *Journal of Geotechnical Engineering* 109, no. 11(1983): 1449–1459.

Robertson, P. K., Campanella, R. G., Gillespi, D. and Grieg, J. “Use of Piezometer Cone Data.” Proceedings, In-Situ ’86, ASCE Specialty Conference, Blacksburg, VA, 1986.

Rockware, Inc. Rockworks 2006.
http://www.rockware.com/product/overview.php?id=165, (accessed August 15, 2008).

Schnabel, P. B., Lysmer, J. and Seed, H. “SHAKE: A Computer Program for Earthquake Response Analysis of Horizontally Layered Sites.” Earthquake Engineering Research Center Report No. 72–12, 1972.

Seed, H. B. and Idriss, I. M. “Soil Moduli and Damping Factors for Dynamic Response Analyses.” Technical Report EERC Report, 70-10, University of California, Berkeley, 1970.

Shekhar, S. and Chawla, S. *Spatial Databases: A Tour*. Englewood-Cliffs, NJ: Prentice Hall, 2003.

Southern California Earthquake Center (SCEC). “Community Modeling Environment (CME)”, 2010.

http://epicenter.usc.edu/cmeportal (accessed September 19, 2010).

Southern California Earthquake Center (SCEC). “ITR/AP: The SCEC Community Modeling Environment—An Information Infrastructure for System-Level Earthquake Research.” SCEC ITR Proposal, 2010b

http://www.scec.org/aboutscec/documents/ITR.FINAL.pdf, (accessed September 19, 2010).

Sykora, D. W. “Examination of Existing Shear Wave Velocity and Shear Modulus Correlations in Soils.” Department of the Army, Waterways Experiment Station, Corps of Engineers, Miscellaneous Paper GL-87-22, 1987.

Tanaka, Y. and Okimura, T. “In-Depth Geotechnical Data Base, Kobe Jibankun, for Seismic Hazard Study.” Research Center for Urban Safety and Security (RCUSS), Kobe University, Japan, 2010.

Techbase International, Ltd. Techbase Geotechnical Engineering Package,
http://www.techbase.com/geotecnical.html (accessed August 15, 2008).

The Carbon Project. Gaia 3 Viewer.
http://www.thecarbonproject.com/gaia.php (accessed August 15, 2008).

The White House ,Office of Management and Budget. “Circular No. A-16.” Revised, August 19, 2002.

Towhata, I. *Geotechnical Earthquake Engineering*. Berlin: Springer-Verlag, 2008.

Turner, L. L. "Examples of Geotechnical Data and Document Management Systems in the United States." Presentation at AGS, June 11, 2008.

http://www.cosmos-eq.org/vdc/PEER-LLProject-2L03.pdf (accessed September 12, 2010).

Verruijt, A., Beringen, F. L. and De Leeuw, E. H. *Penetration Testing: Standard Penetration Test/Cone Penetration Test*. Rotterdam, Netherlands: Balkelma, 1983.

Vucetic, M. V. and Dobry, R. "Effect of Soil Plasticity on Cyclic Response." *Journal of Geotechnical Engineering* 117 (1991): 89–107.

Werner, S. D., Taylor, C. E., Cho, S., Lavoie, J. P., Huyck, C., Eitzel, C., Chung, H. and Eguchi, R. T. "REDARS 2 Methodology and Software for Seismic Risk Analysis of Highway Systems." Special Report MCEER-06-SP08, 2006.

World Wide Web Consortium. "Extensible Markup Language (XML)."

http://www.w3.org/XML (accessed August 15, 2008).

World Wide Web Consortium. "XSL Transformations (XSLT)."
http://www.w3.org/TR/xslt (accessed August 15, 2008).

Yeung, A. K. W. and Hall, G. B. "Spatial Database Systems, Design, Implementation and Project Management." Springer, P.O. Box 17, 3300 AA Dordrecht, The Netherlands, 2007.

Youd, T. L., Idriss, I. M., Andrus, R. D., Arango, I., Castro, G., Christian, J. T., Dobry, R., Liam Finn, W. D., Harder, L. F., Hynes, M. E., Ishihara, K., Koester, J. P., Liao, S. S. C., Marcuson, W. F., Martin, G. R., Mitchell, J. K., Moriwaki, Y., Power, M. S., Robertson, P. K., Seed, R. B. and Stokoe, K. H. "Liquefaction Resistance of Soils: Summary Report from the 1996 NCEER and 1998 NCEER/NSF Workshops on Evaluation of Liquefaction Resistance of Soils." *Journal of Geotechnical and Geoenvironmental Engineering* 127, no. 10 (2001): 817–833.

Zhang, G., Robertson, P. K. and Brachman, R. "Estimating Liquefaction Induced Ground Settlements from the CPT." *Canadian Geotechnical Journal* 39 (2002): 1168–1180.

Zimmermann, R., Ku, W. S., Wang, H., Zand, A. and Bardet, J. P. "A Distributed Geotechnical Information Management and Exchange Architecture." *IEEE Internet Computing* 10, no. 5 (2006): 26–33.

Appendix. Seismic Hazard Analysis for Port of Los Angeles and Port of Long Beach Transportation Network Bridges

Introduction

In this study a seismic risk analysis (SRA) for the Port of Los Angeles and Port of Long Beach road transit network bridges was performed using the REDARS methodology. The objective of the study was to estimate the direct damage to the structures due to earthquake, as well as the economic loss in the post-earthquake state of the transportation network. In REDARS analysis the later risk is typically quantified as economic loss due to reduction in traffic volume and increase in travel time after earthquake.

This report covers the seismic hazard analyses to determine the ground motion levels at each bridge site. The results of this study were used as input data in the REDARS analysis to assess vulnerability of the transportation network components (bridges) to earthquake hazard. The results of the study can help to identify network components that have the most vulnerability and generate the largest economic loss due to considered earthquake scenarios.

REDARS Methodology for Seismic Risk Analysis

A detailed description of the REDARS analysis can be found in “REDARS 2 Methodology and Software for Seismic Risk Analysis of Highway Systems” (Werner et al., 2007). In this section a summary of the methodology based on the above reference is provided.

REDARS 2 can be used to carry out deterministic or probabilistic SRA for any user-defined transportation network in the United States. In probabilistic SRA, the system is analyzed for multiple simulations. Each simulation is defined as the system response to a set of input and model parameters. In deterministic analysis, the system response is evaluated for one set of input parameters that might represent the median level or a randomly selected set of parameters.

REDARS uses a combination of data from various disciplines. These disciplines include geoseismic, geotechnical and structural engineering, construction/repair data, transportation network characteristics and economic models. The SRA methodology in REDARS involves estimating the following steps (Werner et al., 2007):

- ***"Hazards.*** Seismic hazards at the site of each component in the highway system.
- ***Component Performance.*** Each component's damage state and traffic state due to these site specific seismic hazards, in which the traffic state reflects the component's ability to carry traffic at various times after the earthquake as the damage is being repaired.
- ***System Performance.*** System-wide traffic flows (e.g., travel times, paths, and distances)throughout the system, also at various times after the earthquake, that are dependent on each component's traffic state, the redundancies and traffic-carrying capacities of the various roadways that comprise the system, and the trip demands (i.e., the number, type, origin, and destination for all trips that use the highway system).
- ***Losses.*** Consequences of earthquake-induced damage to the highway system, including: (a) economic impacts (repair costs and losses due to travel time delays); increases in

travel times to/from key locations in the region (e.g., medical facilities, airports, centers of commerce, etc.); and (c) increases in travel times along “lifeline” routes within the system, which are previously designated routes that are essential for emergency response or national defense.”

Figure A-1 shows an outline of a typical REDARS analysis for a transportation network. REDARS is a modular program, seismic hazard, component response and economic losses are evaluated in separate modules and the results are used in the subsequent analyses. A complete analysis involves an initialization process, in which input data and model parameters are setup. The second process is system analysis for a given set of parameters, where the response of system components to various seismic hazards (e.g. ground shaking, liquefaction and surface fault rupture) is evaluated. The seismic hazard is evaluated in the Hazards Module, as shown in the flowchart in Figure A-2. Other program modules include System Module, Economic Module and Component Module.

