DISASTER DAMAGE ESTIMATION MODELS: DATA NEEDS VS. GROUND REALITY

by

SUDHA MAHESHWARI

A dissertation submitted to the Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Planning and Public Policy

Written under the direction of

Richard K. Brail

And approved by

____

____

____

____

New Brunswick, New Jersey

October, 2007

ABSTRACT OF THE DISSERTATION

Disaster Damage Assessment Models: Data Needs vs. Ground Reality

By SUDHA MAHESHWARI

Dissertation Director:

Richard K. Brail

Integrated assessment models are being used extensively in the field of disaster damage estimation and assessment. However, there is a great deal of uncertainty involved with the use of these models – not only because of the uncertainty of predicting the occurrence of hazards but also because of the quality of data that are input into these models. The use of these models for real-world decision-making is limited by the data. Poor quality data can lead to poor decisions, particularly at a local level of analysis. This dissertation looks at the issue of model-data interaction and the uncertainty inherent due to the lack of good quality data. The above interaction is researched using the HAZUS™ model (a state-of-the-art damage estimation model) and focusing on building inventory data for two cities: City of Seattle, WA and City of Long Beach, CA. It assesses how the local level building inventory data compares with default building inventory data in HAZUS™ for the two cities above. Finally it looks at how changes in the building inventory data lead to changes in the damage estimation from HAZUS™. In order to understand patterns of variation, both of the above are analyzed at the full city level and at the level of census tracts comprising the cities. The dissertation finds that although a lot of basic

GIS data exist for large cities at the local level, the building inventory data are severely lacking in some required information, accuracy and completeness. Where good data exist, the results show that there is a large variation in building inventory in the default data which leads to an even larger variation in damage estimation. All occupancy classes excepting residential are significantly underestimated and much of the underestimation is concentrated in the commercial, industrial, education and institutional classes. There is even large variation for downtown census tracts and single use census tracts such as ones with universities, etc. Where good data do not exist (as in the case of City of Long Beach), the use of local data is difficult and requires significant expertise and assumptions. In such cases, the use of HAZUS™ should be with a great deal of caution.

Acknowledgements

A large number of people contributed to this project. First, I would like to thank the members of my dissertation committee who provided support and feedback: Richard Brail, Clinton Andrews, Robert Burchell, and James K. Mitchell. My work with Lyna Wiggins led to the conception of this dissertation and I am grateful for her support and guidance throughout my graduate program. The HAZUS™ development team were helpful at various times – particularly Claire Drury, Jawhar Bouabid, and Pushpendra Johari were helpful in providing copies of HAZUS™ (often before pre-release), explaining abnormal behavior with the software, fixing bugs, and processing data outside the software interface. I thank all those GIS professionals at the 19 cities for spending time and discussing their GIS operations with me, in particular the GIS Coordinators at the City of Seattle and Long Beach for allowing me to use their data for the case studies. My friends and co-workers provided sufficient prodding to keep me moving and encouraged. My father always encouraged me to pursue higher education and I am happy to fulfill what was an unfulfilled dream for him, as my grandmother strived for a Ph.D. but was unable to get it due to family hardships. My mother provided critical help and support while I wrote large parts of this document. Other members of my family also supported me with encouragement and inspiration. I am so happy that Arjun and Arushi came into my life at the right time to focus me and provide maternity leaves which were used for writing large chunks of this dissertation. Finally, I want to thank Gavin, who just never gave up on me and didn't let me give up either.

Table of Contents

List of Tables

List of Figures

List of Maps

Map

Chapter 1: Introduction

In recent years, there have been significant advancements in the development of integrated models for hazard assessment and damage estimation. Models such as HAZUS™ (FEMA 2003), CATS™ (Swiatek and Kaul 1999), and TAOS (Watson and Johnson 1999), have attempted to provide decision-makers with tools to assess the impacts of hazardous events on the built environment and human systems. Particularly, they help decision-makers create scenarios of disasters, analyze their impacts, model the costs and benefits of various policy decisions and make rational choices amongst competing alternatives. However, the use of these models can be limited not only by the quality of the tools but also by the quality of data that is input into the models. The quality of data is crucial for local-level decision-making at the scale of a city, county or other local government. The topics of investigation in this research includes understanding the availability of data at the local level for disaster damage assessments, analyzing the variation in local-level data versus estimated data, and the changes in results from the HAZUS™ model.

The cost of natural disasters is constantly increasing for both the developed and developing world. During the last decade (1990-99), there were 460 Presidential declarations of disasters in the United States, double the number of declarations for the previous decade (1980-89). In the last decade, the Federal Emergency Management Agency (FEMA) spent more than $25.4 billion for disasters in comparison with $3.9 billion (current dollars) in the previous decade (FEMA 2005). While the increased costs of natural disasters in developed countries is attributable to higher property losses, the cost of disasters in the developing world is attributable to higher social costs such as loss of life and livelihood, and the complete devastation to the economy from a single event. While it is debatable whether the number of disasters have increased, there is no doubt that there has been an increase in the cost of disasters. The higher cost of disasters is largely because of the increased occurrence of disasters in densely populated urban regions, which is a consequence of urban growth. It has been estimated that there are now 300 cities with more than a million population and 50% of the world's population lives in cities. Cities are exposed to greater vulnerabilities from disasters due to the concentration of people and expensive investment which is particularly vulnerable to damage from natural hazards.

It is important to note that natural disasters are no longer considered to be solely caused by "natural" agents. A complicated intermingling of cause and effect results in "complex" disasters even though the initial trigger may be a natural agent such as an earthquake, flood or hurricane. For example, an earthquake can result in a fire or a flood can cause an oil spill. The complexity of disasters, along with the inter-dependencies of various human-made systems, particularly in urban areas, makes it difficult to understand, manage and mitigate the impacts of such disasters. Decision-makers and urban managers are faced with the challenging task of saving life and property without a proper understanding of the complexity of the natural hazard (i.e. the agent causing the disaster) and its interplay with urban systems which are also intrinsically connected.

The assessment of damage from a disaster over an urban fabric requires the modeling of the built environment. Integrated assessment models can be very useful in this regard. Integrated assessment models are tools that formalize assumptions and bridge different disciplinary domains through synthesis of a broad range of expertise. Integrated models for disaster management can help estimate, analyze and visualize the impacts of various disaster scenarios. They help in understanding the impacts of various policy options and mitigation strategies. They have the potential to assist decision-makers both before and after a disaster. However, the process of integrated modeling by itself is a complex one and modeling the complexity of a disaster situation presents many challenges.

One of the most important challenges in the use of integrated models is the amount and type of data needed to run them. Integrated models for damage assessment are complex and require the input of large amounts of data on building inventory, infrastructure, socio-economic and demographic composition and other environmental variables. The most commonly used integrated hazard assessment tools use geographic information systems (GIS) as an integrative framework with the input and output from models being GIS datasets. The large amounts of GIS data required by most models involve collecting data at various scales, resolution and levels of aggregation. Not only are data not easily available at a local level, very often the available data do not carry adequate information to undertake damage assessments that are useful for public policy purposes at the local level.

An important aspect of the built environment are the building structures and their characteristics such as use, type of construction, area, height, age and value. Data about these characteristics are often available locally and are in GIS format. However, because it is unclear how widely used GIS are in local governments and whether such data are easily available at the local-level, integrated models make assumptions about buildings and their characteristics. These assumptions are based on national-level estimates so that these models can be run without much local-level data. However, studies have shown significant deviation of these estimates from data available at the local level (Nordenson, et al 1999).

The lack of local data can have a large impact on the sustained use of these models, particularly for decision-making at the local level. The local level perspective is important because disaster mitigation policies and first response are undertaken at this level, even though the disasters may not follow any political boundary. There are several implications for the use of non-local data – it can lead to greater uncertainty from the results of the model, may result in flawed mitigation policies, and lead to the poor implementation of policies and priorities. This may eventually cause decision-makers to turn away from these models. Therefore, to reduce uncertainty from these models, the use of local-level data is very important. This research seeks to understand the availability of local-level data for disaster management (particularly damage assessment) for large cities in the United States. In the light of existing data realities, it seeks to understand the challenges associated with using local data. Furthermore, it focuses on the usefulness and sensitivity to local level data of HAZUS™, a state-of-the-art damage assessment model.

The HAZUS™ model was developed in the late 1990s in the United States to undertake disaster damage estimation due to earthquakes. The model was developed and funded by FEMA and National Institute for Building Sciences (NIBS). Subsequently other modules have been added to the model to include floods and wind-based hazards (primarily hurricanes). Like other integrated models, the HAZUS™ model also uses GIS as an integrative framework. It is now referred as HAZUS-MH and stands for Hazards US, Multi Hazards. Like most other complex models, this model is also very data intensive and involves many datasets. When the datasets are integrated with a modeled hazard scenario, the model provides an assessment of damage and helps identify and implement mitigation strategies, as shown in Figure 1.1.

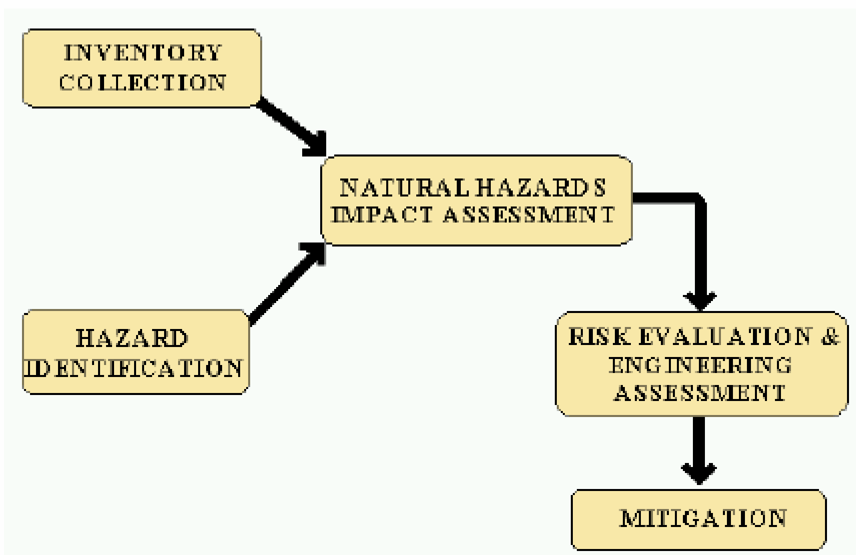


Figure 1.1: HAZUS™: Steps to Assessing Damage
Source: FEMA 2003

Thus, HAZUS™ integrates information on the hazard (such as ground motion and ground failure in the case of earthquakes) with information on general building stock, essential and high potential loss facilities, and other systems such as transportation and utilities, to assess direct and indirect damage as shown in Figure 1.2. Other data such as socio-economic and demographic data are also used to assess fatalities, injuries, and other direct and indirect losses. Most data are aggregated to the census tract and results from the model are also presented at this level of aggregation. The model comes bundled with all the required data, so that even an inexperienced user can easily run a disaster scenario and assess the damage from that scenario. This includes data about the building stock which is derived from census data (for residential buildings) and Dun and Bradstreet data (for commercial, industrial and other buildings). As mentioned before, these datasets can be significantly different from reality, particularly in large cities.

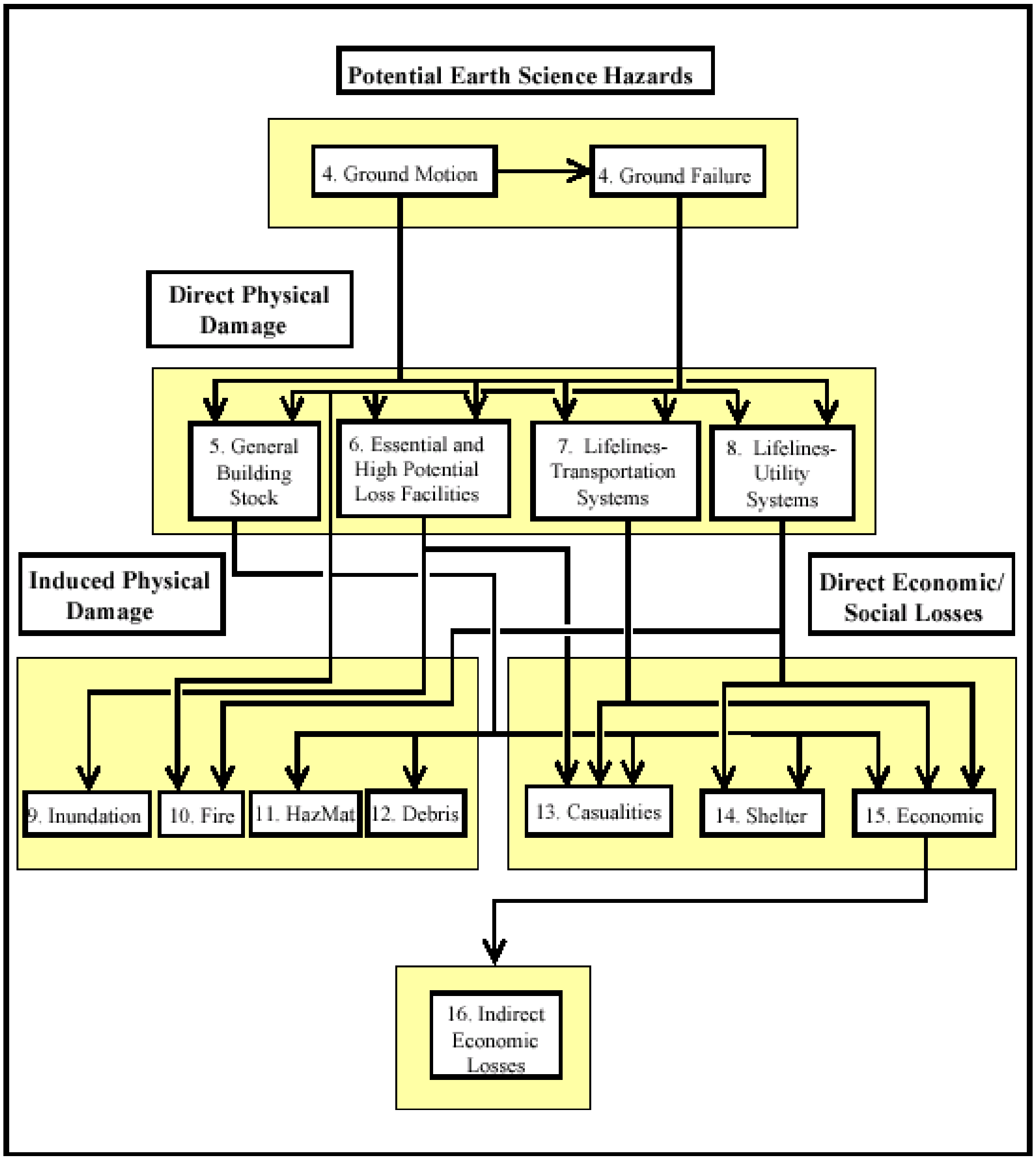


Figure 1.2: HAZUS™ Methodology
Source: FEMA 2003

Therefore, the resultant damage from HAZUS™ can be very different depending on whether local level data or HAZUS™ default data are used. The model is flexible in terms of improving data from local sources and also provides various tools to collect and input local data. The HAZUS™ has a tool called the Building Inventory Tool (BIT) which helps to input local building inventory data into the model. Information needed include the use, type of structure, area, height, age, census tract, and value of the buildings. Likewise, more detailed and accurate hazard information can be input into the model. A detailed discussion of the HAZUS™ model is provided in Chapter 2.

Many sources of local data exist, such as aerial photos, cadastral (parcel) maps along with tax assessor's files, and public works data. However, such data are often not easily available for most cities, are not updated, or are in a format that is difficult to convert into the required format of the model. Furthermore, the data manipulation requires extensive GIS knowledge, as well as knowledge about architecture, civil engineering and other fields. Hence it is not simple to run these types of models with more local data.

One important task of this research was to look broadly at the diffusion of GIS in local governments in large cities. The aim was to understand the state of local GIS (particularly in the context of disaster management) and the availability of local-level data for input into the HAZUS™ model. The diffusion of GIS was analyzed under three main parameters: the organizational structure of GIS implementation at the local-level; the availability of various datasets at the local-level; and the use of GIS for disaster management in local agencies. This was accomplished through a survey/phone interview of respondents from 19 randomly selected cities with populations between 250,000 and 1 million. The survey also provided a basis for selecting two case study cities where local-level data were available for input into HAZUS™. The case studies were undertaken to understand the challenges associated with using local data and to look at the sensitivity of the HAZUS™ model to local level data pertaining to building characteristics and earthquake hazards.

Based on the survey, the two cities identified for further analysis were the City of Seattle, WA and the City of Long Beach, CA. These two cities were chosen as case studies essentially because of their imminent threat from earthquakes and because of their willingness to share their local data. Both cities have been quite proactive in managing their threats from natural hazards and have used GIS data extensively for disaster management. While the implementation of GIS in Seattle was very advanced, Long Beach represented a more “typical” GIS implementation for cities of this size. The GIS organizational structure in Long Beach was much smaller and hence the datasets were not as well developed and documented as Seattle.