Other input data are also shown in Figure A-2 . This figure shows that a REDARS analysis uses a large number of input parameters. Some of the required input data include roadway topology and attributes, bridge locations and attributes, origin-destination (O-D) zones and pre-earthquake trip tables. In order to facilitate model initialization, REDARS has an import wizard which can be used to enter various parameters from external databases. In addition to accessing the database and importing the data, this wizard will guide the user during the process, check the consistency of the imported data and validates the resulting transportation network and connectivity/continuity of the O-D zones. Figure A-3 shows how public external databases are used through import wizard to initialize a REDARS analysis. These databases include National

Highway Planning Network, Highway Performance Monitoring System, National Bridge Inventory, NEHRP Soil Types, Origin-Destination (O-D) Data and O-D Zones.

REDARS can produce results in various forms. They can be used to assess pre-earthquake conditions and evaluate several options for reducing seismic risk. They can also be used in post-earthquake conditions in real time to evaluate the effectiveness of different measures in reducing traffic congestion and improving network condition.

The role of each program module is discussed as follows:

System Module

The system module in REDARS contains the transportation network model, input data defining this network and various performance metrics (e.g. traffic flows and traffic times) at different times after earthquake.

Hazard Module

The hazard module includes the input parameters and models used to determine the seismic hazard within the network. Currently these hazards include ground shaking, liquefaction and fault rupture displacements. It doesn't include the seismic landslide and tsunami hazard at this time. These data include earthquake scenarios to be considered in the analysis, soil types for determining site response and liquefaction hazard, and faults within the network that can develop surface displacements.

Component Module

The component module contains the information on components including their response to earthquake hazard, the components damage state and repair information (e.g. the type, cost and duration of repairs), and the components traffic data. These parameters are time dependent and will be updated for various times after the event.

Economic Module

The economic module models the cost of repairs and economic losses due to increased travel time and reduction in traffic volume. Currently it doesn't include the secondary economic impact due to decrease in economic activity and trip demand.

Analysis Procedure

A regular REDARS analysis involves the following four steps:

Step 1: System Initialization.

During this step these items are initialized: (a) transportation network, (b) component locations and attributes, (c) origin-destination zones, and (d) various modeling and analysis options. These parameters can be initialized using user input, walkthrough tables or import wizard.

Step 2: System Analysis

In this step the system is analyzed for one scenario, including one particular earthquake and one set of network, component and system parameters. For deterministic analysis, either median values are used or a randomly selected set of values of uncertain parameters are used. For probabilistic analysis, the system analysis represents one of the multiple earthquake and system state combinations.

System analysis includes hazard evaluation, direct loss and system state evaluation, transportation network analysis and economic impact evaluation.

Scope of this Study

In this study the ground shaking hazard for bridges due to several earthquake scenarios was evaluated using a deterministic approach. The results from this study were imported to REDARS hazard module to provide ground shaking levels for REDARS analysis. These results were combined with data from other modules (e.g. system, economic and component modules) to provide the data and parameters for SRA.

The latest procedures and sources of data used by the California Department of Transportation (Caltrans) and the United States Geological Survey (USGS) were incorporated in this study. These procedures are currently used by Caltrans for design of new bridges and seismic retrofit of existing bridges in California. The entire procedure can be found in the Caltrans Seismic Design Criteria (SDC) Version 1.5 (Caltrans, 2009b).

In addition to ground shaking, earthquakes cause hazards due to fault rupture, liquefaction, lateral spreading, landslides, and tsunami. This study was limited to estimation of ground shaking hazard and other causes of damage were not considered.

Overview of the Network

Los Angeles metropolitan area has one of the densest road networks in the United States. This network includes a large number of road bridges, including river crossings, viaducts and freeway overcrossing and undercrossings. The transit network of Ports of Los Angeles and Port of Long Beach uses this road network to transport cargo through the entire United States. Any disruption in traffic flow or travel times will impact the economic activity and revenue in these two ports. Therefore study of earthquake impact on this system is significant for the ports.

Bridges are one of the critical components of the transportation network. Major damage to bridges can result in direct loss of life and financial loss due to replacement costs. In addition, bridge inoperability might lead to secondary loss of life and monetary loss due to traffic disruption during and in aftermath of a catastrophe. Repairing damaged bridges after an earthquake is time consuming, expensive and sometimes technically challenging. Due to these facts studying the vulnerability of the bridges due to earthquake event and identifying critical bridges is a major part of risk assessment for a transportation network.

In this study we estimated the ground motion levels at bridge sites in the southern California transportation network due to several considered earthquake scenarios. For this analysis we used a deterministic approach, which means the ground shaking amplitude was evaluated separately for a number of given earthquake scenarios. In a probabilistic approach, on the other hand, multiple scenarios are evaluated and the associated risks are combined to estimate the commutative hazard.

In the following sections the assumptions and procedure used in this analysis is discussed in more detail.

Bridge Locations

Bridge locations for this analysis were provided by METRANS Transportation Center (METRANS). There were a total of 6,353 bridges in the project network. For each bridge location the coordinates in UTM and WGS84 coordinate systems were provided. All bridges were located between latitude 32.5° and 35.7°, and longitude-114.2° and -119.5°. Based on these coordinates all bridges were in UTM Zone 11N. Figure A-4 shows the approximate location of the bridges superimposed on the Southern California road network.

Considered Earthquake Scenarios

Three earthquake scenarios were considered in this study. These scenarios are generated by events on the Newport-Inglewood, Palos Verdes, and South San Andreas faults. The first two faults are the dominant faults in the Port of LA and Port of Long Beach area, and can potentially cause the most significant damage to the road network in the vicinity of the ports. The South San Andreas Fault has the largest magnitude and longest length in the study area, and an event on this fault can impact most of the study area and cause significant damage to the entire network.

Ground Shaking Estimation Methodology

A deterministic approach was used to calculate the seismic criteria for each bridge site. The deterministic approach is appropriate for this study as each scenario is evaluated as a separate event and the response of different components of the system is evaluated simultaneously to the causative earthquake.

The peak ground acceleration (PGA) and response spectrum amplitude at 0.3 second and 1 second was used to estimate the damage to the system components.

A typical deterministic seismic hazard analysis for bridges is performed using the following data:

- Ground motion prediction equations (GMPEs): These empirical equations are used to estimate ground motion at a site due to a given fault rupture scenario. GMPEs are usually developed based on statistical analyses on recorded ground motions from past earthquakes.
- Characterization of fault sources: seismic hazard in the Southern California is governed by shallow crustal earthquakes generated by active (late Quaternary) faults. Faults in deterministic seismic hazard analysis are modeled by simplified planar geometries. Fault planes are defined by surface alignment, dip angle, dip direction, and depth to top and bottom of the rupture plane. Together with maximum moment magnitude and style of faulting, these parameters characterize the seismic hazard associated with each fault source.
- Characterization of the site: sites are characterized by distance from fault, soil profile and a number of other parameters that vary between different GMPEs. The soil profile is usually characterized by the average shear wave velocity of the soil in top 30 meter or 100 feet (Vs30).

In California, there are several seismic design codes in effect that govern design of different types of structures. For regular road bridges, the seismic design is performed according to the Caltrans SDC. Caltrans released a new version of this document (Version 1.5) in September 2009 which introduced several major changes in the seismic design procedure. In comparison to the previous version (Version 1.4, June 2006) there are several improvements in the new code. The two most significant changes are introduction of the new generation of GMPEs, and using a probabilistic seismic hazard analysis based on USGS 2008 United States national seismic hazard maps in tandem with the deterministic approach used by previous versions of SDC.

For this study, we have used the deterministic criteria of the new SDC. This procedure is discussed in detail in Shantz and Merriam (2009). The probabilistic criteria cannot be used in this study because in that approach the ground motion amplitudes at different locations are not due to the same earthquake scenario and therefore cannot happen concurrently.

Ground Motion Prediction Equations

SDC 1.5 uses two new generation GMPEs for deterministic seismic hazard analysis: these are the Campbell-Bozorgnia GMPE (CB) (Campbell and Bozorgnia, 2008) and the Chiou-Youngs GMPE (CY) (Chiou and Young, 2008). These GPMEs are applicable to all fault sources in California. The details of these GMPEs can be found in their respective references provided above. For this study we used these two GMPEs. The deterministic spectrum for each bridge site was calculated as the arithmetic average of median response spectra predicted by CB and CY GMPEs.

Characterization of Fault Sources

Caltrans SDC provides the following parameters for fault sources: fault dip, dip direction, depth to the top of rupture plane and depth to the bottom of rupture plane, style of faulting, and maximum moment magnitude (Mmax). For this study we used the fault parameters provided in the SDC fault database. However, the SDC database does not provide the fault trace coordinates and recommends consulting a geological fault map to determine the site to fault distance. For this purpose we selected the same coordinates used by USGS in 2008 Unites States national seismic hazard analysis to characterize the surface geometry of the three faults considered in the analysis.