The tax assessor’s roll or database (or assessment data as it is commonly called) is one of the most widely available datasets at the local level that captures building characteristic information such as area, height (or number of stories), use, type of construction, assessed value, and market value. Hence the research focused largely on the assessment data and their use for disaster damage assessment. These data were analyzed in-depth for both cities. The two cases aimed to understand the quality and completeness of the assessment data, and the challenges associated with inputting these data into HAZUS™. This research also analyzed the variation of the local-level data vis-à-vis the default data already available in HAZUS™. This analysis was done at the level of the whole city and then at the disaggregate census tract level to understand any spatial patterns of variation. The local data were input into HAZUS™ and the same scenario was run for each city with local data and default data. Three scenarios were modeled for each city and comprised of earthquakes at the same epicenter and same fault system but of three different magnitudes – 5.0, 6.0 and 7.0 events. The results of damage were then analyzed for the three scenarios to understand the sensitivity of the model to local level data both at the level of the city and also at the level of the various census tracts that comprise the city.

The two selected case studies were analyzed by themselves and also compared with each other to uncover any similar patterns that might exist. Consistent patterns may be replicable in other cities or can inform decision-makers in other cities on the reliability of results from HAZUS™ when default data are used. Finally, recommendations are made on the appropriate use of HAZUS™ and on other aspects of local data availability and diffusion, and the challenges associated with using local data in HAZUS™.

A variety of events had an impact on the data analysis for this research. This research started in 2000-2001. At that time, HAZUS™ comprised only an earthquake module that ran on ArcView 3.2 and MapInfo technology. Interviews and surveys were conducted in 2001 to understand the availability of various datasets for input into HAZUS™. Seven interviews were conducted in the summer of 2001. Then came September 11, 2001. The project suffered a setback as the climate in the aftermath of 9/11 was not conducive to calling people and requesting information on GIS datasets. Many cities were removing GIS datasets previously available to the public from their Internet mapping websites. There was a serious debate in the country about the efficacy of public access to GIS datasets given that they could be used for planning acts of terrorism. After a short break, when the immediate reaction had calmed down, the survey was resumed and the rest of the cities were contacted. In 2002, the two case study cities were identified and data collection from these two cities began.

The earthquake module was the only one available in HAZUS™ at the time of the survey. Therefore, this played a large role in the selection of the case cities and also as earthquake was the only hazard that could be analyzed. As data were prepared and input into HAZUS™, this research uncovered many bugs in HAZUS™, particularly in the Building Inventory Tool (BIT). The input data were not being processed correctly, resulting in many census tracts within Long Beach having the exact same amount of square footage for some occupancy classes. The local data processed by the BIT tool in HAZUS™ were also not adding up to the input data. These bugs were reported to the HAZUS™ development team which acknowledged the issues but was not willing to resolve them since a new version of HAZUS™ was under development and would be released soon. Therefore, there was no other option but to wait.

In 2003, a new version of HAZUS (HAZUS-MH) was released. This new version of HAZUS™ incorporated the flood and hurricane module and was developed on the ArcGIS 8.x platform. Various bugs were uncovered through this research in the new version of HAZUS™ as well. Particularly the BIT tool was not usable until the HAZUS™ development team provided a new patch. Once the tool was made to work, and local data could be input, the results were validated to ascertain (as much as possible) that there were no discrepancies. The HAZUS-MH version also contained new data based on 2000 census tracts and updated Dun and Bradstreet data. There were also some changes in the data needs and the tool. In previous versions of HAZUS™, the residential use (or occupancy as HAZUS™ calls it), was divided into 6 categories – single family residential, manufactured housing, multi-family dwellings, temporary lodging, institutional homes, and nursing homes. In the newer HAZUS-MH, multi-family dwellings were further broken up into 6 subcategories – duplex, triplex/quadruplex, apartments (5-9 units), apartments (10-19 units), apartments (20-49 units) and apartments (50+ units). Furthermore, there were two additional fields of information that were required by HAZUS-MH which were previously automatically updated by HAZUS™ when local data were input into HAZUS™ : value of buildings (building value) and the contents in the buildings (content value).

The above changes in HAZUS™ had some impact on this research. First, the data for Seattle were already prepared for HAZUS™ before the breakdown of multi-family dwellings. Therefore, all multi-family dwellings were coded to apartments and it was not easy to break this category down further into the subcategories without redoing the entire data manipulation. The Long Beach data were not prepared before and were broken up into the subcategories. Particularly challenging was the issue of building exposure values. While the building values are available from assessment data, the content exposure information is not available from any local sources (including assessment data). In the absence of these data, HAZUS™ updates the square footage based on local data but keeps the old exposure values. Since the square footage of local data is vastly different from the default values in HAZUS™ (as will be discussed in Chapter 4 and 5), keeping default values of exposure can lead only to minimal changes in the damage assessment. Therefore, it is practically useless to input local data if building and content values cannot be updated to reflect local data or at least improvements in square footage. Since no local sources were available for updating content values, in this research both content and value information were updated using average per square feet exposure information from HAZUS™ based on default values. It was deemed unsuitable to update the building value alone from assessment data because that would mean that exposure values were calculated from two different sources.

Therefore, this research was impacted largely by these various externalities and some of them have contributed to limitations of this research. For example, a lot of time was spent analyzing the integrity of the tool. Besides prolonging the timeline for this research, it also turned out to be a lot more work than anticipated because data were analyzed multiple times and errors in the model were uncovered. Furthermore, newer versions of HAZUS™ are now available and some of the issues discussed in this dissertation may already have been addressed in the newer versions. However, it is important to note that this research has already begun to have its impact since the BIT tool has been improved significantly based on feedback from this research.

This research is presented in the following chapters: Chapter 2 provides a detailed discussion of the context and related literature along with the research framework that guided this research. A discussion of the research design is also provided. This chapter also includes a more detailed discussion of the HAZUS™ model. Chapter 3 discusses the findings from a survey of 19 cities, with respect to local level GIS diffusion and availability of data for damage assessment. Chapters 4 and 5, present findings from the case studies of Seattle and Long Beach respectively. Both case study discussions follow a similar structure – a discussion of the city, followed by a general discussion on the GIS data and organization. The challenges of inputting local data into HAZUS™ are also discussed in each chapter. The two case studies also look at the variation of local data from HAZUS™ default data and compare the results of damage assessments from three scenario earthquakes for each city based on local data and default data. The implications of the findings from each individual case study are analyzed at the end of each of the chapters. Chapter 6 compares the two case studies to uncover any patterns of findings. Finally, Chapter 7 provides a summary of conclusions along with a discussion of future research and limitations of the current research.

Chapter 2: Literature Review and Research Design

2.0 Introduction

The use of integrated models for hazard assessment and disaster management has caught the eyes of disaster managers, researchers, and the scientific community at large. However, there is a great deal of uncertainty inherent in these models – the quality of data input into these models is a large contributor to the uncertainty. Often the data needed are not available at the local level and datasets at the national level do not reflect the reality in most areas, particularly in large cities. This can be a serious setback to the development and use of integrated models for public policy purposes. This is the context of this research and will be explored further in this chapter.

This chapter will examine the literature surrounding the needs and limitations of integrated models for disaster management in the context of large cities. It will develop a conceptual framework for understanding the problems associated with the use of integrated models in light of the reality of data available at the local level. In doing so, it will lay the basis for analyzing the use of the HAZUS™ model for large cities with non-local data and the sensitivity of the model to local level data. It will also examine the use of these models in the context of the availability of local level data in large cities. The chapter starts out by reviewing the context and literature pertaining to use of integrated models for managing disasters in large cities. This is followed by the discussion of the framework under which this research is developed. Research questions based on the literature and the research framework will then be presented followed by a discussion of the research design. The chapter will end with a detailed description of the HAZUS™ model which will be a subject of analysis in this research.

2.1 Research Context and Literature Review

In this section, a review of the current literature pertaining to integrated models in disaster management and their use in large cities in the context of available local data will be presented. Specifically, a review of literature on three topics will be explored – the issue of managing disasters in large cities, the use of integrated assessment models and their applicability in disaster management, and the diffusion of GIS in local government.

2.11 Disaster Management in Large Cities

The management of natural disasters is a complex task but is often simplified by dividing it into various phases: preparedness, response, recovery, and mitigation (Figure 2.1). Various factors add to the complexity of managing disasters: first, the uncertainty associated with hazards makes it difficult to predict their occurrence in space and time. Second, it is difficult to comprehend fully the impact of hazards on human systems and to analyze their effects at various scales and geographies. Finally, public policy related to disaster management is complex owing to the fact that disasters are often treated as "acts of God" and very rare events, making them a low priority for individuals, policy-makers, and politicians.

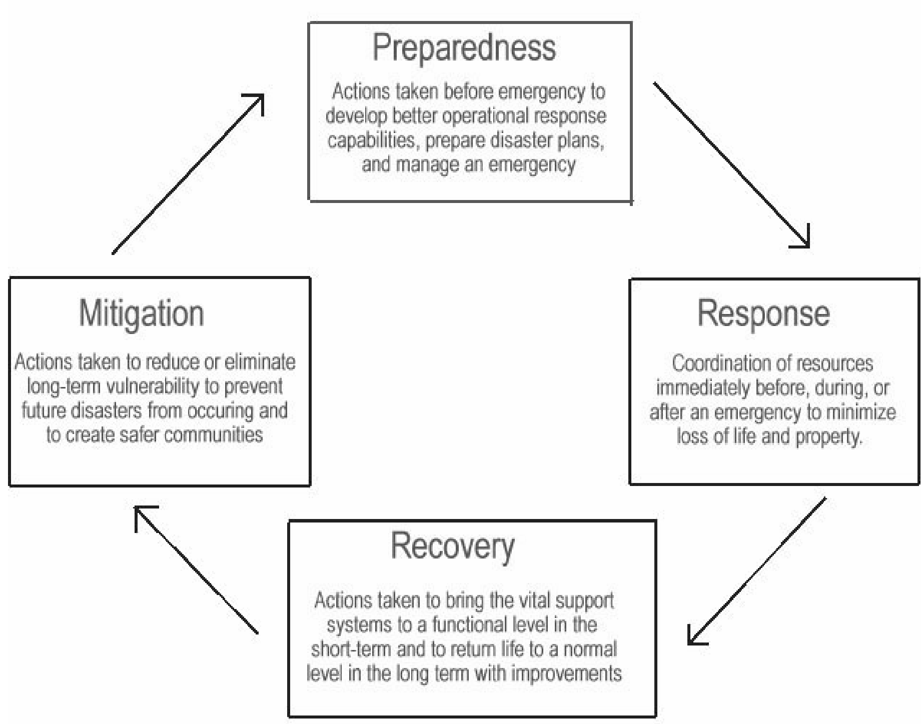


Figure 2.1: Phases of Disaster Management
Source: This figure is compiled from various sources

The complex nature of disasters is exemplified by the fact that the field of disaster research is a subject of study for many disciplines such as geography, geology, engineering, social and behavioral sciences, anthropology, international development, urban planning, public policy and economics. The complexity of disasters and its understanding is further magnified in the context of large cities which have suffered large losses in life and property due to various disasters. The recent past provides ample evidence of the increasing threat of disasters that confront large cities. Among these are: Nagoya, Japan (earthquake), Tangshan, China (earthquake), Bucharest, Romania (earthquake), Adelaide and Melbourne, Australia (bushfires), Mexico City, Mexico (earthquake), Dhaka, Bangladesh (floods), Oakland, USA (fire), Miami/Dade County, USA (hurricane), Los Angeles, USA (earthquake), Cairo, Egypt (earthquake), Kobe, Japan (earthquake) and more recently New Orleans, USA (hurricane).

The number of people in large cities is growing due to the heavy movement of population from rural areas into cities, particularly in the developing world. It has been estimated that there are more than 300 cities with population greater than a million worldwide and this number has grown dramatically in the last 50 years. The United Nations estimates that by the year 2025, 61% of the world's population will be living in cities and there will be 28 "giant metropolitan complexes" of over 8 million people (United Nations Center for Human Settlements1996). The trend of rapid urbanization is not without repercussions for the social, physical, and environmental fabric of the cities. Problems of insufficient and inefficient infrastructure, population growth in marginal lands, increased competition for limited resources and environmental problems such as air, water, and noise pollution accompany most urbanization processes. This affects the normal functioning of the cities, let alone their capacity to cope with catastrophic events. Although many of these trends of urbanization are more common in the developing nations and no longer occurring in the developed world, large cities in the US are having similar issues of crumbling infrastructure, poverty, marginalized population and environmental degradation.

Added to this is the fact that many large cities of the world are located in geophysically hazardous areas such as floodplains, seismically active areas, coastal areas, and cyclone/hurricane prone areas. A study by Degg (1992) shows that 78 of the world's 100 most populous cities are exposed to at least one of the following hazards:

earthquakes, tsunamis, volcanoes, and windstorm; 45 are exposed to more than one of the above hazard. This analysis does not take into account hazards due to flooding which could increase the above numbers significantly. With the current trend of globalization and interconnectedness of economies, the impact of a disaster affecting a large city is seldom confined to that city or its hinterland alone. Ripples of its effect can be felt on the regional, national and often the world economy. The above discussion is brought home by the unfolding of events in New Orleans in the wake of Hurricane Katrina. Therefore, managing disasters in large cities is important and needs to be an integral part of the proper functioning and management of the city.

The question arises, what makes disasters in large cities different from disasters in smaller communities or in rural areas? Although a natural hazard does not follow political boundaries or distinguish between urban and rural area, features of disasters in large cities pose entirely new problems for disaster management as compared to smaller communities (Mitchell 1995). The high population density, complex societal mix, large income gap, abject poverty, and a large informal sector associated with most large cities, lead to a complex intermingling of cause and effect. Not only do large cities pose special problems for the delivery of emergency response services, but recovery in most cities is slow owing to the greater vulnerability of infrastructure networks or what may be termed the lifelines of the city (Mitchell 1995). The concentrated but distributed functionalities such as delivery of food, power, transportation, etc all depend on high tech infrastructure that can break down at the time of a disaster and paralyze the city. Furthermore, many land-use decisions and hazard abatement policies are implemented at the local level.

Many large cities undertake their own emergency management and play the role of first responders to any disaster, particularly in the US.

Therefore, it is important to further the discourse of disasters in large cities and to understand the problems associated with addressing the impact of disasters in large cities. An important aspect of this is to understand the interactions between various urban systems such as transportation, electrical generation-and-distribution, housing, water and other systems which form the “urban system of systems” (Maheshwari 1999). In this complex interdependent “system of systems” (Figure 2.2), a vital aspect is to understand how damage in one or more of these systems can affect other systems and eventually impact the recovery of the city as a whole. This interrelationship between the various interconnected systems is essentially lacking in most of the studies on disaster infrastructure damage (Carrara and Guzetti 1995).

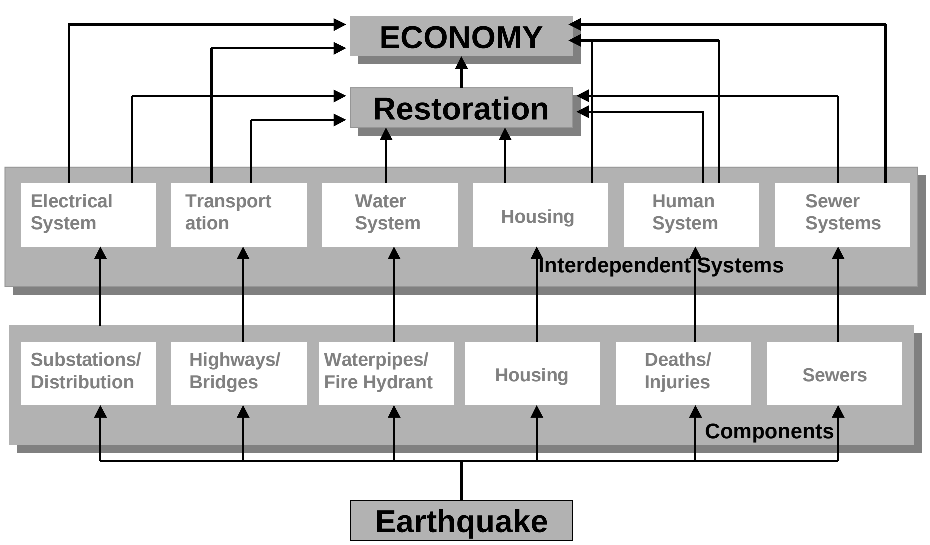


Figure 2.2: Interdependence of Urban Systems

An understanding of disasters in large urban areas from a systems perspective is beyond the scope of a single discipline. Therefore, a multidisciplinary approach is required involving integration of knowledge from disciplines such as geology, seismology, electrical engineering, civil engineering, transportation, sociology, demographics, urban planning, etc. It is also important to translate academic research to real applications that can be used by the decision-makers and can inform public policy.

2.12 Integrated Assessment and Integrated Assessment Models

As mentioned earlier, the field of disaster management is a subject of study for many disciplines. The interdisciplinary nature of disasters has been a double-edged sword. On the one hand it has been instrumental in enriching the field through a multidimensional perspective. On the other, it has contributed to the fragmentation of a holistic understanding of disasters and their impacts as researchers and policy-makers from various fields have talked past each other without exchanging ideas (Alexander 1995). Furthermore, the narrow focus of much of the research in isolated disciplines has limited application for the disaster manager who is responsible for saving lives and mitigating losses due to disasters. Therefore, for multidisciplinary studies such as disaster sciences, integrated assessment provides an excellent opportunity to synthesize a broad range of expert knowledge in advising issues under consideration.