It should be noted that in USGS database some faults have alternative alignments, for example both Newport-Inglewood and Palos Verdes faults have a stitched alternative that model the fault with surrounding smaller faults as a single larger fault. In addition, the main branch of Newport-Inglewood fault has two alternatives. For our study, we used the South San Andreas Fault, stitched model of Palos Verdes fault, and both alternatives of the stitched model for Newport-Inglewood fault.

For South San Andreas Fault only the segments south of Parkfield segments were used. The Parkfield segment and northern segments are not relevant because they are further north relative to the bridge sites. Based on the USGS model only the main branch of San Andreas fault was used and smaller fault traces were not considered.

For Newport-Inglewood fault, the larger of the spectral accelerations from the two alternatives were selected.

It should be noted that Caltrans has revised fault dip angle for on-shore segment of Palos Verdes fault (Palos Verdes Hills segment) from earlier vale of 60 degrees dipping toward south-west to 90 degrees (vertical). We have used the new model which is also consistent with USGS model. Using the earlier model can result in higher spectral accelerations for sites located toward south and west of the Palos Verdes fault.

Table A-1 summarizes the fault source parameters used in this study. The coordinates of the faults are shown in Table A-2.

Site Characterization

Site characterization for CB and CY GMPEs is done using the following parameters:

- Distance from rupture plane, R_{rup}: This is the closest distance from the site to the rupture plane of the fault.

- Distance from surface projection of the rupture plane, R_{JB} (the Joyner-Boore distance): This is the closest distance from the site to the surface projection of the rupture plane. For a strike-slip fault $R_{RUP} = R_{JB}$.

- Distance from fault trace or fault trace extension, R_X: This is the closest distance from the site to the fault trace or fault trace extension. It is measured perpendicular to the fault trace.

- Basin Effect: earthquake record from deep sedimentary sites such as LA basin show amplification at long period comparing to similar sites with no deep basins. This effect is represented in CB and CY models by depth to bedrock. For CB model, depth to bedrock is defined as depth at which the shear wave velocity of the soil is 1000 m/sec ($Z_{1.0}$). For CY model depth to bedrock is defined as the depth at which shear wave velocity of soil is 2.5 km/sec ($Z_{2.5}$). Due to large number of sites, it was not possible to manually estimate $Z_{1.0}$ and $Z_{2.5}$ from contour maps provided in SDC. We used the Community Velocity Model (CVM) Version 4.0 (CVM 2009) directly for bridge sites coordinates. This model which is used to develop SDC basin depth maps calculates the shear wave velocity for a given depth. In order to estimate the $Z_{1.0}$, shear wave velocity was estimated at depths between 350 and 1000 meter at 50 meter intervals. The depth corresponding to 1000 m/sec was estimated from the calculated shear wave velocities using linear interpolation. Similarly for $Z_{2.5}$, shear wave velocity was estimated at depths between 2500 and 6000 meter was estimated at 500 meter intervals. The depth corresponding to 2.5 km/sec was estimated using linear interpolation.

- Shear wave velocity of the site soil (V_{s30}): This parameter is usually measured or estimated as the average shear wave velocity in the top 30 meter (100 ft) of soil. For this study V_{s30} values were estimated based on the bridge site soil profile types obtained from Mr. Tom Shantz of Caltrans. These soil profile types are not based on borehole information at bridge sites, instead they rely on correlations between V_{s30} and surface geology from geological maps. We estimated Vs30 from soil profile type using the correlations provided in Table A-3. Soil profile type was not available for all bridges. The response spectra for sites where soil profile type was not available was determined assuming soil profile Type D and V_{s30} = 270 m/s. Due to large number of bridges better estimation of soil profile type based on soil borings was not possible for this study.

A Note on Shear Wave Velocity Selection

It should be noted that for sites where V_{s30} is less than 150 m/s (500 fps), the C-B and C-Y GMPEs are not applicable. Also if V_{s30} is larger than 1500 m/s (very rare in California), a value of 1500 m/s should be used in the analysis. Caltrans SDC mandates that if (a) V_{s30} is less than 150 m/s, (b) one or more soil layers of at least 5 feet thickness has a shear wave velocity smaller than 120 m/s, or (c) the profile conforms to Type E criteria, a site-specific response analysis is required for determination of final design spectrum. For Type F soil profiles site-specific site response is required for both preliminary and final design spectrum. Due to limitations of this study and lack of availability of actual soil profiles for each bridge, we were not able to perform a site-specific site response analysis.

Ground Shaking Estimation Results

Based on the procedure described above, deterministic response spectra at each site for three faults were calculated. Spectral values corresponding to PGA, 0.3 second and 1.0 second were determined. The results for each bridge site are provided in electronic format. For comparison, the contour plots of the PGA values for each scenario is shown in Figures 2 to 4, for Palos Verdes, Newport-Inglewood and South San Andreas faults, respectively. The PGA in these figures is based on a soil profile type of B/C (V_s = 760 m/s). The values in the PGA values in the ocean should be disregarded. A grid distance of 0.05 degrees in longitude and latitude directions was used to obtain these results. The PGA value at close distance to faults might be somehow inaccurate due to errors in triangulation resulting from distance between grid points.

Comparison with ARS Online

In order to validate the results obtained from this study, the spectral values were compared with values calculated by Caltrans' ARS Online (Caltrans 2009a), a web-based tool which is developed to calculate design ARS curve for bridge sites in State of California. ARS Online provides results for deterministic and probabilistic analyses. Only deterministic results should be compared with the results of this study because probabilistic results are not representative of distinct scenarios studied here. Moreover, ARS Online provides the deterministic results for the controlling faults for each site. Therefore comparison can be made only if at least one of the controlling faults is represented in this study. The comparison results for 7 random cases are shown in Table A-4.

Discussion

Based on the results shown in Table A-4, the following trends were observed:

- In general the values estimated in this study match favorably with values predicted by ARS Online.
- Small variation is results were expected due to different fault trace coordinates used in this study (based on USGS models) and coordinates used by Caltrans. Caltrans digitized fault models were not available. Based on the results shown in Table A-4 both models result in similar spectral accelerations and the difference is insignificant. In addition, Caltrans states that coordinates used in ARS Online are approximate and recommends consulting geological maps for better fault trace interpretations.

- For sites located near Palos Verdes fault (e.g. Bridge 981), ARS Online uses the old Caltrans model with a dip angle of 60 degrees. This model has been revised and the dip angle according to Caltrans and USGS is 90 degrees. We have used the revised model, therefore spectral accelerations calculated in this study are different than values estimated by ARS online. We believe the revised model should be used, therefore the values estimated here are more appropriate.

- Differences observed for estimated parameter $Z_{2.5}$ (depth at which shear wave velocity is 2500 m/s) between values used in this study and ARS Online values. Although both values are based on the same model, different procedures were used to obtain them. Difference was more significant where $Z_{2.5}$ was smaller than 3.0 km. Since at this depth basin factor is 1.0, the impact on spectral accelerations is negligible.

REDARS Analysis Results

REDARS analyses were not completed at the time of preparation of this report. Preliminary results, however, were available and an excerpt of the results is included in this report.

Figure A-8 shows the spectral accelerations at 1.0 second for all bridges in the network, due to the same rupture scenario on the Newport-Inglewood Fault. These values include the local site effects, based on NEHRP generalized soil profile types (Table A-3). The spectral acceleration value at 1.0 second was used to evaluate the damage state and economic loss due to damage to each bridge in the analysis.

Figure A-9 shows the estimated damage state for each bridge in the network, due to ground motions induced by rupture scenario on the San Andreas Fault. A damage state of 1 indicates no damage, while a damage state of 5 indicates complete failure of the structure. This index was used to evaluate the economic loss due to damage to the bridges resulting from considered earthquake scenarios. Table A-5 shows a summary of estimated damage state for bridges in the network.

Conclusion

Deterministic spectral accelerations were estimated using a deterministic approach for bridge sited in Los Angeles metropolitan area. Latest fault models and new generation of GMPEs were used in this study. Based on the results and discussions, we believe the spectral accelerations estimated in this study are appropriate for the intended application, which is risk assessment for the transportation network in Los Angeles area. The results can be improved by using site specific soil data and smaller scale geological maps, however due to the large number of the bridge sites and insignificant impact of these improvements on overall results, we believe these improvements are not necessary and the current study provides results which are reasonably accurate for the intended application.