Assessment involves processes (social or technological) that bridge the domains of knowledge to aid public policy. Integrated assessments synthesize knowledge from many different fields of study, and hence allow decision-makers to understand complex systems such as environment, climate, hazards, and other human and technological systems and their integration. Two forms of assessment have commonly been used – deliberation by interdisciplinary expert panels and formal modeling (Parson and Fisher-Vanden 1997). Integrated assessment models (IAMs) are tools that formalize assumptions and relationships between various factors through mathematical computer modeling and help decision-makers analyze various consequences by changing assumptions and relationships. Although commonly used in climate and energy modeling, integrated assessment has also been used in many policy domains such as environmental impact assessment, risk assessment, hazard assessment, transportation assessment, and technology assessment. The quality of the model is highly dependent on the strength of the assumptions and the quality of data that is used in the model (Consortium for International Earth Science Information Network 1995).

In the field of disaster management, hazard assessment has been commonly used at various levels of sophistication to inform land use planning and management (Burby et al 2000; Deyle et al 1998). Deyle et al (1998) refer to three levels of sophistication in hazard assessment:

> *Hazard identification*, which maps the magnitude, probability and threat of hazards geographically;
> *Vulnerability assessment*, which integrates the threat of hazard with exposure and vulnerability of population and investment and their impending losses due to a disaster event; and
> *Risk analysis*, which provides a complete portfolio of risks by incorporating probability of injury and damage from a full range of hazard events in an area.

While hazard identification is the most simple and commonly practiced means of hazard assessment, vulnerability assessment and risk analysis are more useful in making policy decisions (Bernknopf et al 2001). They integrate various fields of knowledge and allow decision-makers to analyze different scenarios, gather support for various policies or choose amongst competing alternatives. Although integrated hazard assessment models using computer technologies have been discussed for a while (French 1986; Haney 1986; Marston 1986; Masri and Moore; 1995), the degree of integration has evolved from understanding one aspect of a hazard on a particular system to many aspects of a hazard and even multiple hazards on many different systems and their complex intermingling. The tools available these days are much more sophisticated in their ease of use and integration of many disciplines.

Models such as HAZUS™ (FEMA 2003), CATS™ (Swiatek and Kaul 1999), TAOS (Watson and Johnson 1999) and other initiatives such as Urban Security Initiative at Los Alamos National Laboratory (Heiken, et al 2000) have attempted to formalize expert interdisciplinary knowledge. Such integrated hazard assessment models integrate the science related to hazards such as earthquakes, floods, hurricanes (or a combination of these), with information about the vulnerability of building stock, population, and networks such as electrical, transportation, and water distribution. By doing so, they aim to facilitate decision-makers in analyzing holistically the impact of various scenarios of disasters, determine the costs and benefits of various policy alternatives, and assess different types of vulnerabilities posed by different populations and systems. Thus, decision-makers can use these models without having expertise on any or all of the above systems or sciences.

The inductive nature of models such as HAZUS™ and CATS™ can lead to “spuriously precise” results which can often mislead the user of the uncertainty involved in the use of such models (Alexander 2000). Furthermore, their root in positivism leads decision-makers to often overlook more difficult questions and focus on the ones that can be answered by the model. For example, decision makers may focus on the amount of debris that will be generated after a disaster rather than complex issues of magnified vulnerabilities due to gender, age, race and socio-economic status that have a significant bearing on disaster response and management. Alexander (2000) argues for deductive models where the relationships and processes in disasters are not rigidly embedded in the models but specified “a priori”. However, Alexander agrees that deductive modeling requires some inductive modeling to form relationships and processes and in doing so, one uncovers changes in the relationships such that the two activities (deductive and inductive) loop iteratively (Alexander 2000). Therefore the importance of inductive models such as HAZUS™ cannot be disregarded. They are particularly useful in providing a framework to organize existing knowledge, understand gaps in the knowledge and hence identify areas for further research. They also help in understanding the uncertainties related to the modeled systems and highlight ones that need further research.

The use of integrated assessment models for hazard assessment is largely dependent upon the degree of reliability of the model, ease of use, and applicability of the models for analyzing policy imperatives at various scales and units of analysis. The reliability of the models is dependent on the strength of assumptions, quality of data, and some clear articulation of the degree of uncertainty inherent in most integrated assessment models, including hazard assessment models. The prediction of “spuriously precise” (Alexander 2000) results as discussed earlier can lead to the masking of uncertainty associated with these models rather than informing users of the uncertainty.

Uncertainty is inherent in hazard assessment models because of the limited scientific understanding of the causal processes of natural events that lead to disasters and their impact on the built environment. This is further compounded by the quality of data input into the models and the level of aggregation (i.e. geographic scale of analysis). While it is not easy to reduce the uncertainty attributable to the lack of understanding of scientific phenomenon (and certainly outside the domain of the urban planning discipline), the uncertainty due to data quality and aggregation can be better understood, addressed and controlled to a large extent. The need for reliability is higher for smaller geographical scales where results/outputs from models inform decisions about implementing hazard abatement policies and formulating strategies for saving life and property. Therefore the need to provide some explanation of uncertainty and reduce uncertainty is very important.

Most existing models for damage estimation and assessment are very data intensive and need to model the physical built environment in great details. One of the major components of the built environment is the building stock whose characteristics (physical, use, and economic value) determine the amount of direct and indirect economic loss from any natural hazard. The impact of building inventory in vulnerability assessment is crucial. The damage to buildings results not only in the largest proportion of direct economic losses due to disasters, but also determines the number of casualties, need for shelter, hospitals and other emergency services. They are also large contributors to indirect damages and cost due to work interruption, debris removal and other economic impacts. Figure 1.2 shows the interrelationships between various modeled systems in HAZUS™ and their contribution to direct and indirect losses.

Although the building inventory is central to damage assessment, there are few sources of data that can provide a comprehensive picture of the building stock for any modeled geography. A report produced by the Applied Technology Council discusses three levels of building inventory: level 1 involves the use of existing facility-specific databases, level 2 involves the synthesis of building inventory from economic data such as number of employees and annual production amounts, and level 3 involves the coarse estimation based on population and other parameters (ATC 1985). The reliability of the model decreases as one move from Level 1 to Level 3 data. Many damage assessment models have used default data that is derived from various sources such as expert panel knowledge and inferred from secondary sources such as population, employment, and economic output, etc. The use of this generalized, coarse level data for estimating the housing and building stock of a region makes such models suitable only for use at larger scale applications.

However, datasets about the built environment, such as parcels and tax assessment data are now commonly available at the local level for much of the United States, particularly for larger urban areas. These datasets provide a fairly accurate picture of the built environment at the most disaggregate level possible, and the use of such data can render formal models useful for small-scale applications as well. In fact the HAZUS™ model has become widely used at the local level for analysis of disaster outcomes to inform policy decisions. Through the use of local level data, one can at best reduce the uncertainty associated with large-scale aggregation or at least understand the uncertainty to manage such events better. Without the latter, there is a danger that practitioners will either derive erroneous conclusions or will simply not use these models.

Therefore, it is crucial to understand the impact of disaggregate data (available at the local level) in comparison to the use of generalized estimates on hazard assessment models. It is also important to understand the sensitivity of the widely used hazard assessment models to local level data.

Information on the use, condition, height and type of structure of buildings is important to model the built environment and its response to a natural hazard. While information on square footage based on the type of use is available through a variety of local sources, the information on the type of structure is more difficult to get from any standard database. The ATC-13 report uses various methods to infer the distribution of structure types based on occupancy or use of the building for a given zip code from information about the age and height of the buildings (ATC-13 1985). Such matrices are generated through the use of tax assessment data and workshops of building experts in various regions (HAZUS User Manual 2003). If such matrices were to be generated for many geographical regions (rather than for East, West and South as is the case in HAZUS™), they would provide a better estimation of structure type. However, such a task is quite difficult and hence for a large project like HAZUS™, encompassing the entire country, regional matrices are drawn up and used as defaults.

While the use of generalized estimates is suitable at large scales encompassing several counties, they do not represent the reality of urban areas and cities. Research has shown that often for large cities, real data is significantly different from generalized default data (Nordenson et al 1999; Wiggins 2000). Nordenson et al (1999) show the extreme case of the Wall Street census track in New York City by comparing the default data in HAZUS™ with locally collected inventory as shown in Table 2.1 and 2.2 below.

Table 2.1: Comparison of Building Inventory by Occupancy Class: HAZUS™ Default vs. Local

Occupancy	Hazus™ Default		Wall Street Census Tract	
	(Square feet)	(Count)	(Square feet)	(Count)
Residential	201,800	10	991,914	5
Commercial	19,599,500	649	38,574,963	57
Industrial	567,800	31	-	-
Agriculture	-	-	-	-
Religious	250,700	17	18,468	1
Governmental	-	-	-	-
Educational	52,800	3	-	-
Total	20,672,600	710	39,585,345	63

Source: Nordenson et al (1999)

Table 2.2: Comparison of Building Inventory by Structure Type: HAZUS™ Default vs. Local

Structure Type	Hazus Default (count)	Wall Street Census Tract (count)
Wood	154	3
Steel	264	49
Reinforced Concrete	46	4
Precast Concrete	22	3
Reinforced Masonry	61	3
Unreinforced Masonry	163	1
Mobile Homes	-	-
Total	*710*	*63*

Source: Nordenson et al (1999)

Such changes in the building inventory can result in damage estimates that are as much as 76 percent off the generalized default values or off by a factor of 4 (Nordenson et al 1999). Therefore, the use of these models with default data, particularly for cities, can lead to very erroneous results and can lead to flawed outcome in public policy. Therefore, to implement policies at the level of a local government or at a smaller scale, there is need for accurate data at a low granularity of aggregation i.e. block-level census data and parcel-specific data on landuse and structure (Deyle 1998). Also, for large cities a better understanding of uncertainties associated with using regional averages is needed since they deviate more from national estimates.

While the Wall Street census tract in New York City is an extreme case that is not particularly reflective of most cities and also not reflective of all census tracts in New York City, the above results establish the variation of local data from default data in HAZUS™ for one census tract. Furthermore, they point to the importance of analyzing the variation of the local data at the scale of an entire city to understand where large discrepancies occur and how they impact the results from the model with all other parameters remaining constant. This will help determine areas where national averages are adequate and where the need for local data is imminent. An analysis of such spatial patterns will help decision-makers decide where to spend their limited resources to improve the data that are input in these models. Also, if resources do not allow the use of local level data, it will help decision-makers assess the uncertainty of results obtained when default data are used and analyze spatially where such uncertainties exist so that they can make more educated policy decisions.

In this section, it is established that integrated models can be very useful for disaster management in large cities. The role of technologies such as geographic information systems (GIS) and remote sensing have been instrumental in the advancement of integrated hazard assessment at all levels of sophistication, by providing an appropriate means of representing, analyzing, and visualizing information which is inherently spatial in nature. GIS is used as the integrative framework for input and output of local data in most of the recent integrated models such as HAZUS™, CATS™ and TAOS. However, the need for local level data (particularly GIS data since most data input into models are GIS data) is important for the increased reliability of these models.

The next section examines the literature surrounding the availability and diffusion of GIS in local government.

2.13 Geographic Information Systems and its Diffusion in Local Government

Disaster management is one of the most spatially-oriented of all management sciences (Morentz 1986; Drabek 1991). The importance of geographic information systems and other spatial technologies in the management of disasters is well established by its use in many disasters and disaster-related research (Eichenbaum 2002, Greene 2002, Dymon 1999, 1993, 1990, Coppock 1995, Waugh 1995). The diffusion of GIS in local government has been rapid in areas of public utilities, property assessment and taxation, and planning. Although the use of GIS for disaster management has spanned all phases of the disaster management cycle (Figure 2.1) and all types of hazards (Radke et al 2000), these technologies have not diffused widely in disaster management in local government applications. This is largely due to the above-mentioned perception that disasters are rare events for which major investments in technologies are not warranted. However, the use of spatial technologies in the management of the events of September 11, 2001, as well as other disasters such as Hurricane Andrew and Hurricane Fran (Dymon 1999, 1993), Oklahoma City bombing, etc. has elevated the role of GIS for managing adverse events in everyone’s eyes.

Local governments are now leveraging their GIS investments to collect more data on the location of various critical infrastructures or are now investing heavily to acquire such systems. At the same time, the federal government is striving to create it’s own inventories of various datasets for urban regions, as evidenced by efforts to create the National Map (US Geological Survey 2001) and 120 Cities Project (Elber 2002). GIS-based integrated damage estimation and assessment tools such as HAZUS™ (FEMA 2003), and CATS (Swiatek and Kaul 1999) are being promoted by the Federal Emergency Management Agency (FEMA) and other federal agencies for use by local agencies. The use and integration of local level data is strongly advocated, and often assumed in such efforts without a proper understanding of the ground reality of GIS development at the local level. Without a proper assessment of what already exists and how local governments are using their GIS systems, such top down efforts can often lead to duplication of efforts and be counterproductive to real local needs (Elber 2002). Furthermore, such efforts can have little benefit at the local level and risk rejection by local disaster managers and first-responders.

There are a few studies on the diffusion of GIS (Masser and Onsrud 1992; Masser, Campbell & Craglia 1996; Chan and Williamson, 1999). However, there are limitations of these studies for this research since none of them focus on particular datasets or use of GIS for disaster management. The Framework Data Survey was conducted by the Federal Geographic Data Committee (FGDC) to assess the diffusion of GIS in local government throughout the US in the late 1990s (Sommers 1999). This survey also has limited utility for this research because of methodological fallacies. The goal of the survey was to assess the status of Framework data for the entire country including federal, state, local, private, and other GIS data producers. In order to include all 50 states, state-level survey coordinators were established but no consistent methodology was used to determine the sampling frame for each state. Each coordinator was free to distribute the survey to GIS users and in the various levels of government without any sampling or follow-up methodology (Tulloch and Fuld 2001). These methodological limitations restrict the widespread application of the results of this survey (Tulloch 2000, Tulloch and Fuld 2001). Similar to the FGDC's effort, states have also undertaken their own surveys to assess GIS activity at the local level (Shanley et al 2001). Various other studies conducted elsewhere in Europe and US have either taken a broad look at diffusion of GIS in local government (Masser 1992; Wiggins 1992) or have focused on diffusion in particular cases (Rumor 1992). Furthermore, most of these studies, conducted in early to mid nineties are now outdated with the rapid evolution of this technology. Other surveys conducted by private agencies such as Gartner (2002) have methodological limitations with respect to sampling and sampling frame, and also focus on a broader aspect of GIS diffusion rather than on particular datasets (Gartner 2002).

There are no studies that focus on questions such as what datasets are commonly available for a specific type of geography (e.g. large urban areas), what kind of organizational structure does GIS operate in, and how GIS is being used to support disaster management needs. Hence there is still a critical need to look in-depth at the diffusion of GIS in the case of large US cities and to analyze the use of GIS applications and GIS data in the context of disaster management.

2.2 Research Framework

Based on the above review of the literature, it is apparent that GIS-based hazard assessment models can be very useful for disaster management at the level of cities.

However, the appropriate and sustained use of these tools in local government is largely dependent on a delicate balance between three factors:

- Reliability of the models which is largely determined by the assumptions and scientific strengths of the models and the data used to drive these models, particularly in the context of smaller scales such as cities;
- Availability of local-level GIS data (i.e. the degree of diffusion of GIS in local governments), particularly in the context of disaster management; and
- Ease of use and applications for analyzing public policy.

These three factors can be analyzed in the framework of a three-legged stool, every leg serving a critical function in maintaining equilibrium for the sustainable use of integrated models for hazard assessment (Figure 2.3).

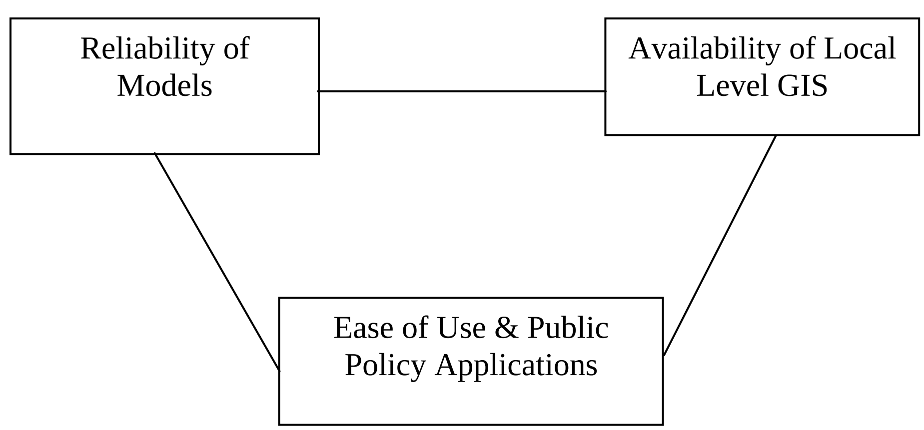


Figure 2.3: Factors Affecting the Sustainable Use of Integrated Models for Hazard Assessment

The issue of reliability of integrated models, the scale of analysis and quality of data can be analyzed as follows: the reliability of integrated models needs to be higher as the scale of analysis increases (i.e. analysis at the scale of a city or small county) since local-level decision-making requires greater details. Also the quality of data improves the reliability of the model. As advancements are made in hazard assessment sciences, the reliability of hazard assessment models is more and more dependent on the input and integration of local-level data, particularly for decision-making at smaller scales. Likewise, as the diffusion of technologies such as GIS increases at the local level, hazard assessment models get more advanced, and the utility of such models in local-level decision making increases dramatically. Therefore it is important to assess the diffusion of GIS at the local level to understand the ground reality of availability of GIS data. The use of scientifically advanced models with inadequate and poor data can lead to so much uncertainty in the results that the appropriate use of these models is jeopardized. However, as shown in Figure 2.3, more advanced models and more data can also mean the need for high degree of expertise in the use and manipulation of data, which can be a deterrent in the use of hazard model. The ease of use and the application of the models to analyze scenarios that inform public policy decisions related to competing alternatives will also determine the degree of acceptance of these models.