Table A-1. Fault parameters used in deterministic seismic hazard analysis.

Fault Name	Mmax	Fault Type	Dip (degree)	Depth to Top of Rupture Plane (km)	Depth to Bottom of Rupture Plane (km)
Palos Verdes	7.3	RLSS	90	0	13
Newport-Inglewood Alt 1	7.5	RLSS	90	0	13
Newport-Inglewood Alt 2	7.5	RLSS	90	0	13
South San Andreas SM+NSB+SSB+BG+CO	7.8	RLSS	90	0	12

Table A-2. Fault Coordinates Based on USGS Model.

Fault Name	Coordinates	
Palos Verdes (USGS Stitched Model)	33.9702	-118.5570
	33.9021	-118.4960
	33.8637	-118.4390
	33.8175	-118.4000
	33.7895	-118.3340
	33.7472	-118.2540
	33.6919	-118.2330
	33.5835	-118.1460
	33.5449	-118.1190
	33.4816	-118.0800
	33.4428	-118.0600
	33.3546	-117.9860
	33.3184	-117.9430
	33.2798	-117.9170
	33.2066	-117.8380
	33.0732	-117.7420
	33.0269	-117.6870
	32.9520	-117.6160
	32.9251	-117.5740
	32.8655	-117.5200
	32.8085	-117.5070
	32.7524	-117.4670
	32.7235	-117.4330
	32.6786	-117.4130
	32.6238	-117.3440
	32.5940	-117.3150
	32.5519	-117.2980
	32.4949	-117.2660
	31.8900	-116.8400

Table A-2. Continued

Fault Name	Coordinates	
Newport-Inglewood Alt. 1 (USGS Stitched Model)	32.5603	-117.1473
	32.6033	-117.1505
	32.6478	-117.1654
	32.7099	-117.1621
	32.7290	-117.1704
	32.7595	-117.1976
	32.8011	-117.2100
	32.8355	-117.2413
	32.8545	-117.2636
	33.0189	-117.3252
	33.0858	-117.3769
	33.0971	-117.3961
	33.1083	-117.4110
	33.1225	-117.4234
	33.1559	-117.4291
	33.2163	-117.4870
	33.2515	-117.5474
	33.4024	-117.6882
	33.5080	-117.7989
	33.5910	-117.9146
	33.6127	-117.9340
	33.6745	-117.9930
	33.7045	-118.0436
	33.7179	-118.0630
	33.7355	-118.0757
	33.7649	-118.1138
	33.7887	-118.1504
	33.8267	-118.2057
	33.8321	-118.2128
	33.8438	-118.2322

Table A-2. Continued

Fault Name	Coordinates	
Newport-Inglewood Alt. 1 (USGS Stitched Model) (Continued)	33.8848	-118.2643
	33.9115	-118.2881
	33.9306	-118.3038
	33.9333	-118.3182
	33.9479	-118.3288
	33.9616	-118.3531
	33.9888	-118.3603
	34.0024	-118.3672
	34.0433	-118.3896
Newport-Inglewood Alt. 2 (USGS Stitched Model)	32.5603	-117.1473
	32.6033	-116.1473
	32.6478	-117.1654
	32.7099	-117.1621
	32.7290	-117.1704
	32.7595	-117.1976
	32.8011	-117.2100
	32.8355	-117.2413
	32.8545	-117.2636
	33.0189	-117.3252
	33.0858	-117.3769
	33.0971	-117.3961
	33.1083	-117.4110
	33.1225	-117.4234
	33.1559	-117.4291
	33.2163	-117.4870
	33.2515	-117.5474
	33.4024	-117.6882
	33.5080	-117.7989
	33.5910	-117.9146
	33.6060	-117.9247

Table A-2. Continued

Fault Name	Coordinates	
Newport-Inglewood Alt. 2 (USGS Stitched Model) (Continued)	33.6780	-117.9949
	33.6954	-118.0326
	33.7512	-118.0927
	33.8204	-118.1951
	33.8503	-118.2157
	33.8715	-118.2479
	33.9132	-118.2811
	33.9566	-118.3315
	34.0433	-118.3896
South San Andreas CH+CC+BB+NM+SM+NSB+ SSB+BG+CO (USGS Model)	33.3501	-115.7119
	33.7883	-116.2463
	33.8485	-116.3830
	33.8481	-116.4265
	33.8847	-116.5169
	33.9070	-116.5849
	33.9176	-116.6239
	33.9442	-116.6858
	33.9374	-116.7786
	33.9532	-116.8014
	33.9591	-116.8198
	34.0114	-116.8735
	34.0338	-116.9024
	34.0738	-117.0139
	34.0928	-117.0677
	34.1500	-117.2220
	34.1731	-117.2742
	34.2328	-117.3887
	34.2709	-117.4510
	34.3163	-117.5490
	34.4029	-117.7536

Table A-2. Continued

Fault Name	Coordinates	
South San Andreas CH+CC+BB+NM+SM+NSB+ SSB+BG+CO (USGS Model) (Continued)	34.5479	-118.1039
	34.6985	-118.5090
	34.6985	-118.5090
	34.7733	-118.7673
	34.8072	-118.8876
	34.8076	-118.8901
	34.8290	-119.0301
	34.8639	-119.2100
	34.9157	-119.3629
	34.9441	-119.4029
	34.9878	-119.4711
	35.0475	-119.5583
	35.1607	-119.7068
	35.3142	-119.8660
	35.4139	-119.9703
	35.5333	-120.0868
	35.7520	-120.3001

Table A-3. Correlation between soil profile type and V_{s30}.

Soil Profile Type	V_{s30} (m/s)	Number of Bridges
A	2,250	0
AB	1,500	0
B	1,100	737
BC	760	539
C	560	1,155
CD	360	2,830
D	270	5,820
DE	180	114
E	135	65
EF	90	0
F	70	0

Table A-4. Comparison between results of this study and Caltrans ARS Online.

GID / Bridge No.	Location (Lat./Lon.)	Fault[(1)]	Parameters used in this study[(2)]	S_A from this study (g) PGA 0.3 sec 1.0 sec	Parameters from ARS Online	Deterministic S_A from ARS Online (g) PGA 0.3 sec 1.0 sec
6 53 0145	33.78979 -118.22400	NI	R_{RUP} = 4.30 km R_x = 4.30 km R_{jb} = 4.30 km V_s = 560 m/s $Z_{1.0}$ = 430 m $Z_{2.5}$ = 4.95 km	0.488 1.097 0.659	R_{RUP} = 4.41 km R_x = 4.41 km R_{jb} = 4.41 km $Z_{1.0}$ = 480 m $Z_{2.5}$ = 2.00 km	0.445 0.969 0.556
14 53 0969	33.79043 -118.28125	PV	R_{RUP} = 2.74 km R_x = 2.74 km R_{jb} = 2.74 km V_s = 560 m/s $Z_{1.0}$ = 410 m $Z_{2.5}$ = 2.47 km	0.488 1.061 0.586	R_{RUP} = 2.15 km R_x = 2.15 km R_{jb} = 2.15 km $Z_{1.0}$ = 415 m $Z_{2.5}$ = 2.00 km	0.517 1.128 0.634
796 53 1860	34.72849 -118.16973	SSA	R_{RUP} = 15.92 km R_x = 15.92 km R_{jb} = 15.92 km V_s = 560 m/s $Z_{1.0}$ = N/A $Z_{2.5}$ = N/A	0.252 0.541 0.299	R_{RUP} = 15.66 km R_x = 15.66 km R_{jb} = 15.66 km $Z_{1.0}$ = 74 m $Z_{2.5}$ = 2.00 km	0.255 0.547 0.304
981 53 2618	33.76242 -118.23830	PV	R_{RUP} = 2.23 km R_x = 2.23 km R_{jb} = 2.23 km V_s = 270 m/s $Z_{1.0}$ = 420 m $Z_{2.5}$ = 2.88 km	0.463 0.923 0.860	R_{RUP} = 4.01 km R_x = 4.01 km R_{jb} = 3.64 km $Z_{1.0}$ = 450 m $Z_{2.5}$ = 2.00 km	0.422 0.854 0.753
981 53 2618	33.76242 -118.23830	NI	R_{RUP} = 7.54 km R_x = 7.54 km R_{jb} = 7.54 km V_s = 270 m/s $Z_{1.0}$ = 420 m $Z_{2.5}$ = 2.88 km	0.356 0.742 0.617	R_{RUP} = 6.21 km R_x = 6.21 km R_{jb} = 6.21 km $Z_{1.0}$ = 450 m $Z_{2.5}$ = 2.00 km	0.381 0.787 0.682
2471 53 1142	33.85753 -118.28515	NI	R_{RUP} = 3.32 km R_x = 3.32 km R_{jb} = 3.32 km V_s = 270 m/s $Z_{1.0}$ = 510 m $Z_{2.5}$ = 2.93 km	0.444 0.895 0.877	R_{RUP} = 3.70 km R_x = 3.70 km R_{jb} = 3.70 km $Z_{1.0}$ = 510 m $Z_{2.5}$ = 2.00 km	0.435 0.880 0.851