This dissertation aims to address issues associated with two of the three factors mentioned in Figure 2.3 that determine the use of integrated assessment models for hazard assessment in the context of large cities – the issue of reliability of the models and the availability of local data. The third factor, related to the ease of use of models is beyond the scope of this dissertation. This dissertation deals broadly with issues concerning data availability for doing disaster damage assessments in larger cities in the United States and on the appropriate use of such methodologies in assisting at decision-making, planning and policy. Specifically, it focuses on the availability of building inventory data (particularly from tax assessment rolls) for earthquake damage estimation using the earthquake damage estimation model HAZUS™ in the context of large cities in United States. The dissertation also aims to understand the sensitivity (and hence the reliability) of the HAZUS™ model to local-level data pertaining to building inventory and the challenges associated with using local data. The HAZUS™ model is used in this study because it currently represents the state-of-the-art in earthquake damage estimation. This model is developed by the Federal Emergency Management Agency (FEMA) in conjunction with National Institute of Building Sciences (NIBS) and RMS Inc, and is intended to be applicable for the entire United States. Furthermore, with the release of the wind and flood module, this model is increasingly being used in various contexts throughout the whole country. The next section discusses in detail the specific questions that this dissertation will seek to investigate.

2.3 Research Questions

Based on the research framework discussed in Section 3.0, this research will address the following questions:

1) Given the data requirements of hazard assessment models for disaster damage estimation, what is the state of GIS in large cities for sustained use of integrated models at the local level?

This research seeks to understand the state of development of GIS in large cities, both organizationally and in terms of development of key GIS datasets. It seeks to investigate the availability of various datasets including those needed to determine the characteristics of buildings in the city – this includes the height, age, use, construction material, etc. The research looks at the extent to which GIS has been used for disaster management in large cities. It analyzes the various factors that induce or hinder the use of GIS for disaster management. This question helps in understanding how useful integrated models will be for local-level decision-making. An understanding of ground reality for local level data also informs the development of these models and other federal initiatives. This research also aims to identify strategies for inducing change at the local level for appropriate GIS development which can be more useful in managing disasters (while also serving other local needs and mandates).

2) How do default estimates for building inventory in HAZUS™ compare with local data for building inventory?

This research focuses on the local data for buildings and compares it with default data available in HAZUS™. It analyzes the strengths and weaknesses of the tax assessment data and look at ways that the tax assessment data can be improved using other datasets. It looks at the problems associated with transforming local data (particularly tax assessment data) into a format required by HAZUS™. Finally, it analyzes how much the default data in HAZUS™ deviates from locally available data. The analysis is conducted with the city as the unit of analysis. It looks at spatial patterns of variation of default data from local data over the tapestry of the city. Therefore, the research investigates whether the default data deviates more in certain parts of the city or in certain types of areas such as residential, commercial, and industrial areas or if there are specific patterns of deviation. Overall, this question seeks to inform local decision-makers about the issues associated with using local data, the degree to which the default data in HAZUS™ is different and where in the city (if any) the default data is significantly different from the ground reality. This can help determine appropriate strategies for integrating local data and reducing the uncertainty in HAZUS™. It will also help identify areas of future research with respect to improving local data.

3) How sensitive is the HAZUS™ damage estimation model to improvements in building inventory data from local sources?

This research analyzes the changes in damage estimates based on local level building information. It looks at how changes in inventory affect various outputs such as total loss, building loss, shelter needs, amount of debris, etc. It seeks to understand the sensitivity of the HAZUS™ model to earthquakes of various magnitudes to see if the changes in loss estimates is higher for certain magnitude events. It also analyzes the spatial variation in damage and losses at the city level and at the census tract level across the city. This question will help decision-makers understand the degree of uncertainty in the HAZUS™ model, and help them determine where efforts should be put to improve data. It will also help understand the sensitivity of the HAZUS™ model to improvements in data and will help determine areas of further research to improve damage assessment models.

2.4 Research Methodology

To address the above questions, the research design consisted of a four-step process:

In the first step, 19 cities with population between 250,000 to 1 million in the United States were selected randomly. This represented 30% of the cities in the US in this population range. A survey/interview of chief GIS officials in city agencies dealing with GIS was conducted. This survey had a two-fold objective.

1. To understand the level of GIS diffusion in large cities in the US (both in terms of GIS data as well as the organizational structure of GIS implementation), with a focus on the use of this technology for disaster management, and
2. To identify cities that could be used for further case study for this research.

A survey questionnaire (Appendix B1) was mailed to the GIS coordinator with a follow-up time for a telephone interview. Through this combination of mailed questionnaires and phone interviews, the following information was solicited about GIS development and diffusion in all the cities:

In the first part, the interviewees were asked to provide information about themselves (title, GIS expertise, knowledge of HAZUS, etc), their organization (size of the jurisdiction in terms of number of parcels and area, department which housed GIS, type of GIS organizational structure in the jurisdiction, number of personnel with GIS job descriptions or those that used GIS as an integral component of their job), inter-local agreements for spatial data sharing, and methods of GIS data dissemination (such as use of Internet-based mapping, etc).

The second part of the questionnaire was designed to identify the key players and custodians of various spatial datasets in the organization in the event that information about some datasets was not available with the manager/coordinator. This would be expected in cities where GIS was managed in a more distributed organizational framework (i.e. many departments undertaking their own GIS data and application development without any centralized coordination). These datasets included parcels, tax assessment data, road centerline, building footprint, orthoimagery, topography, utilities and critical facilities such as location of schools, hospitals, emergency management facilities and hazardous waste sites.

In the third part of the questionnaire, detailed questions were asked about each of the above datasets. Such questions pertained to the role of the respondent in the creation, maintenance, update or distribution of the particular dataset, its completion, positional accuracy, currency, format, cost, its use for disaster management, and the willingness of the respondent to share that data for more detailed analysis using HAZUS™. The survey questionnaire is provided in Appendix B1.

In the second step of this process, two cities were selected from the above nineteen cities for more detailed analysis. The selection of the cities was made based on the findings of the survey about availability of GIS information and access to it. For the selected cities, local-level data were collected including parcels, building footprints, tax assessment, and digital orthoimages. These data were then analyzed to understand the challenges associated with using the local data for input into HAZUS™. Since the tax assessment data are a primary source of information about building characteristics that are widely available throughout the country, this dataset was a central focus for further investigation. The emphasis was on the quality of data, the level of aggregation in the data, the degree of integrity of the data, and the amount of information available. The data was then prepared to be input into HAZUS™ to evaluate the difference between HAZUS™ default data and local-level data. This evaluation was done for the entire city and for the various census tracts that comprise the city to uncover how local data deviated from default data.

In the third step, the HAZUS™ model was run for the selected case cities for three scenarios which involved earthquakes of different magnitude at the exact same location. The variation of the output results from the model was analyzed in detail for each city under consideration. The focus was on the variation of the results due to better building inventory data from local sources as compared to the default data. Results for building damage, economic loss, casualties, and shelter capacity were evaluated to see which of these results changed most significantly with the input of more local data. Results were analyzed at the city level and at the level of the individual census tracts that comprised the city. Finally, in the fourth step, the two case studies were compared to understand if there were any consistent patterns of variation between the two cities and if there are ways to improve the default data in HAZUS™ rather than going through the elaborate process of inputting local data.

Based on the above four steps, conclusions are made on whether damage assessment at the city level is feasible or even meaningful given the data limitations. The conclusions also include analysis of whether improvements in data collection are warranted if HAZUS™ is used at this level of analysis. Furthermore, the research will recommend where better data absolutely needs to be collected and if so, the appropriate target areas for collection of better data.

Since the object of analysis in this research is the HAZUS™ model, the next section provides a detailed description of the model, including data needs, and sources of data and the issue of uncertainty in the model.

2.5 The HAZUS™ Model: Standardized Earthquake Loss Estimation Methodology

The HAZUS™ model, developed by FEMA and NIBS (National Institute of Building Sciences) is meant to help decision-makers at the local, state and regional level understand the impacts of various earthquakes, assess the level of damage and losses, test various alternative mitigation strategies and prepare for response and recovery (Figure 1.1).

The methodology involves the integration of various models and knowledge domains to assess the level of damage and loss for a given earthquake scenario as is shown in Figure 1.2. The user can select a scenario for an earthquake (the model accommodates both deterministic and probabilistic analysis). The ground motions based on the selected scenario are computed and applied to various systems such as building stock, essential facilities, transportation and utility lifelines. Through the use of various damage functions for each of these systems, direct physical damage, induced physical damage as well as direct and indirect economic losses are calculated. The model is modular, i.e. analysis can be performed by various levels of expertise and at various levels of detail. For example, detailed ground motion maps from a past earthquake can be provided or a scenario can be simulated. Likewise, the model comes with some default data on building stock based on regional estimates (ATC 1985) but local data can be input into the model for more accurate analysis.

As can be seen from Figure 1.2, building inventory is a major component of the loss estimation methodology. It is not only one of the largest contributors for direct economic and social loss, but also impacts the induced physical damage such as amount of debris, hazardous material and work interruption. Consequently, it contributes largely to indirect economic losses. Because it is computationally difficult to model the impact of the earthquake on every building in a region, the model aggregates the building inventory to the census tract. Various other lifelines are also aggregated to this unit of analysis. The model breaks up the building inventory into 33 specific and 7 general building uses (called occupancy classes in HAZUS™) and 36 different structural building classifications (called building type in HAZUS™). The classification and their description is provided in Appendix A, Table 1 and Table 2. Although the model is packaged with some default data and default mapping schemes, local level data can be input into the model for more accurate assessments. The Building Inventory Tool (BIT) is used to input local-level building data into the model.

The default building data in HAZUS™ is derived from various sources – the default square footage estimates for occupancy classes RES1 (Single Family Residential), RES2 (manufactured housing), RES3 (multi-family housing including duplex, triplex, quadruplex and multi-unit apartments), and RES5 (institutional dormitory) are based on census data (on number of dwelling units or number of people for that occupancy class). The square footage information for the remaining occupancy classes is obtained from a building square footage inventory database from Dun and Bradstreet Company, and based on SIC (Standardized Industrial Classification) codes mapped to NIBS occupancy classes (HAZUS™ Users Manual 2003), shown in Appendix A Table 3.

For local level data to be integrated into HAZUS™, information on building occupancy (or land use), age, square footage, height of building, the type of structure, building value, content value and seismic design level are required. The inference on the type of building structure is made based on various matrices in HAZUS™. The Building Inventory Tool (BIT) changes the square footage of the various occupancy classes in the default data and also creates new occupancy matrices that reflect local data.

Based on the building inventory and hazard parameters (such as peak ground acceleration, spectral response in the case of earthquakes), various damage functions are used to compute the probability of various types of damage – slight, moderate, extensive and complete. Fragility curves, developed for every building type, define the probability of being in a certain damage state based on the size of the earthquake. Figure 2.4 shows an example of a fragility curve. Most inputs can be modified by the user including building inventory, soil characteristics, fragility curves and other parameters used to calculate losses. Results can be displayed both spatially (in the form of maps) or in tabular formats and reports.

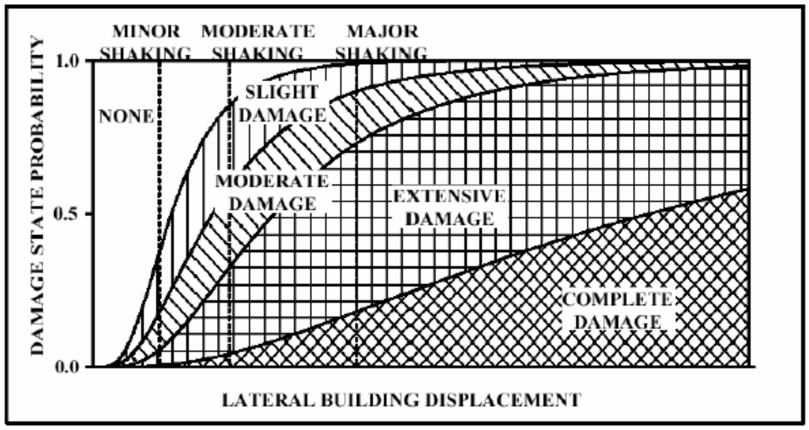


Figure 2.4: Example of a Fragility Curve in HAZUS™

While the use of default data requires minimal effort by the user, modification of the defaults or input of local-level data requires significant expertise in GIS, databases and in other aspects of the system being modified. The issue of sensitivity and uncertainty is discussed in the User Manual as:

> "Any region or city studied will have an enormous variety of buildings and facilities of different sizes, shapes, and structural systems constructed over years under diverse seismic design codes. Similarly, many types of components with differing seismic resistance will make up transportation and utility lifeline systems. Due to this complexity, relatively little is certain concerning the structural resistance of most buildings and other facilities. Further, there simply are not sufficient data from past earthquakes or laboratory experiments to permit precise predictions of damage based on known ground motions even for specific buildings and other structures. To deal with this complexity and lack of data, buildings and components of lifelines are lumped into categories, based upon key characteristics. Relationships between key features of ground shaking and average degree of damage with associated losses for each building category are based on current data and available theories. While state-of-the-art in terms of loss estimation, these relationships do contain a certain level of uncertainty. Ranges of potential losses are best evaluated by conducting multiple analyses and varying certain input parameters to which the losses are most sensitive." (HAZUS 2003)

Chapter 3: State of Local Data in 19 Cities

3.0 Introduction

The use of GIS in local units of government has increased significantly in the last decade. While local land management functions may have provided the initial impetus for the investment in GIS, other application areas such as public safety and disaster management have now become the driving force for continued investment in these technologies. However, there are very few studies that focus on the diffusion of GIS in large cities and particularly in the context of disaster management and disaster damage assessment. Most existing studies, as discussed in Chapter 2, focus either on GIS diffusion in general or on anecdotal applications of GIS for a specific disaster and particularly in the disaster response phase. There is no systematic study of the diffusion of GIS at the local level (in terms of the organizational setup and the availability of GIS datasets) and its application for disaster management and disaster damage assessment. Therefore, this research sets out to conduct a survey of 19 cities to assess the extent of GIS diffusion in local government, particularly in the context of disaster management in large cities. This survey also aims to help identify cities that would be suitable candidates for further in-depth study of the sensitivity of the HAZUS™ model to local level data.

Thus the purpose of the survey is twofold:

1. To understand the level of GIS diffusion in large cities in the US (in terms of availability of GIS data as well as the organizational structure of GIS

implementation), with a focus on the use of this technology for disaster management.

2. To identify two cities that can be used for further case study for the HAZUS™ model.

This chapter discusses the methodology used for this survey with respect to the selection of cities to be surveyed, the implementation of the survey, and the design of the survey instrument. This is followed by a discussion of the findings of the research with respect to the GIS organizational structure and the state of development of critical GIS datasets. For each category there is a brief description of the topic and it's relevance to disaster management followed by a discussion on the findings of the research and its implications. Finally, the chapter concludes with a discussion of the implications of the overall findings for the use of GIS in disaster management and disaster damage assessment in the context of large cities.

3.1 Methodology

3.11 Selection of Study Cities

The term "large" as applied to cities has been used rather loosely in this research so far. It is now important to define what constitutes a "large" city for this research. For the purposes of this research, the population size of the city (in terms of 1990 census population) was the sole criterion in the selection process. This research focused on cities with population size between 250,000 to 1, 000,000 according to the 1990 census. The US Census (1990) lists 200 cities with a population of over 100,000 – of which 8 cities have population over a million, 56 cities have population between 250,000 and 1 million, and the rest with population less than 250,000. The 8 cities with population greater than a million were too large and unique cities to be considered suitable for further case studies since the results could not be generalized. Furthermore, they were considered to be difficult to manage from a data analysis perspective because of their large spatial extent. On the other hand, many of the cities with population between 100,000 and 250,000 were either smaller freestanding cities that lacked a diverse economic base or edge cities of larger metropolitan regions. For example cities such as Ann Arbor, MI (population 109,592), New Haven, CT (population 130,474), and Madison, WI (population 191,262), were primarily university-based cities. Others such as Pasadena, CA (population 131,591), Arlington, VA (population 170,936), etc. were integrally linked to larger metropolitan regions or a larger city and posed special challenge in analysis as a separate entity.

The 56 cities that fell in the population range of 250,000 - 1 million are mostly large urban centers with a diverse economic base, sufficiently mixed land use, and strong municipal governments that are responsible for provision of public utilities. These cities are also usually responsible for managing disasters in their jurisdiction and are first responders to disasters. Hence this threshold of population served as a good basis for defining the pool of cities from which the sample of survey cities could be drawn. Although the threat from an impending disaster (particularly natural disaster) or the occurrence of major disasters in a city could be a criterion for the selection of the survey sample, this was not used at this stage of the survey. This was because the focus of this survey was on the general development of municipal GIS and not necessarily on the development of GIS in cities with hazard susceptibility.