Table A-4. Continued

GID / Bridge No.	Location (Lat./Lon.)	Fault[1]	Parameters used in this study[2]	S_A from this study (g) PGA 0.3 sec 1.0 sec	Parameters from ARS Online	Deterministic S_A from ARS Online (g) PGA 0.3 sec 1.0 sec
3523	33.71032 -115.38103	SSA	R_{RUP} = 50.41 km R_x = 50.02 km R_{jb} = 50.41 km V_s = 360 m/s $Z_{1.0}$ = N/A $Z_{2.5}$ = N/A	0.118 0.262 0.150	R_{RUP} = 50.16 km R_x = 50.16 km R_{jb} = 50.13 km $Z_{1.0}$ = 264 m $Z_{2.5}$ = 2.0 km	0.118 0.263 0.151

Notes:

1) PV: Palos Verdes, NI: Newport-Inglewood, SSA: South San Andreas
2) Basin depths: If $Z_{1.0}$ < 300 m or $Z_{2.5}$ < 3.0 km basin factor is 1.0, N/A means estimated basin depth is shallow and basin factor of 1.0 is used in analysis.

Table A-5. Damage state estimation for bridges in the network.

DS	Newport-Inglewood	Palos Verdes	South San Andreas
1	61,903	63,256	63,080
2	691	117	176
3	476	107	122
4	350	48	115
5	111	3	38
SUM	63,531	63,531	63,531

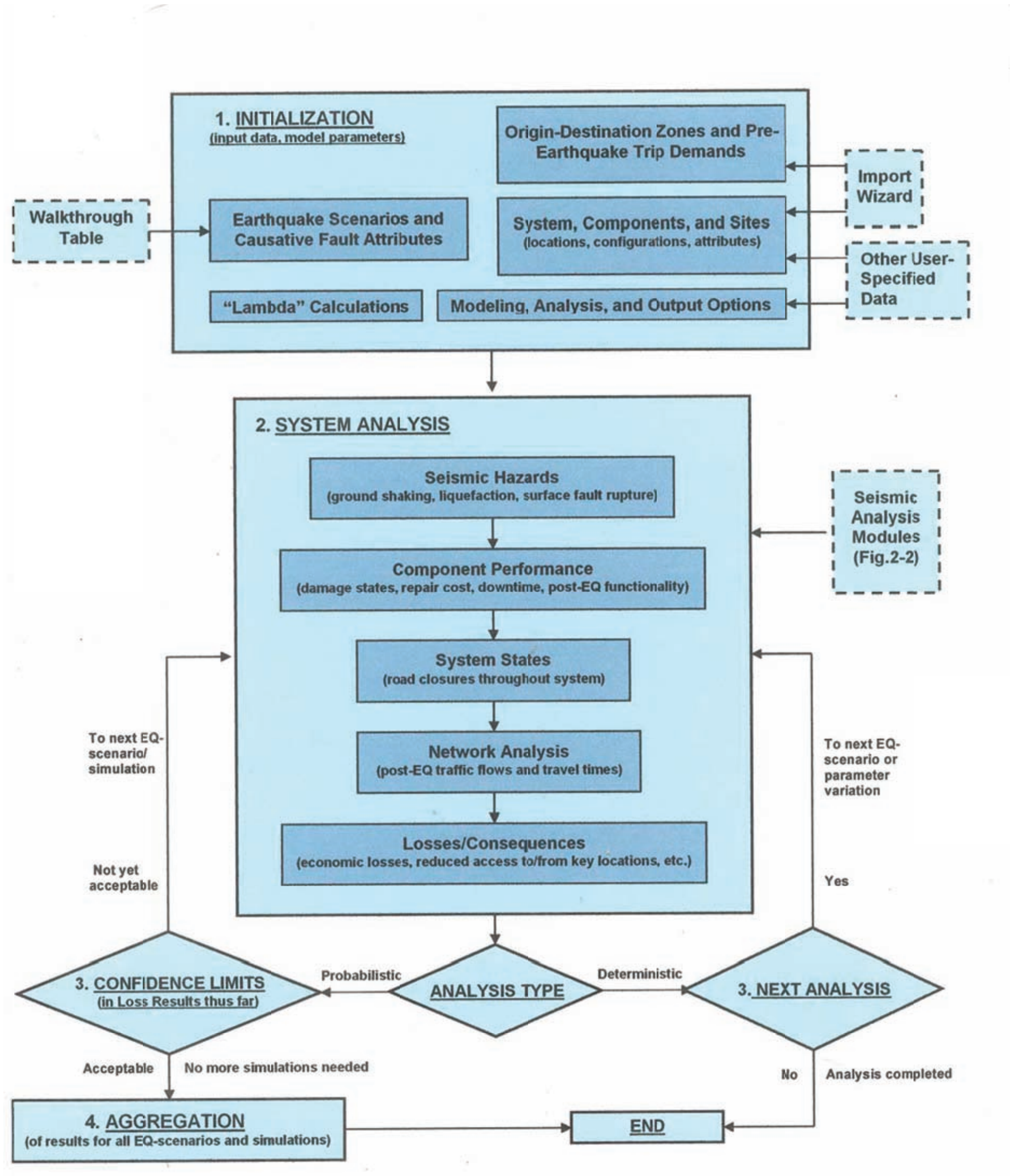


Figure A-1.REDARS 2 methodology for seismic risk analysis of highway systems. (Werner et al., 2006)

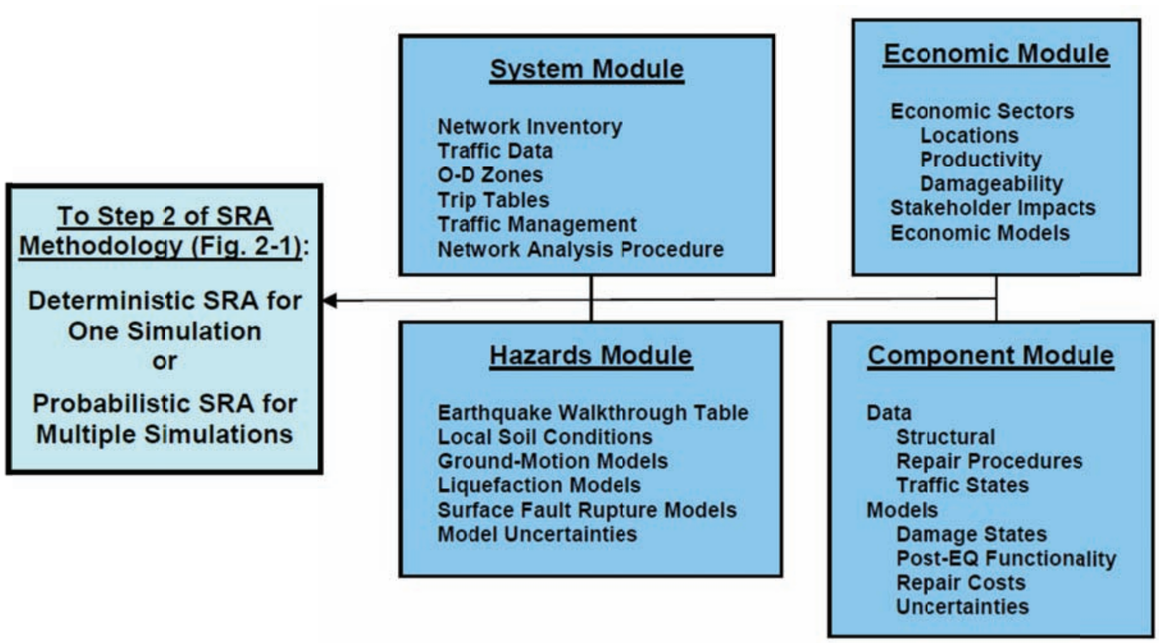


Figure A-2. REDARS 2 Seismic Analysis Module. (Werner et al., 2006)