A total of 19 cities were randomly drawn from the population of 56 cities (30%) for the survey. A sample size of 19 provided a good range of cities to draw generalizations from and also to provide sufficient candidates for further case studies. The randomly drawn sample yielded the following cities (as shown in Figure 3.1):

Figure 3.1: Location of Selected Cities for Survey

Atlanta (Georgia), Birmingham (Alabama), Colorado Springs (Colorado), Honolulu (Hawaii), Jacksonville (Florida), Las Vegas (Nevada), Long Beach (California), Miami (Florida), Milwaukee (Wisconsin), Minneapolis (Minnesota), Newark (New Jersey), Oklahoma City (Oklahoma), Omaha (Nebraska), Portland (Oregon), San Antonio (Texas), Santa Ana (California), Seattle (Washington), and Wichita (Kansas). As expected, this sample represented a good spatial distribution across the country as shown in Figure 3.1.

3.12 Implementation

In order to assess the diffusion of GIS in the selected cities, a survey questionnaire with a follow up telephone interview was chosen as an appropriate method. This was driven by two assumptions based on the size of the cities chosen for this research:

1. A majority of the chosen cities would have well developed GIS, and
2. Most well developed GIS programs are likely to have a GIS Coordinator or a GIS Manager who would be an appropriate point of contact.

It was further assumed that once such a person was identified, much of the information about the GIS program including the availability of various datasets, the organizational structure, and the use of GIS for disaster management would be available through this person. Extensive research on the Internet was used to identify the GIS Coordinator/Manager. Where this information was not easily available, some appropriate departments such as planning, information technology, public utilities, engineering, etc. were identified and a few phone calls identified the presence or absence of the right person for interview. Where such a person was not present, information was gathered from different sources in the different departments or from a person knowledgeable about the overall GIS program. The identified persons were contacted and all the persons agreed to participate in the research. A questionnaire was prepared to guide the research and a copy of this questionnaire was provided to the participant before the actual telephone interview. This provided the respondents with prior knowledge of information they would be required to provide and the opportunity to gather the necessary information in advance. The telephone interview took over an hour for each respondent. Some of the participants chose to fill out the obvious questions in the questionnaire and were available for phone interview for further discussion. In such cases the telephone interview took less than an hour.

3.13 Questionnaire Design

The questionnaire was designed to solicit information on three aspects of GIS development:

In the first part, the respondent was asked to provide information about themselves (title, GIS expertise, knowledge of HAZUS™, etc), their organization (size of the jurisdiction in terms of number of parcels and area, jurisdiction, department in which GIS was housed, type of GIS organizational structure in their jurisdiction, number of personnel with GIS job descriptions or those that use GIS as an integral component of their jobs), inter-local agreements for spatial data sharing, and methods of GIS data dissemination (such as use of Internet-based mapping, etc).

The second part of the questionnaire comprised a table designed to identify the key players and custodians of various spatial datasets in the organization in the event that information about some datasets was not available with the manager/coordinator as would be expected in a more distributed organizational framework.

In the third part of the questionnaire, detailed questions were asked about each of the datasets. Such questions pertained to the role of the respondent in the creation, maintenance, update or distribution of the particular dataset, its completion, positional accuracy, currency, format, cost, its use for disaster management, and willingness to share data for more detailed analysis in this research. Most of the datasets that were analyzed were ones identified as part of the Framework Data by FGDC (parcels, road centerlines, tax assessment data, orthoimagery, topography, utilities information, location of schools etc.) although some datasets were added due to their importance in disaster management, e.g. building footprints, location of fire stations, police stations, etc.

Although for the most part, the design of the questionnaire adhered to the Framework Data Survey (www.fgdc.gov), in all the above sections special emphasis was placed on analyzing the impact of GIS on disaster management and vice versa. For example, in the first part, respondents were asked about the type of hazards that pose an imminent threat to their city. Answers were also solicited on their knowledge of emergency management operations within the organization. In the third part, information was gathered regarding the use of every dataset in emergency/disaster management and the phase of disaster management that such efforts occurred. These questions provided a good understanding of the degree of GIS diffusion in management of adverse situations. A copy of the survey instrument is provided in Appendix B1.

It is important to note that this research was started in July 2001. Interviews with 7 cities were conducted before September 11, 2001. After the events of September 11, the environment was not conducive to contacting people over telephone and seeking answers to questions regarding the GIS framework, data availability and data sharing. Hence there was a brief lapse in this part of the research and the research was resumed in November 2001. There was no perceived bias introduced due to the September 11 events although the use of GIS for managing disasters had been elevated in the minds of most participants and their higher management. Some interviewees reported that they had to remove access to data through the Internet due to public security concerns.

3.2 Research Findings

The findings of this portion of the research are discussed in this section and are presented in three categories: the characteristics of the respondent, the characteristics of the organizational and the data availability and use for disaster management.

3.21 Respondent Characteristics

A single respondent was used for each city. In most cities, the GIS Manager or someone of similar job classification was contacted. Most of these people contacted were the actual respondents for the interview. However, some of them looked over the questionnaire and delegated the questionnaire to someone else in their organization. All the respondents were familiar with GIS operations in their organizations. Respondents from 11 cities were GIS managers/coordinators/directors, whereas respondents from 3 cities were GIS analysts. Respondents from 3 cities were IT engineer/supervisor/design engineers, 1 respondent was a city surveyor and 1 was a planner. All the respondents considered themselves to be GIS experts and rated themselves as users of GIS familiar with all GIS functionalities – i.e. creating and manipulating spatial databases, spatial analysis and display. Only 6 of the respondents had ever heard about HAZUS™ and none of these six respondents had ever used HAZUS™.

3.22 Organizational Structure

The organizational structure under which GIS is implemented in an enterprise is very important in the use of GIS for disaster management for all aspects of the disaster life cycle which includes phases of preparedness, response, recovery and mitigation. The GIS organizational structure is a major contributing factor to the effective use of GIS in any phase of the disaster management cycle. As mentioned in Chapter 2, the understanding of the impacts of an event is a complex process and requires the integration of knowledge and datasets from various disciplines and sources. Hence the nature of relationships between different departments within an organization, the integration and standardization of hardware, software and data are key factors to the effective use of GIS for disaster management. Broadly speaking, the organizational structure encompasses not only the structure within the enterprise but also the linkages between the enterprise and other GIS consumers and producers in the region. It extends also to mechanisms for public access to data and maps. The survey was designed to solicit information on all three aspects.

To understand the role of the GIS organizational structure in disaster management, it is important to understand the progression of organizational structures involved with GIS production and consumption along with the progression of GIS technology. The implementation of GIS started with the use of GIS software on isolated computer workstations in the back offices of many departments within an enterprise. As technology evolved, and more spatial data were readily available and more applications of GIS unfolded, spatial technologies diffused through the various departments in the organization. Many different departments started their own spatial data development and often established their own GIS programs. This resulted in GIS implementation taking place in different departments in the same enterprise with very little inter-departmental coordination regarding data and software standards. Eventually this led to redundancy, duplication of efforts and higher costs. This implementation was commonly called the "Departmental GIS" approach.

The problems with the "Departmental GIS" approach paved the way for a more coordinated effort in managing the GIS needs of the whole organization. This led to an "Organizational GIS" or an "Enterprise GIS" approach whereby all GIS efforts in an enterprise were coordinated; a GIS Manager or GIS Coordinator position, responsible for coordinating the GIS efforts throughout the enterprise, was created. Furthermore, a centralized database was used to store all the enterprise datasets in a standard format that could be used interchangeably and was accessible to all users across the various departments. Some enterprises went as far as having a central GIS division that consolidated all GIS data, applications and support for all the departments.

In the late 1990s and early 2000, the Internet unleashed a new model of GIS commonly called “Community GIS” or “Societal GIS” whereby the power of the World Wide Web made sharing of data between organizations easier, and provided an avenue for easy access of geographic information to the public. Although, this model does not strictly comply with an organizational structure, it merely extends the organizational structure, whether Departmental or Enterprise. It allows for regional data sharing and analysis along with increasing public participation.

The cities in the sample display some variation in terms of organizational evolution discussed above. Of the 19 cities, 13 cities have moved to an “Enterprise GIS” organizational implementation and 6 cities are still in a “Departmental GIS” mode. While many cities have migrated to the “Enterprise GIS” organizational structure, there are different types of enterprises. Of the 13 cities that have an “Enterprise GIS” implementation, 3 cities have chosen to implement a very centralized enterprise structure. The City of Birmingham has one such structure whereby GIS is a highly centralized operation - all editing, update and data capture are done by a Central GIS body which consists of 8 staff people (technicians, programmers and analysts). Data queries and access are provided to several departments using customized ArcView applications created and maintained by the centralized GIS. Similar is the case with Oklahoma City and City of Wichita. The other 10 cities have typical “Enterprise Systems” where there is a central repository of GIS data and a central GIS division in some department that coordinates the GIS efforts over the entire enterprise. However, individual departments maintain custody of datasets that they create or datasets that are pertinent to them. The cities of Atlanta, Honolulu, Jacksonville, Las Vegas, Long Beach, Milwaukee, Minneapolis, Portland, San Antonio, and Seattle are cities that have implemented this type of enterprise solution.

There are 6 cities that are still in a “Departmental GIS” structure and have no centralized data repository or standards for GIS development. These cities include Colorado Springs, Miami, Newark, Omaha, Pittsburgh and Santa Ana. In these cities various departments create their own datasets. Although they may adhere to some standards that are set for GIS implementation, such standards are not well established or policed and data sharing is not easy between the departments. For example, the City of Pittsburgh has no central GIS - there is a central Computer Information Systems Department (CIS) which maintains computers, servers, access, and networks but not GIS data or standards. Many departments in the City of Pittsburgh use GIS and have a steering committee. Data standards exist and everyone complies with them as best as they can. The Water and Sewer Department is outside of the City organization and there is no data sharing between the two. Most needs for GIS-based applications or for customization originate from the departments and consultants are hired to implement them. A similar structure is seen in the City of Newark and Omaha. In cities without an enterprise structure, loose and unofficial structures may exist whereby individuals know what is happening in other departments even though there are no formal structures in place. It is important to note that it was more difficult to get information on GIS in cities with a “Departmental GIS” organizational structure.

No cities have completely deployed a “Community GIS” model, although some cities do provide access to GIS data over the Internet. Of the studied cities, only 5 cities (26%) provide access to GIS data through Internet GIS browsers at the time of the survey. These cities include Honolulu, Jacksonville, Milwaukee, Portland, and Seattle. Another 9 cities (including Atlanta, Birmingham, Las Vegas, Long Beach, Minneapolis, Oklahoma, San Antonio, Santa Ana, and Wichita) had Intranets or were working on Internet applications that were to be released for public access in the near future. No web-based GIS applications were being planned for 5 cities – these include Pittsburgh, Newark, Colorado Springs, Omaha and City of Miami. Interestingly enough each of these cities has a "Departmental GIS" structure. The City of Pittsburgh and the City of Colorado Springs provide some static maps via Internet. For the cities that have interactive GIS websites, there is a variety of access levels. For example, while the City of Seattle provides access to their data through a parcel viewer, the City and County of Honolulu provides all their data for viewing through an Internet map browser and even allows users to download any of their datasets.

It is interesting to note that 11 cities have a position of GIS Coordinator/GIS Manager/GIS Director/GIS Project Manager (Birmingham, Honolulu, Jacksonville, Las Vegas, Milwaukee, Minneapolis, Oklahoma City, Portland, San Antonio, Seattle, and Wichita). All the 11 cities have either a tight centralized structure or an Enterprise GIS implementation. The cities of Atlanta and Las Vegas have an Enterprise GIS but not an official position for a GIS coordinator or Manager. However, in the above cities, the person interviewed served as a coordinator/manager as in the other 11 cities even though there was no formal position. In Atlanta the functions of a GIS manager are performed by an IT Engineer position in the Public Works Department. In Las Vegas a Senior GIS Analyst/QA QC position in the Information Technology Department coordinates GIS activities throughout the enterprise. The organizational characteristics of the sample are summarized in Table 3.1. Table 1 in Appendix B2 provides a summary of the organizational characteristics for every city surveyed.

Table 3.1 Organizational Characteristics in Surveyed Cities

Type of Organization	**Number of Cities**
Enterprise	13
Department	6
Community GIS	
Internet	5
Intranet	9
Neither	5
GIS Coordinator Position	**11**
Department Housing Enterprise GIS	
Information Technology	9
Public Utilities	2
Planning & Permitting	2
Regional GIS Consortium	**9**
Cities Responsible for Emergency Management	**11**

(sample size = 19)

The centralized GIS activities (or the Enterprise GIS) have been housed in a variety of departments in different cities. Of the 13 cities that have an Enterprise GIS, a majority (9 cities including Jacksonville, Las Vegas, Long Beach, Milwaukee, Minneapolis, Oklahoma, Portland, San Antonio, and Wichita) have their GIS operations in the Information Technology Department. Two (Atlanta and Seattle) have their GIS in Public Works or Public Utilities and 2 cities (Birmingham and Honolulu) are in the Planning and Permitting Department. Information Technology Department appears to be the most common because of the technological nature of GIS involving expertise on servers, networks, and programming and because of the need for a highly technical workforce. Public Works and Public Utilities or Planning and Permitting Departments may be suitable choices where such departments have historically been the champions of GIS in the organization and creators of some of the biggest datasets in the organization.

Only 9 cities reported that there was a consortium/group that coordinated the development of geographic data in the region and only 1 out of the 9 cities did not participate in this consortium. A total of 6 cities out of these nine have an Enterprise GIS organizational structure. In other words, 50% of the cities that have Enterprise GIS participate in a consortium that coordinates geographic data. On the other hand, only 30% of the cities that have Departmental GIS participate in such consortiums. When asked to list the top 3 geographic data coordinating groups, respondents provided a range of organizations such as regional commissions, county agencies, metro agencies, state agencies, local chapter of URISA (Urban and Regional Information Systems Association) and one project-based inter-local agreement to acquire digital orthoimagery.

It is obvious that an Enterprise GIS implementation is very suitable for the use of GIS for disaster management. GIS can be used most effectively in all phases of disaster management if all the data in the organization resides in a central server and in a standardized format. Furthermore, the core GIS team can be the resource to manage the deployment of the technology in the response and recovery phases, a key need in the aftermath of a disaster (Greene 2002). This team is aware of datasets from all over the enterprise and can leverage the use of all the technologies. The GIS Manager/Coordinator can serve as the coordinator and implementer of all data sharing policies and software. The proper agreements for data sharing and data coordination are very important for disaster management since disasters obviously know no geographical boundaries and often require analysis beyond jurisdictional boundaries. Such agreements can be negotiated much better through Enterprise organizational structures rather than through individual departments in an enterprise.

However, a tight centralized structure may be detrimental to the effective use of the technology since the emergency managers may not be fully aware of the power of the technology and what it can do for emergency management. Furthermore, it is important for the first responders to know the usefulness of any technology. For example, in the case of Hurricane Andrew in Miami-Dade County and World Trade Center bombings, as the first responders became aware of the kind of the questions that they could get GIS to help them, the needs for maps and analysis emerged (Dymon 1993). A less-centralized enterprise structure can “enable” users by providing the right organizational environment where they wouldn’t have to understand the intricacies of GIS development but could focus on applications and uses pertaining to their business needs. Providing easy access to data enables the government to provide information to citizens about the response and recovery. Furthermore, it provides citizen groups and other non-profit organizations access to data and tools that can help them in analyzing issues of equity, response and long-term recovery.

3.23 Availability and Use of Core Datasets

As mentioned earlier, the management of disasters requires knowledge from many different disciplines. Hence there are many different datasets that are needed for any meaningful use of GIS for disaster management. While some datasets are pertinent only to certain hazards (e.g. earthquake potential maps, flood boundaries, or hurricane zones, etc.), there are some core datasets that are essential for a multi-hazard approach to disaster management and particularly for disaster damage assessment. These include parcels, road centerlines, etc and are created and maintained by most municipal and governmental agencies. The following discussion lists these datasets, their use in disaster management in general and damage assessment in particular, and the state of development of these datasets in the surveyed cities. This information is also summarized for the full sample in Table 3.2 and presented for each city in Table 2 in Appendix B2.

Cadastral (Land Parcel) Data and Related Attributes

Most municipal or other local governments responsible for collecting taxes keep records of every parcel of land that is taxed in the jurisdiction. This includes a graphical representation of a parcel and related information in a database or spreadsheet format. The related information includes legal description, and information on ownership, address, assessed value, sales transactions, and number of rooms, type of construction, height and square footage. Parcel data were historically in paper format, were later converted to CAD (computer aided design) format and now converted to GIS format. The conversion to a GIS format allowed the parcel fabric to become intelligent by joining to other databases such as tax assessment and building permit data, which are commonly maintained by all taxing agencies. This provides the power of spatial analysis and data queries that were never possible before.

When combined with tax assessment data, parcel data is one of the most valuable sources of data for disaster damage assessment. It is one of the only sources that provide detailed information on the use of land at the finest granularity. Furthermore, if information is recorded appropriately, it provides the most basic, accurate and detailed information on every building, which is required to assess the damage from any disaster or to engage in mitigation activities for proactive disaster planning. The use of GIS parcel data is so integral to many government functions that many states have invested in statewide parcel coverage (e.g. Oregon, Texas, New Mexico, Florida, Kansas, Nebraska, Pennsylvania, North Carolina, South Carolina, Maryland, and New York) and most urban areas have these data.

As expected, this research indicates that parcel data are readily available for most large cities. All cities surveyed in this research have parcel data available in digital format. While most of the cities maintain their parcels in some commonly used GIS format, the City of Newark is the only city that uses Microstation (primarily a CAD product) for maintaining their parcels. A total of 15 cities use ArcView, ArcInfo, or ArcGIS for their parcel creation and maintenance (i.e. an ESRI product, which is one of the largest GIS vendor and the technology on which HAZUS is based). Only 2 cities also have applications or use MapInfo products, 4 use AutoCAD, and 4 cities use Microstation for parcels. In addition, the City of Minneapolis uses Oracle Spatial, and the City of Portland deployed SDE (Spatial Data Engine) geodatabases.

The time at which parcel data were first automated (digitized) varies from city to city – in 5 cities before 1990, in 6 cities between 1991 and 1995, and 4 cities between 1996 and 2000, and in 2 cities post 2000. The respondents for 2 cities did not know the vintage of their parcel data. A majority of cities do regular updates of parcel data with 16 cities updating their data at least once within a year of the survey. A total of 11 of these cities update their data daily, weekly or monthly. Only 2 cities had never updated their parcel data which reflected a snapshot from the time of automation (1 city with data as old as 1990) and 1 city has data that is more than a year old (18 months old).

In all the surveyed cities, parcel data were created and maintained at the local level (municipal or county). Of the 19 cities surveyed, 9 cities created and maintained their own parcel data, 2 cities didn't create their own parcel fabric but were responsible for maintenance, 1 city created its parcel data but the County took over maintenance. For 11 cities, the County was responsible for creation and maintenance of parcel data.

One of the key elements of the use of parcel data is the ability to link these data to various databases that are maintained by many local government departments such as assessment data, permit data and derivative datasets such as land use and zoning. These provide intelligence to the parcel data which is critical for disaster management and damage assessment. For example, to assess the level of damage from flooding, it may be crucial to analyze how many houses have basements, and what grade they are at. Similarly, to assess the level of damage from a forest fire or to predict the spread to fire, it is important to know the roofing material of houses and also land use. Of the 19 cities studied, 16 cities have all tax assessment data linked to the parcels while 3 cities have partial tax assessment data. Only 3 cities have photographs of every building and 6 cities have collected custom building information. A total of 14 cities have collected custom land use (different from that used in tax assessment data) and 13 cities have collected custom zoning information. Permit data are available in digital format for 7 cities and for 2 cities partial permit data are available (i.e. from the year that this data was collected digitally).

Parcel data have been widely used for disaster management purposes in most of the surveyed cities. Of the researched cities, only Atlanta, Las Vegas, Pittsburgh and Santa Ana reported that parcel data had never been used by the organization for disaster management purposes to the best of the respondent's knowledge. Most cities reported a range of uses for this data for disaster management - from post-disaster damage assessment (Honolulu, Long Beach, Miami, Portland, Seattle, Wichita), to evacuation (floods and snow in Birmingham, hurricane in Jacksonville), disaster response management (Omaha, Portland, San Antonio), floodplain management in Minneapolis, and Y2K response in Milwaukee. However, only a few cities had used parcel data and other GIS data in the pre-disaster planning and hazard identification purposes. The cities of Seattle, Long Beach and Portland are worthy of mention - Seattle had used parcel data for landslides and earthquake mapping, Long Beach for pre-disaster damage assessment, and Portland for what-if scenarios.

Building Footprints

Many jurisdictions have invested in developing a building footprint layer either approximately (through heads-up digitizing) or very accurately (through photogrammetric feature extraction). The building footprint dataset can provide more detailed information about buildings in a parcel and can be very useful for the purposes of response. Information regarding entrances, exits, fire escapes, elevators and stairwells, roof layout, material and type of construction, year built, number of floors, number of occupants and use of every floor, location of water mains and other data such as floor plans, and photographs can be tracked in detail. Cities such as Chicago (not part of this survey) are now requiring owners of large buildings to submit detailed floor plans of every floor and linking them to GIS for disaster management purposes (City of Chicago 2002). This dataset can also be particularly useful for parcels that have multiple buildings in a single parcel as is seen in the case of residential and commercial condominiums, institutional campuses, research and industrial parks, etc. In large cities most of the above cases are frequent and the information at a parcel level may not be sufficient to understand the impact of a disaster and accurately estimate the damage from a disaster.

While digital parcel data were largely available for all cities, the same is not the case for the building footprint data. Only 12 of the researched cities had a building footprint layer that encompassed the entire jurisdiction whereas 3 cities had building footprints digitized for either a small part of the City (e.g. CBDs in Honolulu and San Antonio) or only certain buildings (e.g. Miami). The building footprint data are planned for the City of Jacksonville whereas no plans are in place for the acquisition of this data layer for the cities of Omaha, Santa Ana and Wichita.

These data are not deemed crucial for city management purpose as they do not directly serve tax assessment purpose and have not been available historically in a paper format. However, many cities use Sanborn Maps, which were created for large cities by a private company called Sanborn Map Corporation. However, information on the use of such maps was not sought in the questionnaire. Regardless, the building footprint dataset is more difficult to develop in the absence of any historical paper maps that can be automated. Accurate planimetric data can be captured through photogrammetric feature extraction but this requires high skill and a large financial investment. Heads up digitizing from ortho-rectified aerial imagery can develop a very crude layer for building footprints as has been done in the case of Portland. Another method is the one deployed in Las Vegas where good spatial representation is captured for all buildings except residential buildings, which are represented by a single point. This significantly reduces the level of effort and yet yields a product that can be useful.

For the 15 cities that had building footprints (full or partial), only 4 cities undertook regular updates –Birmingham and Las Vegas did monthly updates whereas Milwaukee and Colorado Springs did daily updates. In all other cities these dataset represent a snapshot in time. The Cities of Atlanta and Portland are currently in the process of updating their building footprint layer and Newark and Pittsburgh had done some updates to this dataset since its creation. This dataset was created in the nineties in 9 cities whereas it was created in the eighties in 3 cities.

As far as attributes are concerned, few cities tracked any particular attributes about the buildings and for most of the cities the only data that can be attributed to the buildings is the assessors' database, either through joining a parcel id number or through spatial join. This presents limitations in parcels that have multiple buildings in them or buildings that span multiple parcels. Of the cities that tracked other information, Atlanta kept permit information, Birmingham Planning Department attributed building use information through a survey, and Las Vegas did a similar survey to attribute information regarding building characteristics such as capacity and year built. The City of Portland is notable for having a photo of every building, and had collected custom information on use, building characteristics and zoning. Seattle had information on the elevation of every building and structure type (such as building, deck, etc), that was collected as part of the compilation of building footprints. The City of Colorado Springs tracked addresses and multiple addresses for all buildings. Honolulu tracked custom building characteristics, land use, and permit information even though the building footprint layer only covers less than one third of the extent of the jurisdiction.

Of the 12 cities that had building footprint data, 7 had used this dataset for disaster management purposes and 4 had never used it for the same, and one respondent was not aware whether it was used for disaster management or not. The cities that had used this data for disaster management included Birmingham, Colorado Springs, Long Beach, Newark, Oklahoma City, Portland and Seattle. The use ranged from proactive disaster management planning in Colorado Springs (planning and simulations), Long Beach (disaster damage assessment using a proprietary tool called VRISK), and Newark (mapping for Office of Emergency Management) to post-disaster response in Oklahoma City (1995 Oklahoma Building bombing and 1999 tornado response), Portland (integration of 3D CAD drawings to assess flood damage), and Seattle (maps for Nisqually Earthquake response).

Streets

Like parcels, datasets on roads/streets are commonly used enterprise datasets and are available for most GIS programs, even ones that are not very advanced. The street centerline dataset is one of the foundational datasets that is developed and other datasets are based on it. Address ranges are usually available with this dataset.

As expected all the studied cities had street centerline data available although the City of Newark and City of Minneapolis reported the street centerline data as a work in progress, expected to be completed by 2002. For all the cities, the extent of the dataset was the jurisdiction boundary and two cities reported that they created and maintained the street centerlines for an area beyond their jurisdiction. In 3 cities the street centerline data were developed prior to 1990, in 6 cities these data were completed between 1990 and 1995, and in 5 cities these data were developed post 1995. Respondents in 5 cities were unaware of the vintage of these data. The degree of updates to these data is also variable – 10 cities maintain these data regularly (i.e. daily, biweekly or monthly). Others have a periodic or intermittent update strategy. All cities carry address ranges with their street centerline data.

With respect to the use of this dataset for disaster management, the research points to widespread use of the street centerline dataset with 14 cities (74%) reporting the use of this dataset for disaster management purposes. However, much of that use is related to E911 activities (and Computer Aided Dispatch) with 7 cities (37%) describing their use of street centerline data for E911 dispatch only. The rest reported uses other than or in addition to E911 activities. Many of those uses relate to disaster response such as hurricane evacuation (Jacksonville), tornado response (Oklahoma), emergency routes and evacuation (Portland) and post-flood response (San Antonio). Only a handful of cities had used this dataset for proactive planning for disasters. The City of Wichita had undertaken tornado frequency analysis with this dataset, Seattle had used it for earthquake and landslide mapping, and Long Beach had used it for VRISK.

Ortho Imagery

Aerial imagery in the form of paper photographs has traditionally been used for city planning and management. In recent years these paper photographs have been replaced by digital, low-altitude, high-resolution imagery taken from cameras in aircraft. The images are referenced to the earth and corrected for displacement due to tip and tilt of the camera, and the uneven surface of the earth. Such imagery has now come to be a major resource for any GIS operation. It provides accurate location of various features on the earth along with a clear picture of the location of natural features such as rivers, lakes, shorelines, forests, and wetlands, and human-created features such as roads, buildings, parking lots, etc. It is a powerful dataset for visualization, particularly for politicians, decision-makers, planners and public to understand the impact of a disaster, particularly in the post-disaster phase. Imagery taken when the disaster is in progress or immediately afterwards can also provide very accurate estimates of the extent of the disaster and the losses associated with it. In recent events in New York City and other floods and hurricanes, the media used ortho imagery acquired from airplanes or satellites to convey the extent and magnitude of damage to the public.

As expected, this research shows that all cities had some form of ortho imagery associated with their GIS programs. All surveyed cities had imagery for the whole extent of the jurisdiction except for Birmingham and Honolulu, which had imagery for less than a third of the jurisdiction. While this was not completely unexpected for Honolulu since the unit of analysis was the whole county, the revelation was somewhat perplexing for Birmingham, which seemed to have a very strong GIS program. The GIS Manager attributed the lack of orthoimagery to lack of funding. For most cities, the available digital orthoimagery represented a recent snapshot of the cities with imagery in 15 cities acquired between 1998 and 2001. However, the imagery in some cities dated back to mid-1990s. Las Vegas and Portland reported that they had a program of acquiring imagery every year. In order to have updated imagery every year, Las Vegas has decided to compromise on the ortho-rectification of the imagery since their use revealed that a georeferenced aerial imagery served their purpose well and could be acquired for a much lower price. The City of Minneapolis reported a program to acquire imagery for one half of the city every year.

An interesting aspect of ortho-imagery is the issue of licensing – 4 cities reported that they licensed their imagery from a private vendor, which severely restricted the use of the imagery. This can have a significant impact on the use of the imagery in the response phase of disaster management when many different organizations come together to manage the event and most would require aerial imagery. Particularly, this imposes severe restrictions on the use of the imagery by federal, state, or other local and regional governments that may be responding to the same disaster.

With respect to the use of imagery for disaster management, more than 50% of the cities reported some kind of use – from cartography and visualization (Birmingham, Portland and Wichita) to use in planning (Honolulu), and response (Oklahoma, Seattle and Wichita). Most of the uses pertained to visualization or to analyze before and after situations due to a disaster.

Topography

Information on topography is crucial for many engineering applications and work performed by many departments of the local government. Slope can play a critical role in determining the feasibility of new construction, erosion control, degree of flooding, run-off, landslide potential, and soil dynamics, etc. Elevation data is acquired in the form of digital elevation models (DEM), digital terrain models (DTM), or contour lines through photogrammetric compilation.

Given the widespread use of digital topographic data for local government functions, it is hardly surprising that only 2 cities (Miami and Santa Ana) do not have topography data. Although the City of Miami (via the Miami County) was planning on acquiring LiDAR data, Santa Ana had no plans to acquire any kind of topography data in the near future. All the other cities had some elevation data even though the spatial extent of this dataset was variable – 11 cities had data for the full extent of the jurisdiction whereas 5 cities had these data for less than two-third of the extent of the jurisdiction. While for the most part this dataset was relatively new in most cities, in 4 cities these data were created prior to 1990, in 4 cities these data were created between 1990 and 1995 and in 8 cities after 1995.

Only 7 cities reported the use of these data for disaster purposes. Of these, Long Beach used these data to report inaccuracies in FEMA'S Flood Insurance Rate Maps (FIRMs) and Oklahoma City used these data for flood mitigation. Portland used the data for planning and permitting, flood modeling, and slide potential studies. In Minneapolis, the Sewer Division under Public Works used the topography data for similar purposes.

Other Data

Although the focus of this research was core data, information on other data was solicited from responders. These datasets included transportation, utilities such as water, wastewater, and communication networks. Information was also requested on the availability of datasets related to location of schools, sites with hazardous materials, and location of police, fire and emergency response facilities. However, very basic information on these datasets was requested and an in-depth analysis was beyond the scope of this research.

Some respondents (4 cities) reported lack of knowledge about data on utilities and pointed to other knowledgeable people for further information. This could be because of the fact that utilities' data is not one of the core datasets and is usually in departments that have their own GIS. However, many respondents were knowledgeable about the utilities data. A total of 12 cities had complete water network data, and 3 cities reported that these data were a work in progress. On the other hand 10 cities had wastewater coverage, in 3 cities these data were work in progress and 2 cities reportied there were no plans to develop these data. There was less data pertaining to communication networks with only 3 cities having communication data and 2 of these had only data for city-owned infrastructure, and not private communication networks.

Only 6 cities reported using these data for disaster management purposes ranging from planning to mapping and responding to disasters. The City of Long Beach had all fire hydrants mapped and available in fire trucks and stations through mobile mapping. The City of Portland used water network data in various capacities in mock drills. None of the interviewed officials discussed the use of these data in the protection of their water and utility infrastructure. Neither was the issue of acquiring these data from private utility companies in times of disaster or for proactive disaster management raised by any of the respondents. The questionnaire did not explicitly address such issues.

With respect to point data such as location of schools, hazardous waste sites, and emergency response facilities, 9 cities had mapped sites with hazardous wastes, and 6 cities did not have this mapped. Respondents from 3 cities didn't know about this dataset and only 1 city reported that this dataset was under construction. The statistics are slightly better for educational and emergency critical facilities such as police, fire and EOCs. A total of 14 cities had mapped these facilities, 3 cities did not and 2 didn't know.

Table 3.2 summarizes information about the above datasets and Table 3.3 summarizes the use of various datasets for disaster management purposes in the sample cities.

Table 3.2: Availability of Various Datasets in Surveyed Cities

Name of Dataset	**Developed (# of cities)**	**Planned/ Under development**	**Not available/ no plans**	**Don't know**
Parcel + Tax Attributes	19			
Building Footprints	12- Full Extent 3 – Partial Extent	1	3	
Street Centerline	17	2		
Ortho Imagery	17– Full Extent 2– Partial Extent			
Topography	11– Full Extent 5– Partial Extent	1	1	
Water Network	12	3		4
Wastewater	10	3	2	4
Communication	3			
Hazardous Waste Sites	9	1	6	3
Educational	14		3	2
Critical Facilities	14		3	2

Table 3.3: Use of Dataset for Disaster Management in Surveyed Cities

Name of Dataset	# of cities that have used dataset for disaster management	Types of Uses
Parcel + Tax Attributes	15	Post-disaster damage assessment, evacuation, disaster response management, floodplain management, Y2K response. Isolated use for pre-disaster damage assessment, what-if scenarios and hazard identification.
Building Footprints	7	Post disaster response (Oklahoma City bombing, flood damage assessment, and earthquake response) and pre disaster planning (simulations, damage assessment)
Street Centerline	14	Primarily for E911 dispatch. Other uses include emergency routes and evacuation, post disaster response (tornado and flood). Handful of uses for proactive planning (tornado frequency analysis, earthquake and landslide mapping, and predisaster damagae assessment.
Ortho Imagery	10	Cartography and visualization. Some planning and response
Topography	7	Validation of flood insurance rate map, flood mitigation, planning and permitting and flood modeling
Utilities	6	Fire hydrants mapped and available through mobile computing in fire trucks, mock drills

Metadata

An important aspect of creating data is to document information about the data, commonly known as metadata. Metadata for geographic data includes information about the extent of a dataset, projection, data type, and data dictionary for tabular data, currency, keywords, and information about the creator and their contact information and many other pieces of information regarding the lineage of the data. Good metadata is key to appropriate use of datasets not only between agencies but also within an agency. Many standards exist for creating and publishing metadata such as FGDC (Federal Geographic Data Consortium), ISO, etc. and most GIS software now have tools that automatically capture a portion of the information. Capturing the rest of the information can be a lengthy and cumbersome process and hence is often ignored by many agencies.

Recognizing the importance of metadata, many agencies developed their own methods (prior to the federal and other standards) for documenting their data in the form of “homegrown” metadata.

This research reveals that 15 cities had some metadata. However only 5 of these cities had metadata that complied with standards such as FGDC, and the rest 10 cities had “homegrown” metadata. Half of these cities with “homegrown” metadata planned to convert their metadata to some standard metadata format. There was no metadata available for 4 cities and 3 out of these four cities were cities with a Departmental GIS organizational structure.