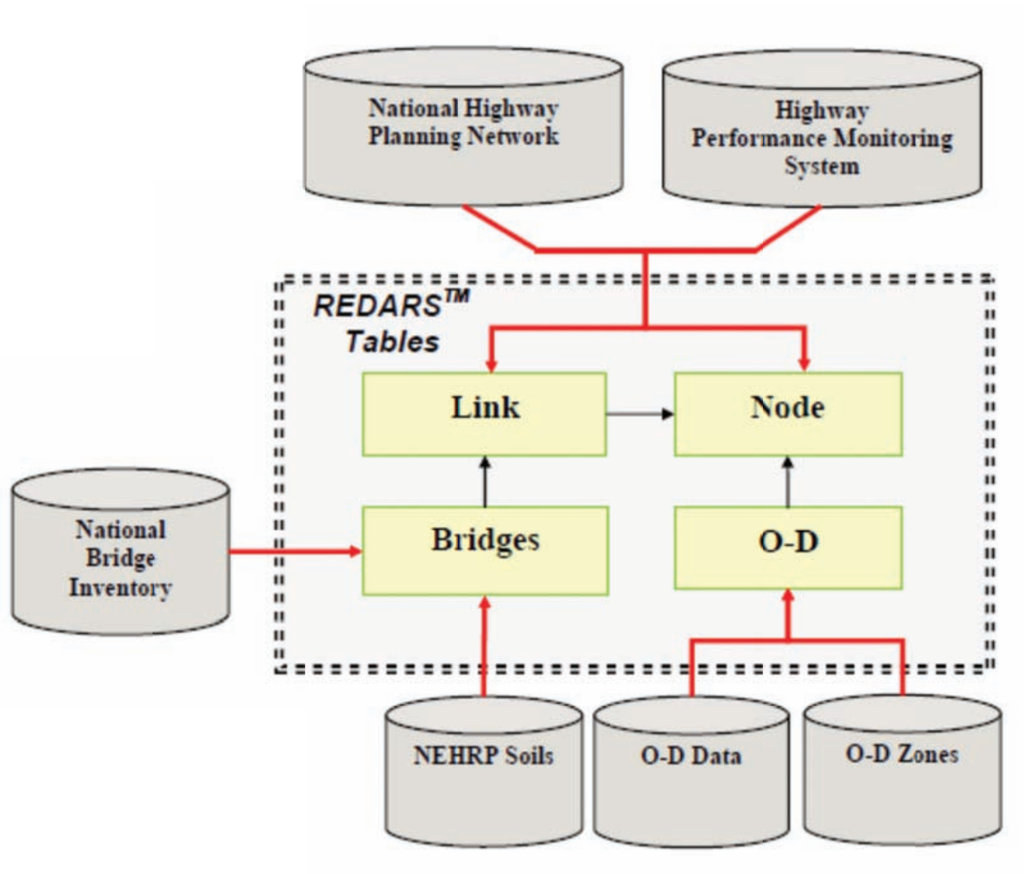


Figure A-3. REDARS 2 analysis initialization using external databases to provide input parameters.

(Werner et al., 2006)

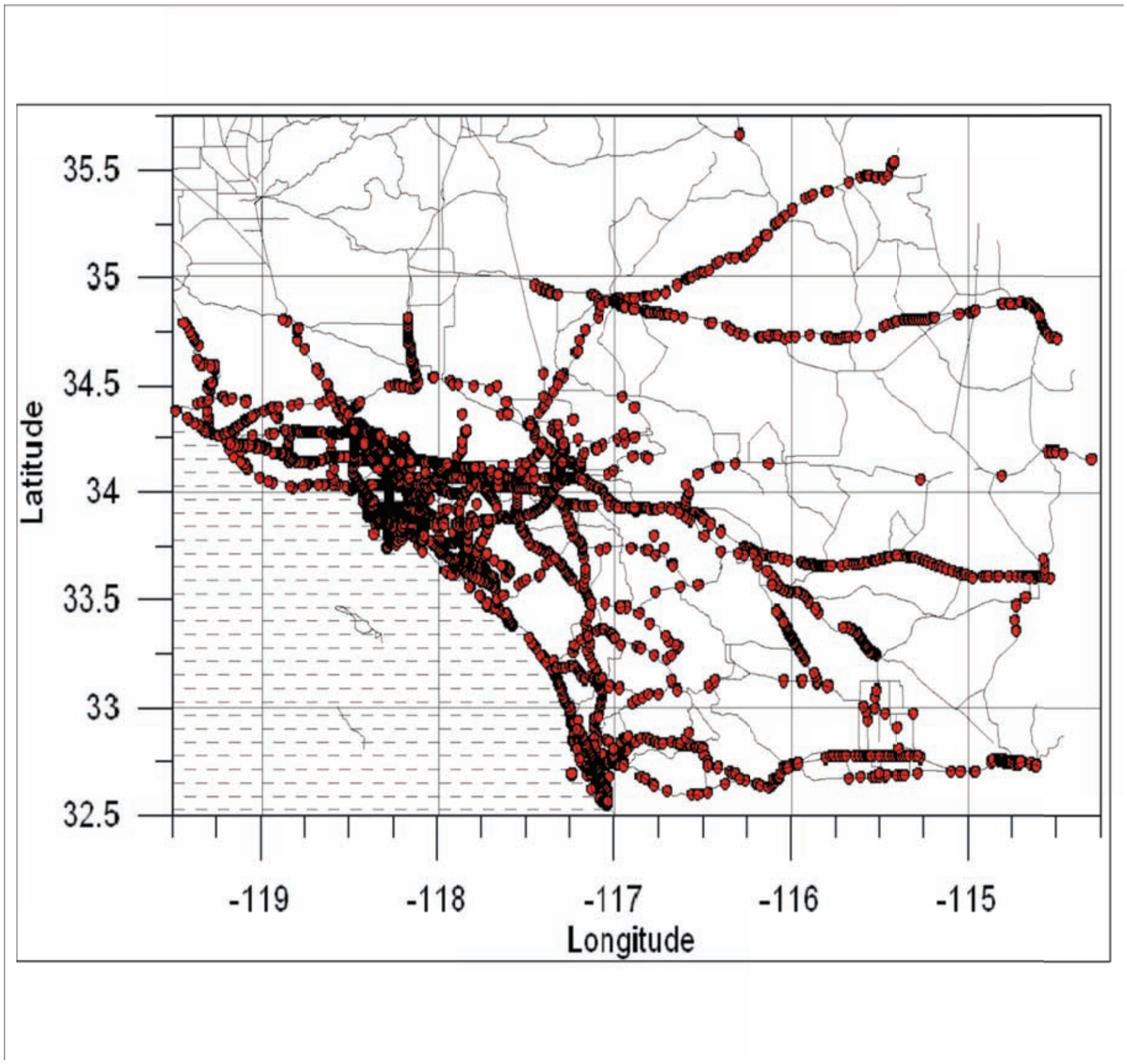


Figure A-4. Location of the Bridges Included in the Study.