Data Cost

An important aspect of data sharing is the issue of whether to charge for the data or share the data free of charge with other agencies and the public. Although the question was posed regarding this for every dataset, most respondents gave the same answer for most datasets. Furthermore, many respondents were not clear about their pricing structure and either didn’t have a structure in place or were in the process of creating one. Of the 19 cities that were surveyed, 11 cities sold their data based on various cost models and pricing structures. Many of the cities that sold data did not actually charge other governmental agencies. In addition, 2 cities, charged only for the labor cost and cost of media to provide data to outside agencies. A total of 6 cities did not sell their data. For the cities that did not sell their data, it is not clear whether the data was distributed freely or whether the data was not easy to acquire.

3.3 Implications of Findings on Use of GIS for Disaster Management

This research clearly shows that GIS is widely diffused at the local government level of large cities. Most core layers such as parcels, tax assessment information (which includes a wide range of information very crucial to disaster management and damage assessment), street centerline, topography, and orthoimagery are available for a majority of the cities. All the above data are critical for use in disaster management and particularly to support local level decision-making and the needs of the first responders. However, information on building footprints, utilities, and critical facilities is still not as widely diffused and efforts need to be made to acquire these datasets for large cities. However, such efforts should either be at the local level or at least planned in conjunction with local agencies. The State and Federal government should be involved in setting standards for these data and often funding their collection through grants and cost-sharing rather than collecting these data themselves as has been proposed by many federal programs. Local participation will ensure that the data is maintained, updated and verified for currency and accuracy. Furthermore, it is evident that most cities have fairly advanced GIS programs and hence are well prepared to gather such data. Where cities do not have a good GIS program, assistance from the State and Federal sources can provide the impetus in establishing such programs so that they are sustainable in the long term. Nevertheless, any development of tools and models, or acquisition of data must take into consideration the programs that exist at the local level and the presence of some key datasets at most local levels of large cities. In the absence of this, tools and models will not serve the needs of the local emergency managers and will lead to early rejection rather than diffusion and adoption.

This research also shows that although most cities have advanced GIS programs, the use of GIS for disaster management has been limited. The use of GIS has been mostly in the response and recovery phase, after a disaster has struck. Although some examples exist demonstrating the use of GIS for planning, simulations, and proactive analysis to understand the likely impacts of hazards, such applications are few and restricted to a handful of cities. This could be because of the low priority given to planning for unlikely catastrophic events or the lack of any imminent threat from such events. However, with the heightened awareness of the possibility of such threats as well as the knowledge of the use of GIS in managing disasters, this is likely to change. Many GIS programs are now an integral part of emergency management agencies in local government and also actively participate in any EOC (Emergency Operations Center) activation. However, for the most appropriate use of the technology, it is important that GIS is used not only to respond, but to undertake hazard assessment and vulnerability assessment, to identify people and infrastructure at risk, prepare and test emergency plans, and mitigate against losses from impending disasters. To do this, it is essential to assess the needs for GIS for emergency/disaster management and make sure that the GIS implementation, both organizationally and with respect to appropriate data availability, can serve these needs.

Some organizational changes may be necessary, but if needs are assessed appropriately, such changes may be beneficial to the entire organization and not just for disaster management. For example, the needs of disaster management may suggest an

Enterprise GIS organizational structure rather than a Departmental structure but the benefits of this can be realized for the entire organization. It is important to note here that the findings of this research do not demonstrate that cities with Enterprise GIS organizational structure are more likely to use GIS for disaster management. On the contrary, some cities with a Departmental structure (such as Colorado Springs, and Newark), have used GIS for disaster management more than some cities with Enterprise GIS. A complex set of factors and their interplay determine the effective use of GIS for disaster management, a topic that needs further exploration. However, an Enterprise GIS structure tends to remove a lot of obstacles for the use of GIS in disaster management by standardizing data development (with respect to format, projection, accuracy and integration), facilitating data access through central servers or through internet, facilitating he establishment of data sharing policies, and ensuring the maintenance of GIS expertise in the organization.

An important aspect of using GIS for disaster management is the issue of sharing data across jurisdictional boundaries and the establishment of policies and standards that will promote such data sharing. This research points to the lack of regional and interlocal agreements for GIS use for disaster management. Such agreements have to be put in place before the occurrence of a disaster and should include agreements for data sharing with private utility companies in the aftermath of a disaster. Of course, such data sharing will require good metadata, another area where a lot needs to be done. Finally, the power of using GIS on the Internet to inform the public about any impending disaster, and events as they unfold is something that is yet to be realized in most places. The balance between informing the public and compromising information related to security presents challenges, but must be achieved for the greater good of society.

3.4 Selection of Case Studies

Respondents from 13 cities out of 19 said they were willing to share their data in exchange for result of analyses of an earthquake event using HAZUS™. Respondents from 6 cities were not so sure about this, 4 because an earthquake analysis provided them little value since there were no earthquakes in their region. Based on the findings of this survey, two cities were identified as the case cities for further research: City of Seattle, WA and City of Long Beach, CA. Both cities had an imminent threat from an earthquake and were very proactive in managing disasters. Both cities had the datasets needed for further analysis with HAZUS™ and were willing to share these data for this research. The City of Portland was also considered as a candidate for further research. However, the City of Portland was used for the initial validation of the HAZUS™ model and hence it was not considered a suitable candidate.

While the City of Seattle had a very advanced GIS organization, the City of Long Beach had a GIS implementation that was more "typical" and hence was important to analyze for the purposes of understanding the challenges associated with both types of cities. Other cities such as Atlanta, Birmingham, Las Vegas, Milwaukee, Minneapolis, Omaha, Wichita, Pittsburg, and Jacksonville were not suitable candidates for an earthquake analysis since and earthquakes were not an imminent threat for these cities and coordinators from some of these cities were not too keen on sharing data for such analysis.

The next two chapters (Chapter 4 and Chapter 5) discuss in details the GIS in the two selected case cities, the challenges associated with inputting local data into HAZUS™, the deviation of local data from default data, and the difference in loss estimates based on the local data and default data. Inferences will be made based on each of the case studies and then the results from the two case studies are compared in Chapter 6 to analyze any trends in the differences, both in building inventory and in losses from earthquakes of various magnitudes.

Chapter 4: City of Seattle Case Study

4.0 Introduction

The City of Seattle was chosen as a case study because of the highly evolved GIS program in the City, the willingness to share data and the imminent threat of earthquakes to the City. This chapter will present the findings from the Seattle case study. In this section, a general introduction of the City is provided. This is followed by an overview of GIS in the City and the GIS data provided by the City for this research. A detailed discussion of using this data and preparing it for HAZUS™ follows. This discussion focuses on data availability, completeness and accuracy and on suitability for input into HAZUS™. The next section discusses the variation of local-level data from data available in HAZUS™ - the section will explore the variation at the city-level as well as at the level of the various census tracts in the City of Seattle. The results from running various scenarios of earthquakes will then be analyzed – again looking at the variation of results using local-level data as compared to default data in HAZUS™ at the scale of the city and at the census tract level. Finally, the chapter will summarize the findings from this case study.

The City of Seattle is located in the State of Washington, about 113 miles from the US-Canada border on the northwestern corner of United States. The city was founded in 1869 and is a thriving center of arts, culture, commerce and technology for the U.S. Pacific. It has an area of 84 square miles, a population of 563,374 and a population density of 6,715 (2000 US Census). It is the 23rd largest city in the US based on population. The city's population was 530,844 in 1970, and declined to 493,846 in 1980 and then has been steadily increasing to 516,259 in 1990 and 563,374 in 2000. As per the 2000 US census, about 70% of the population is white, 13% is Asian, 8.4% is African American, and 1% is American Indian and Alaska native.

Seattle ranks as one of the best US cities to locate a business (Seattle Datasheet 2005). Consequently, Seattle and the greater Seattle region have emerged as a leader for aerospace, computer software, bioinformatics, electronics, telemedicine, etc. (Seattle Datasheet 2005). It is the corporate house of Boeing, Costco, Microsoft, Weyerhaeuser, Washington Mutual and other corporations. The city also boasts of a very high quality of life with a number of parks and recreational venues such as stadiums, arboretum, theaters, aquarium, and waterfront activities (Map 4.1). Downtown Seattle is not only home to a number of offices, but 4% of the city's total population also lives downtown.

The City of Seattle has been affected by many earthquakes – 9 earthquakes of magnitude between 5.0 and 6.5 on the Richter scale have their epicenter within a 35 mile radius of the City of Seattle between 1932 and 1965. More recently, on February 28, 2001, the Nisqually earthquake (magnitude 6.8 on the richter scale) struck the region (with epicenter 35.7 miles SSW of Seattle) and affected the City of Seattle albeit only minimally since the epicenter had a depth of 32.5 miles. Other hazards affecting the city include landslides and volcanoes.

4.1 GIS Description and Data Quality

The City of Seattle has an advanced GIS program under the Information Technology Division of the Seattle Public Utilities. The respondent for the survey was the Corporate Data and GIS Manager of the Seattle Public Utilities who leads the GIS development for the City of Seattle. At the time of the survey there were 25 GIS professionals working in the GIS program (at the Seattle Public Utilities), 15 in the Seattle City Light, 3 in Parks, 4 in Design, Construction and Land Use, 2 in Strategic Planning, 3 in Police, 3 in Fire and 2 in Fleets and Facilities Department in the City. With a true Enterprise GIS organizational implementation, many of the core GIS activities (such as creation and maintenance of base GIS data) takes place in the Information Technology Division of the Seattle Public Utilities. Many departments in the city administration also have key GIS personnel who are responsible for maintaining their own data derived from the base layers. There are also a large number of people throughout the city who are viewers and users of the data through GIS browsers or Internet browsers.

The GIS group is responsible for the creation/procurement and maintenance of many core enterprise datasets such as cadastral (parcel data), street centerlines, building footprints, ortho photography and digital terrain data. They are also responsible for the creation and maintenance of utility data such as sewer and drainage infrastructure, and water distribution systems. The tax assessment and attribute data is created and maintained by the King’s County Department of Assessment. It is distributed to the City departments by the GIS group. The City also distributes its GIS data and map products through a central distribution location called the Seattle Public Utilities GIS Map Counter (SPU GIS Map Counter). The City has well-defined distribution policies and costs. The city sells predefined and prepackaged data through the GIS Map Counter. The required GIS data were made available free of charge for research purpose only by the Corporate Data and GIS Manager. Various datasets were made available: parcels, building footprints, orthoimagery, census tracts, city boundary, contour data and location of schools, hospitals, etc.

Parcel data for the City of Seattle (approximately 203,000 in number encompassing 123 census tracts) are created and maintained by the Seattle Public Utilities whereas the assessment data are maintained by the King's County Department of Assessments. Both are updated on a daily basis although for publication purposes, the data have a time lag. Since prepackaged data were made available by the City of Seattle, the GIS data dated to October 2000 whereas the assessor's data were from 2001. Hence there were some inconsistencies in the data but most of these were confined to differences in the owner of the property or a few parcels with no match to attribute data from assessor's data. This number was low (less than 1%) and hence was considered insignificant for the purpose of this research.

The City of Seattle provided a building footprint dataset that was compiled by the city in 1993 through photogrammetric techniques. This data remained static (i.e. with little updates) since 1993 and hence had little utility excepting as a source of validation. The building footprint data carried information about the peak elevations, which was useful for validation purposes. The city also provided access to a high-resolution (6-inch pixel resolution) color orthoimagery. This was acquired in 1999 and was licensed to the City of Seattle from a private data provider. This imagery also helped immensely in validation and other verification purposes. Other data provided by the City of Seattle included location of churches, bus stops, public hospitals, public libraries, police stations, schools (public and private), liquefaction zone hazard areas, surface geology, flood prone areas, street networks, street trees, etc. The City also provided metadata along with the data.

The GIS parcel data for the City of Seattle carried some core parcel attributes and 99% of the parcels had assessment data attached to them. Through various tables, the information recorded included legal description of the property, type of property, taxpayer information, address, present use and square footage of the property. Detailed information included zoning, water and sewer districts, views to different sides of the city, number of buildings on the site, current use, hazard information such as hundred year flood plain, seismic hazard, landslide hazard, steep slope hazard, and many more. Although, fields related to each of the above attributes were present, they were not always populated with data. For example, only 14 records out of approximately 203,000 records were populated for seismic hazard. Therefore, by inspecting the data dictionary or metadata, it may appear that there are a lot of data for the City of Seattle, but in reality much of the fields are not populated. This is common for other cities too, where the data dictionary and metadata lists a lot of fields but the fields are rarely populated.

Although some core attribute data were attached to the parcels data layer, much of the building-specific information was tracked in separate tables. Primarily, there were separate tables for residential buildings (with 1, 2 or 3 living units), commercial buildings (including apartments), apartment complexes, and condominiums with one-to-many relationships between the parcel data and the multiple buildings on the parcel (i.e. one parcel with multiple buildings in it). Information on building square footage, construction, age and height (number of stories) was tracked in these building tables. A master table was compiled at the building level with data required by HAZUS™: information on the square footage, height, age, type of structure, use and census tract for every building at the centroid of the parcel. Although information on the value of the building was available in the table, there were no data on the value of content in the buildings. The default HAZUS™ values for building and content exposure, updated to reflect updated square footage information was used. The rationale for this is discussed in the next section.

The master table compiled from the residential and commercial tables was found to be missing buildings from about 2269 records (including 1488 records which actually had a building from the building footprint layer and a use other than vacant or null). These records were added to the master table by assigning square footage based on the building footprint square foot multiplied by the number of stories (based on the height of the building footprint from the building footprint layer). This meant that some of these records did not contain any information on year built and type of structure. However, through this method, the square footage was improved.

The next section will discuss in detail the quality of data for all the HAZUS™ required fields and manipulation of data for input into HAZUS™.

4.2 Data Preparation for HAZUS™

As mentioned earlier, for inputting local data into HAZUS™, the following attributes were required: use of building (or occupancy), type of building (i.e. the type of construction), square footage of building, number of stories, year the building was constructed, census tract in which the building is located, and the value of the building and value of contents in the building. The quality of information regarding each of the above attributes varied considerably. Various checks and data manipulations were done to validate the quality of the data and improve it as are described below. Table 4.1 below provides a summary of the percentage of data that is populated for the essential attributes.

Table 4.1: Data Input into HAZUS™ for City of Seattle

Field Description	**Records populated**	**% populated (out of 159,355)**
Use/Occupancy Type	159,026	99.8%
Type of Structure/Building Type	157,086	98.6%
Area/Square Footage	159,352	100%
Height/Number of Stories	159,355	100%
Year Built	157,086	98.6%
Census Tract	Derived from other sources and 100% populated	
Building Value	Updated through HAZUS™ defaults at the census tract level	
Content Value	Updated through HAZUS™ defaults at the census tract level	

Use/Occupancy Type

Since the use of every building was recorded in great detail, it was not very difficult to associate the building use to the building occupancy classes in HAZUS™. Since 99.8% of the records were populated with information on the use of the buildings, this attribute was spot checked for consistency and accuracy, both of which were found to be acceptable. An important aspect of the Seattle data was that all tax-exempt buildings (such as government, educational and religious buildings) were appropriately recorded and all information about them maintained in the assessor’s database. However, records for the University of Washington parcels were missing attribute information. These parcels comprised one census tract with many buildings in a few parcels. It is unclear why this information was missing from the assessment data, even though information about other universities and educational facilities were available in the database. The data for buildings in the University of Washington parcels were used but their area and height were estimated in various ways as will be discussed later.

Although information about occupancy/use was tracked in great details in the Seattle datasets, some uses did not translate easily to HAZUS™ use descriptions. For example, uses such as “UTILITY” or transportation hubs such as “airport buildings”, “bus stations”, etc were unclear uses in the HAZUS™ occupancy classes. It is important to note, that there are no consistent schemes used to track the use of land across the country for assessment purposes. Therefore, there are many different classifications in use for this purpose.

Type of Structure/Building Type

A major challenge was in the classification of building structure type information from the assessor’s data into the 36 categories used by HAZUS™. The type of structure information was recorded differently in the residential and commercial building tables. The residential building table had a field called “Percent Brick Stone” with a value from 0 to 100. For this research, buildings with 50 percent or more brick or stone construction were classified as “Masonry” and those that were less than 50 percent brick/stone were classified as wood construction. The commercial building table had a field called “Construction Class”, which had 5 coded values for structural steel, reinforced concrete, masonry, wood frame and prefab steel. This information was complete for the entire table with only 8 records missing data. Therefore, (even though the Seattle data did not incorporate all the 36 structure types required by HAZUS™), the data were considered reliable and were associated as best as possible to the generalized classification of structure types in HAZUS™. The HAZUS™ software was allowed to use various parameters to distribute the generalized structure type information into the more specific structure types used in HAZUS™. Appendix C (Table 1) shows the structure type data in the local assessor's data and the classifications in HAZUS™ that they were mapped to.

Area/Square Footage

Square footage information was tracked in many different fields in various tables. For example in the residential table, separate fields tracked the square footage for the different floors, basement, garage, etc. In the commercial table there were two fields with information on gross square feet and net square feet. Apartment complexes carried their own average unit size which when multiplied by the total number of units, yielded the total square footage. This could also be acquired from the commercial table (since each apartment complex in the apartment table had multiple apartments in the commercial table with the corresponding square footage information). Various checks were performed on the data to make sure that the information in the various fields in different tables was consistent and all the data triangulated well.