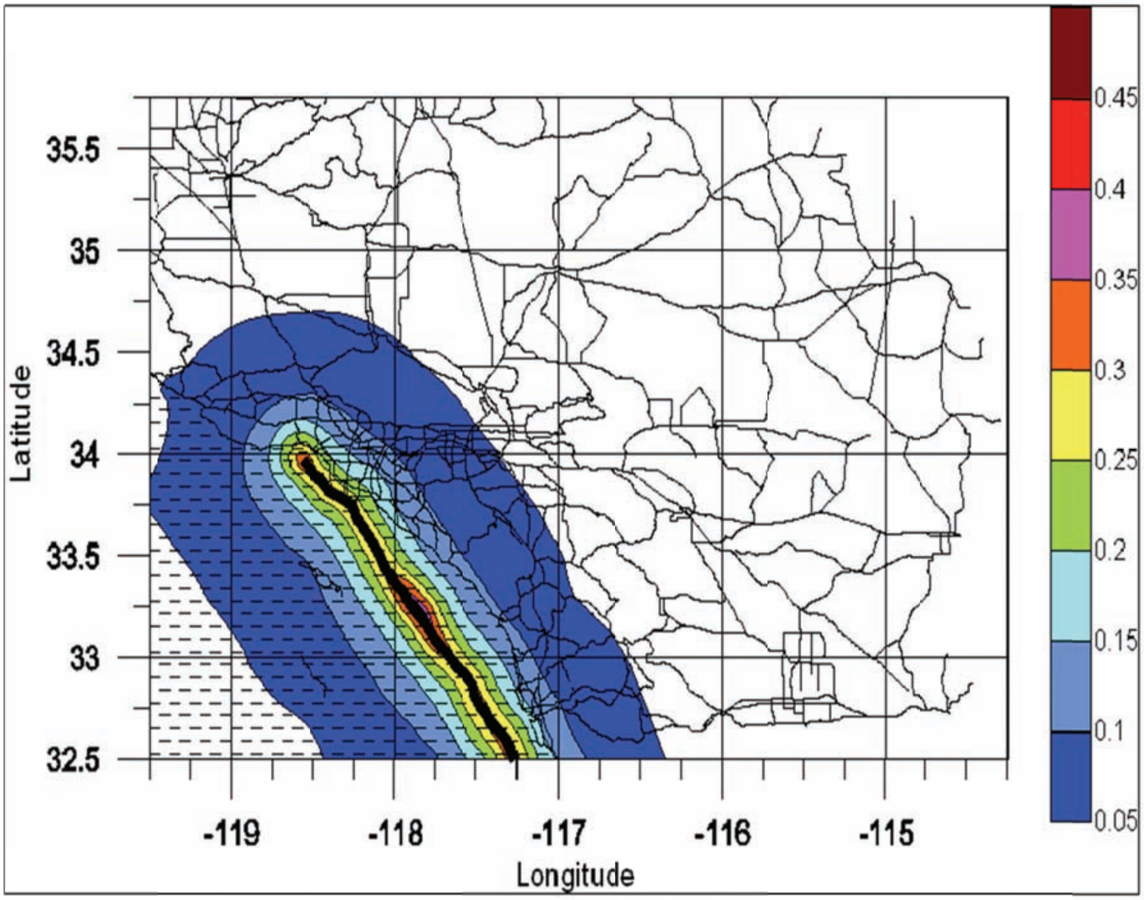


Figure A-5. Palos Verdes Fault PGA Values for Soil Profile Type B/C (Vs=760 m/sec).

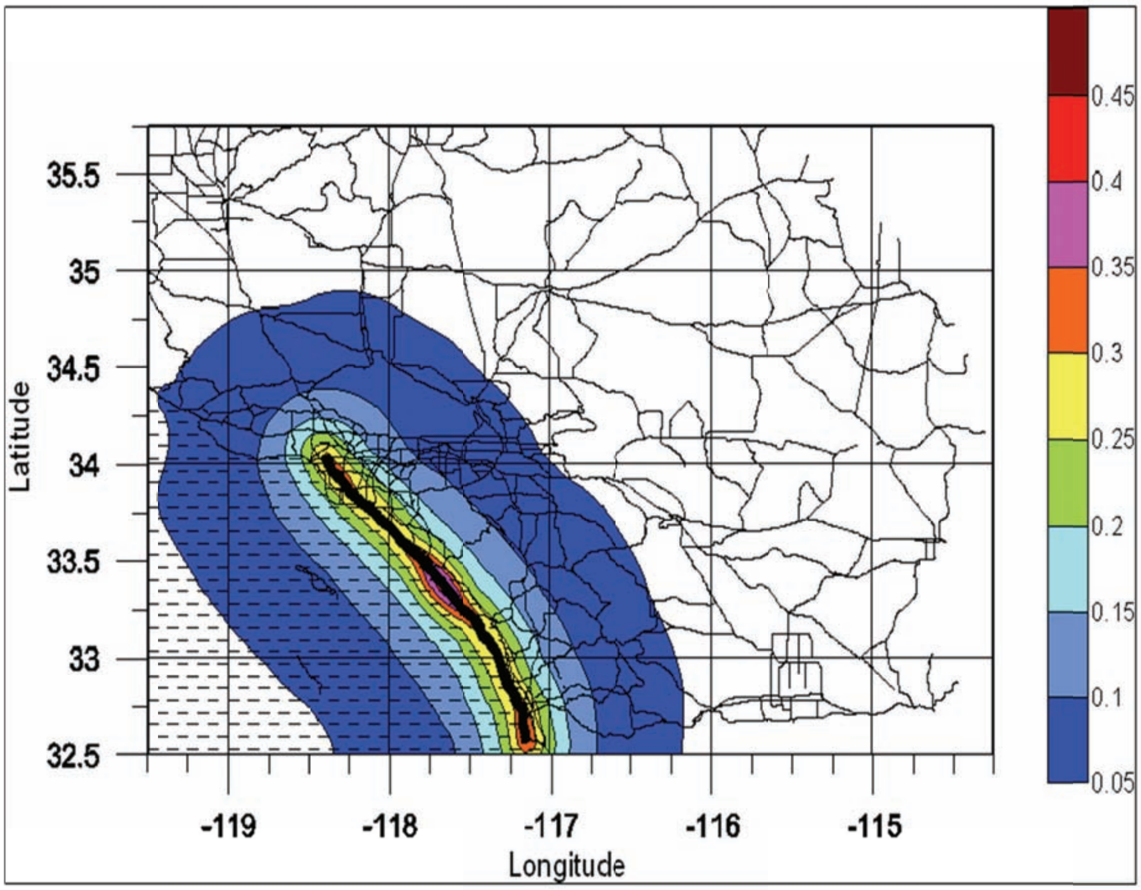


Figure A-6. Newport-Inglewood Fault PGA Values for Soil Profile Type B/C (Vs=760 m/s).

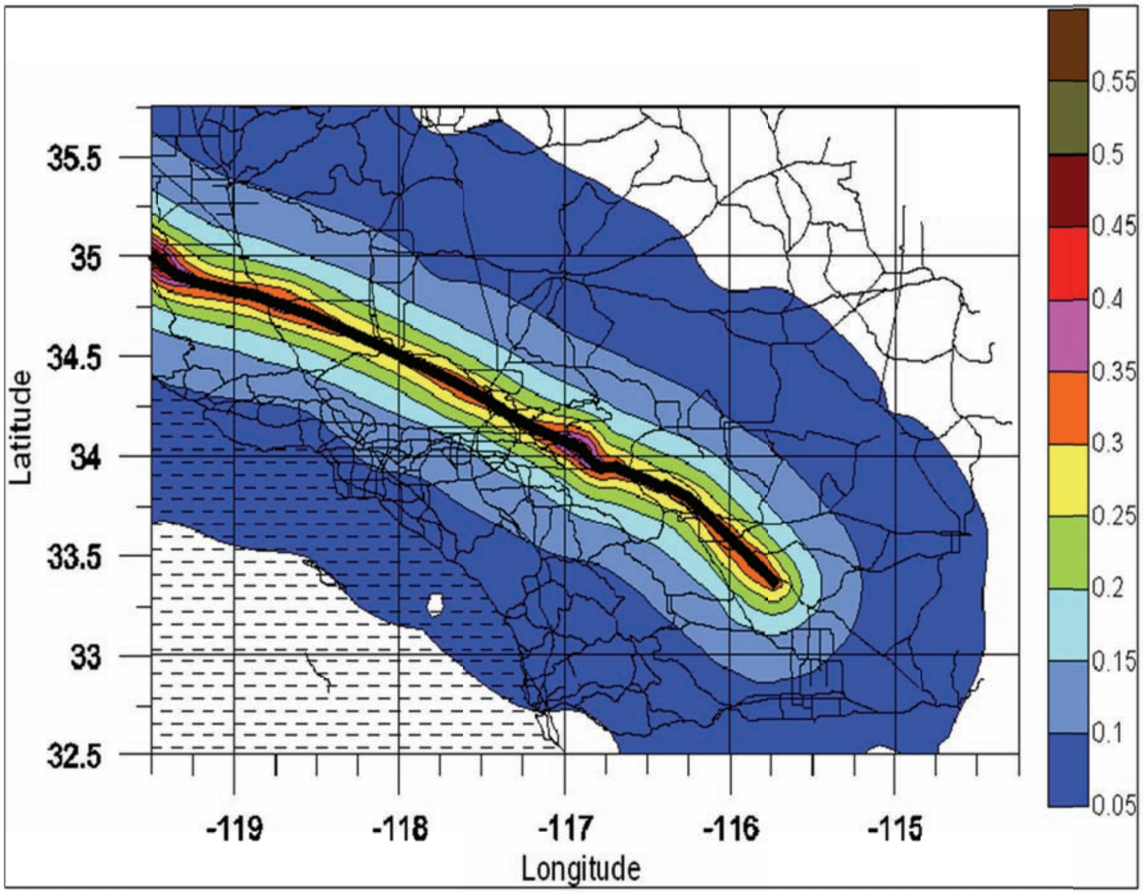


Figure A-7. South San Andreas Fault PGA Values for Soil Profile Type B/C (Vs=760 m/sec).

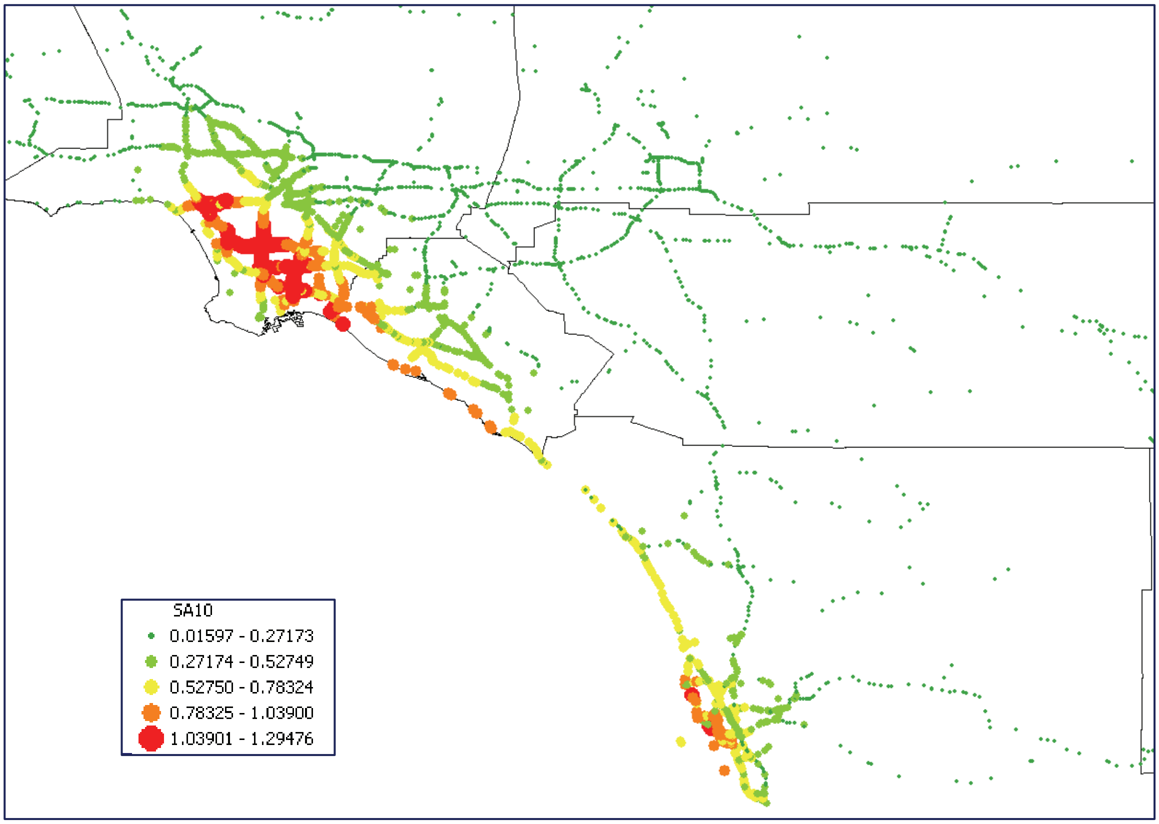


Figure A-8. The figure shows the spectral acceleration values at 1.0 second for the Newport-Inglewood Fault rupture scenario for the studied bridges. This spectral acceleration was used to assess the damage to the bridges due to earthquake. (Cho, 2010)

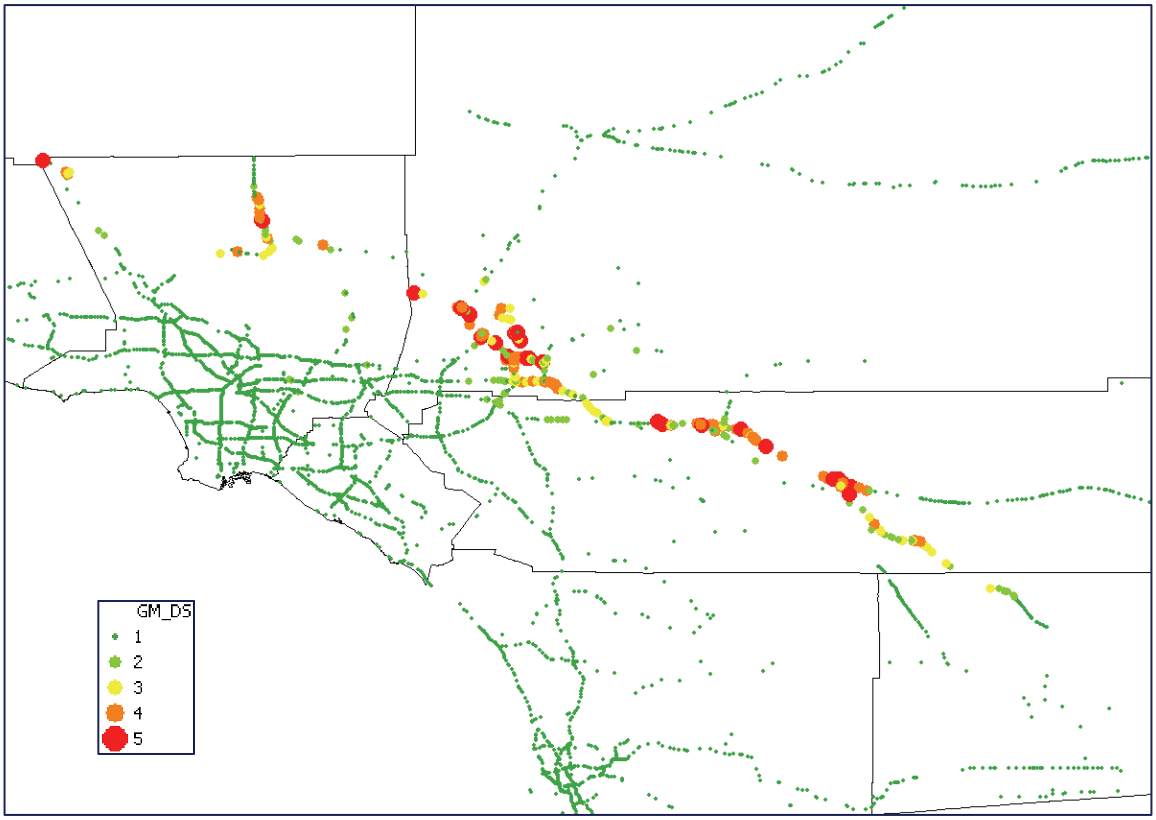


Figure A-9. The figure shows the Damage State for each bridge due to ground motion from the San Andreas Fault rupture scenario. A damage state of 1 means no damage, while 5 means complete failure. This parameter was used to assess the economic loss due to damage to each bridge structure. (Cho, 2010)

Appendix References

California Department of Transportation (Caltrans) (2006) “Seismic Design Criteria (SDC), Version 1.4.”

California Department of Transportation (Caltrans) (2009a) “Caltrans ARS Online,” http://dap3.dot.ca.gov/Shake_stable/index.php/

California Department of Transportation (Caltrans) (2009b) “Seismic Design Criteria (SDC),” Revision Date 8/12/09.

Campbell, K. and Bozorgnia, Y. (2008) “NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PG and PGD, and 5% Damped Linear Elastic Response Spectra for Periods Ranging from 0.01 to 10 s,” Earthquake Spectra, Vol. 24, pp. 139-172.

Chiou, B. and Youngs, R. (2008) “An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra,” Earthquake Spectra, Vol. 24, pp. 139-172.

Cho, S. (2010), Internal Web Presentation for METRANS/Megacities Project.

Shantz, T. and Merriam, M. (2009), “Development of the Caltrans Deterministic PGA Map and Caltrans ARS Online,” July.

Southern California Basin Models (Community Velocity Model V.4) (2009), http://www.data.scec.org/3Dvelocity/

Werner, S. D., Taylor, C. E., Cho, S., Lavoie, J.P., Huyck, C., Eitzel, C., Chung, H., and Eguchi, R. T., 2006, REDARS 2 Methodology and Software for Seismic Risk Analysis of Highway Systems., Special Report MCEER-06-SP08.