As mentioned above, about 2269 records were missing square footage information and 1488 parcels out of these had a building footprint based on the building footprint data layer. For these records, the total area was estimated by multiplying the area of the building footprint by the number of stories (which was calculated based on the height of the building from the building footprint layer). Of course, this was approximate and subject to some uncertainty. But in the absence of any other source, this was the best way to estimate the square footage and since the number of parcels concerned was a small percentage of the total number of parcels, this method was suitable.

Height/Number of Stories

Information on the number of stories for every building was also good with very few records containing unlikely values. For example, the commercial building table comprised of 12 buildings with 99 stories, and the residential table comprised 2 buildings with 30 stories (even though this table contained information for residential buildings with up to 3 units!). All of these records were discarded as errors and since the number was very small relative to the size of the entire database, the error due to this was considered negligible. It is important to note, however, that even though the larger fallacies were easily apparent and could be trapped through automated QA/QC, the inaccuracies arising out of data that were mistyped, such as 11 instead of 1, or 21, or 31, etc. were more difficult to track and not dealt with. However, any database is expected to have some level of inconsistencies and instead of cleaning up the data completely, it was deemed more suitable to have a “typical” dataset.

As mentioned earlier, the height of some buildings were estimated based on the value in the building footprint layer and on the assumption that each floor is 10 feet high and rounded to the nearest whole number. Of course, for older buildings, the assumption of height of floors to be 10 feet could itself lead to some overestimation.

Year Built

The information on the age of the buildings was tracked fairly meticulously. There was no missing information or grossly erroneous information. The earliest year in this field was 1900 and it was assumed that all buildings built prior to 1900 were also given the same year of construction owing to limitations in database or software. Since the City was established only in 1869, this was not considered to be a huge problem.

Census Tract

Although the assessment data carried a field for census tract, this field held 1990 census tract value. The boundaries of 2000 census tracts provided by the City were used to assign census tract information to the parcel data.

Building and Content Value

While most assessment datasets carry information about the assessed value (sometimes broken up into land value and building value), it is uncommon for assessment data to carry information about the value of contents in the building (since content value is not taxed). There are no other easily available sources for content value at the local level. In previous releases of HAZUS™ (when this research was started), value information was not a required field. Both building and content value were updated based on updated square footage information derived from local data and an assumed per square feet value of building and content for different uses. In more recent releases of HAZUS™, building and content exposure data are required by the software if any other local data is to be input. If value of building and content is not input, HAZUS™ does not adjust for the improvement in square footage from the local level data. The building and content exposure is replaced by default values (which is based on default square footage information). The software provides no interface to populate these values based on content value per square foot, nor does it provide any estimates of content value per square foot.

Since the local data for Seattle were prepared before the current release of HAZUS™ and no content value was available from assessment data, it was deemed more appropriate to take value per square feet (for both building and content value) in HAZUS™ and update the values based on the improved square footage. Therefore, using the current version of HAZUS™ meant that these values had to be calculated outside of HAZUS™ and input into HAZUS™. The building value and content value per square foot for various uses was established using default data. In some cases, where data were missing, average values were used. These values were then used to estimate content value for the updated local data in database calculations performed outside of HAZUS™ user interface. The HAZUS™ development team provided guidance for this.

There were various difficulties encountered in doing this. For many census tracts, there were zero values for square feet in default data for various occupancies. This meant that per square feet values could not be calculated for these census tracts and the corresponding exposure could not be calculated for updated data. For such cases, the per square feet exposure was calculated for all non-zero census tracts in each occupancy class, and the average of this was used for the zero square footage census tracts. In HAZUS™, there are no data for the occupancy class Commercial 10 (COM10) which is parking. This meant that exposure per square feet (for building exposure, content exposure) could not be calculated for this occupancy. Therefore, for both default data and local data, the exposure for parking structures was missing.

Once the master table with all the buildings and their associated information was compiled from various sources into one table, this information was processed using the Building Inventory Tool (BIT) in the HAZUS™ software. This tool allows users to create their own Occupancy Mapping matrix and replace the default schemes with new schemes generated from local level data. This also generates new square footage information, which the user has to replace for the default square footage information in HAZUS™. Since the exposure information (i.e. content and building exposure) was not updated using the tax assessment data, these data were not updated in HAZUS™ using the BIT tool. As mentioned above, these data were updated outside of HAZUS™, and input into HAZUS™.

Once the local-level data were input into the model, scenarios for various magnitude earthquakes on the Seattle North Zone fault were simulated using default building data and local building data, keeping all other factors constant to analyze the differences in results due to the difference in building inventory. Before understanding the impact of local data on the results from HAZUS™, it is important to understand the deviation of default data in HAZUS™ from local data. This is discussed in the next section, followed by a discussion of the damage results from various scenarios.

4.3 Building Inventory Data Variation – Default vs. Local Data

This section analyzes the variation in default data in HAZUS™ as compared to the local data. The variation is first analyzed at the city level and then analyzed at the census tract level to understand the spatial variation across different parts of the city. This will help identify strategies to improve data and appropriate use of the results from the HAZUS™ model.

4.31 Variation at the City Level

A summary of variation of the default data from the local data is presented in Table 4.2 and Table 4.3. As can be seen in Table 4.2, in terms of square footage, the default data in HAZUS™ underestimates the total square footage for the entire City of Seattle by approximately 33% which is about 150 million square feet. The variation across the different occupancy classes is much larger. While the residential square footage is somewhat well estimated by HAZUS™ defaults, the percentage difference in commercial, industrial, government, agriculture and educational square footage are high.

Table 4.2: Variation in Square Footage by General Occupancy Classes in City of Seattle

Occupancy Class	Default Data (in thousand sq ft)	Local Data (in thousand sq ft)	%Difference over Default	Total Difference (in thousand sq ft)	Difference (% of total)
Residential	357,949	339,112	-5.3%	-18,837	-12.6%
Commercial	78,775	181,142	130.0%	102,368	68.5%
Industrial	11,898	23,691	99.1%	11,793	7.9%
Agriculture	602	127	-79.0%	-476	-0.3%
Religion	2,983	7,100	138.0%	4,117	2.8%
Government	1,149	10,259	792.8%	9,110	6.1%
Education	2,360	43,730	1753.0%	41,370	27.7%
Total	**455,712**	**605,160**	**32.8%**	**149,448**	**100.0%**

*Notes: The numbers here do not denote absolute percentage of difference that is attributable to each occupancy class because some numbers are negative.

The breakdown of the residential general occupancy into specific occupancies and their variation is shown in Table 4.3. Square footage for single family housing, duplex and manufactured housing are overestimated by HAZUS™ whereas triplex/quadruplex, and apartments are underestimated. It is important to note, that the residential table in the Seattle data comprised of only 1, 2 or 3 units residential structures (although some uses were 4-Plex indicating 4 units). All other residential structures were captured in the commercial and apartment tables. Although local data provided the breakup of residential occupancy into various categories based on the number of units this was also a recent change in HAZUS™ (as mentioned in Chapter 1), and local data were compiled prior to this change. Thus, all apartments were mapped to the RES3F occupancy and Table 4.3 shows some of the RES3 occupancy classes as 0 values. Also, the RES3F occupancy therefore shows an over-inflated increase. However, when all the multiple dwellings are added, HAZUS™ only underestimates multi-family housing by 4.3%.

It is important to look at both percentage difference as well as real difference since a high percentage difference on a low default value may contribute only minimally to total change. Thus, as seen in Table 4.2, HAZUS™ overestimates agriculture occupancy by 79% but that amounts to only about half a million square feet. On the other hand, a 5.3% overestimation in the residential occupancy results in the overestimation of almost 19 million square feet.

Therefore, as shown in Table 4.2, HAZUS™ significantly underestimates the square footage for commercial, industrial, government, religion and educational occupancy classes. Of the overall total change in square footage, commercial accounts for 68.5% of the total difference, and education accounts for more than 27.7% of the total difference. The underestimation in commercial occupancy class in City of Seattle may be attributable to the high density of commercial occupancy in the downtown. The underestimation in the commercial occupancy is concentrated primarily in retail, wholesale, professional and technical services, hospital, entertainment and recreation, and parking (the Dun and Bradstreet data in HAZUS™ does not include parking and hence the default shows zero values) as shown in Table 4.3. Personal and repair services, and banks are overestimated by HAZUS™.

The underestimation of the educational occupancy class in HAZUS™ default is also fairly large in terms of absolute difference and may be attributable to the concentration of educational activities in the city, including the presence of many universities and colleges. Furthermore, even the tax assessment data do not provide a complete assessment of educational facilities since many public education facilities do not pay taxes. Some of the underestimation of the education occupancy may also be attributable to overestimation of the local data based on the fact that much of the square footage for the University of Washington were estimated by approximate means as discussed in previous section. However, the local data do not make any distinction between educational facilities based on their educational levels. The breakdown of change by specific classes (Table 4.3) shows that there are no universities and colleges in Seattle based on the real data. Although some other queries could be performed to better distinguish between schools and higher education facilities, this was not undertaken for this research.

The default data in HAZUS™ also underestimates the amount of square footage for industrial land use by 99%. This amounts to 11.8 million more square feet of industrial space in the City than is estimated by the default data. This underestimation is concentrated in heavy industry, light industry and high tech industry with some compensation by overestimation in foods/drugs/chemicals and construction (Table 4.3). Similarly, the occupancy classes, government and religion are also underestimated by HAZUS™.

As mentioned in Chapter 2, the source of the data for the various occupancy classes is a significant factor in this underestimation – the HAZUS™ models estimates the square footage information for residential occupancy from Census data and for commercial, industrial and other uses, from Dun and Bradstreet data (Appendix A, Table 3). The various SIC codes that are used to pull data from Dun and Bradstreet data may not be complete, the data tracked by Dun and Bradstreet may not capture all properties, or all the square footage for the properties, and hence the underestimation. It is difficult to assess the source of discrepancy in the default data since the HAZUS™ User Manual does not discuss in detail the methodology used by Dun and Bradstreet to compile this data

Table 4.3: Variation in Square Footage by Specific Occupancy Classes in City of Seattle

HAZUS Specific Occupancy	**Description**	**Default Data** (in thousand sq ft)	**Local Data** (in thousand sq ft)	**Percent Difference over Default**	**Total Difference** (in thousand sq ft)
RES1	Single Family Dwelling	237,548	215,584	-9.2%	-21,964
RES2	Manuf. Housing	613	594	-3.1%	-19
RES3A	Duplex	14,504	12,811	-11.7%	-1,693
RES3B	Triplex / Quads	9,132	9,838	7.7%	706
RES3C	Multi-dwellings (5 to 9 units)	15,094	0	-100.0%	-15,094
RES3D	Multi-dwellings (10 to 19 units)	17,844	0	-100.0%	-17,844
RES3E	Multi-dwellings (20 to 49 units)	23,756	0	-100.0%	-23,756
RES3F	Multi-dwellings (50+ units)	22,171	82,286	271.1%	60,115
RES4	Temporary Lodging	3,184	11,185	251.3%	8,001
RES5	Institutional Dormitory	13,453	5,722	-57.5%	-7,731
RES6	Nursing Home	651	1,092	67.9%	442
COM1	Retail	15,931	30,268	90.0%	14,338
COM2	Wholesale Trade	11,014	45,251	310.8%	34,237
COM3	Personal and Repair Services	7,342	1,434	-80.5%	-5,908
COM4	Professional/Technical Services	29,758	63,506	113.4%	33,748
COM5	Banks	1,127	792	-29.7%	-335
COM6	Hospital	1,727	8,525	393.6%	6,798
COM7	Medical Office/Clinic	4,892	5,144	5.1%	252
COM8	Entertainment & Recreation	6,714	12,903	92.2%	6,189
COM9	Theaters	271	2,654	880.9%	2,384
COM10	Parking	0	10,666	Zero default value	10,666
IND1	Heavy	4,061	14,011	245.0%	9,950
IND2	Light	3,468	6,397	84.5%	2,929
IND3	Food/Drugs/Chemicals	1,634	1,160	-29.0%	-474
IND4	Metals/Minerals Processing	201	191	-4.8%	-10
IND5	High Technology	49	427	780.9%	379
IND6	Construction	2,486	1,505	-39.5%	-981
AGR1	Agriculture	602	127	-79.0%	-476
REL1	Religious	2,983	7,100	138.0%	4,117
GOV1	General Services	1,114	10,259	821.0%	9,145
GOV2	Emergency Response	35	0	-100.0%	-35
EDU1	Grade Schools	1,079	43,730	3953.2%	42,651
EDU2	Colleges/Universities	1,281	0	-100.0%	-1,281
Total		**455,712**	**605,160**	**32.8%**	**149,448**

The variation in the overall citywide building count is much less than the variation in square footage. The building count data are provided in Table 4.4. Although the square footage is grossly underestimated by default data in HAZUS™, the building count is **overestimated** by HAZUS™ default data but only by 2.7% or 4,350 buildings. Therefore, per the local data, Seattle has more square footage but in lesser number of buildings, reflecting the higher densities seen in most cities. This overestimation in building count can be explained by the method that HAZUS™ uses to estimate building count for default data. The HAZUS™ methodology is square footage based – to estimate the number of buildings in each occupancy class, HAZUS™ simply takes the square footage information and divides it by the average size of a building in that occupancy. Since the size of residential buildings is smaller and there is high proportion of residential buildings in the city, a small variation in the average size can lead to a large variation in the count. Thus, as seen in Table 4.4, the overestimation in the count of buildings is largely in the residential occupancy class. This is the reason why the total building count is overestimated in HAZUS™, even though building counts for all other occupancy classes are underestimated.

Table 4.4: Variation in Building Count for General Occupancy Classes in City of Seattle

Occupancy Class	Default Bldg. Count	Local Bldg. Count	Percent Difference over Default	Total Difference
Residential	159,548	147,469	-7.6%	-12,079
Commercial	3,590	8,707	142.5%	5,117
Industrial	260	1,416	444.6%	1,156
Agriculture	5	17	240.0%	12
Religion	121	643	431.4%	522
Government	98	371	278.6%	273
Education	24	673	2704.2%	649
Total	**163,646**	**159,296**	**-2.7%**	**-4,350**

The other occupancy classes are underestimated also because many census tracts show small square footage for certain occupancy classes. Therefore, if the amount of area for a particular occupancy is less than the average building size for that occupancy, then the building count is 0 for that census tract. Hence the building counts are underestimated to a large extent. For example, it is rather implausible that the entire City of Seattle has only 24 buildings for education (with only 2 buildings that are grade school) and 98 buildings for government occupancy classes. Even more implausible is the fact that there are only 47 buildings in the City of Seattle that are used for retail trade (COM1 occupancy) as shown in Table 4.5. The parameters for the average building size could easily be changed in previous versions of HAZUS™ but are not available through the interface in the current version of HAZUS™. Table 4.5 shows the breakup of the building counts and their variation over specific occupancies.

The distribution of square footage and building count by structure type is summarized in Table 4.6. Since detailed structure information was not available in the assessor's data to appropriately map to HAZUS™ required values (as discussed in the section above), this data cannot be used more effectively. The data input into HAZUS™ only mapped the general building types rather than the specific ones. Matrices in HAZUS™ called occupancy to building type matrices were allowed to further classify buildings into subcategories. Also, where no structure data were available, HAZUS™ was allowed to assign a structure type based on information about occupancy, height, and year built. As can be seen in the table below, HAZUS™ overestimates the number of wood structures – since a majority of wood structures are smaller residential buildings, this overestimation is at par with the overestimation of single family residential (11%

Table 4.5: Variation in Building Count for Specific Occupancy Classes in City of Seattle

Occupancy Class	Description	Default Bldg Count	Local Bldg Count	Percent Difference over Default	Total Difference
RES1	Single Family Dwelling	148,491	132,119	-11.0%	-16,372
RES2	Manuf. Housing	528	130	-75.4%	-398
RES3A	Duplex	4,458	6,477	45.3%	2,019
RES3B	Triplex / Quads	2,582	3,008	16.5%	426
RES3C	Multi-dwellings (5 to 9 units)	1,363	0	-100.0%	-1,363
RES3D	Multi-dwellings (10 to 19 units)	1,016	0	-100.0%	-1,016
RES3E	Multi-dwellings (20 to 49 units)	323	0	-100.0%	-323
RES3F	Multi-dwellings (50+ units)	250	5,194	1977.6%	4,944
RES4	Temporary Lodging	18	242	1244.4%	224
RES5	Institutional Dormitory	495	269	-45.7%	-226
RES6	Nursing Home	24	30	25.0%	6
COM1	Retail	47	2,987	6255.3%	2,940
COM2	Wholesale Trade	314	2,155	586.3%	1,841
COM3	Personal and Repair Services	642	275	-57.2%	-367
COM4	Professional/Technical Services	282	1,499	431.6%	1,217
COM5	Banks	269	102	-62.1%	-167
COM6	Hospital	28	65	132.1%	37
COM7	Medical Office/Clinic	657	282	-57.1%	-375
COM8	Entertainment & Recreation	1,335	1,082	-19.0%	-253
COM9	Theaters	16	67	318.8%	51
COM10	Parking	0	193	#DIV/0	193
IND1	Heavy	118	780	561.0%	662
IND2	Light	81	352	334.6%	271
IND3	Food/Drugs/Chemicals	26	105	303.8%	79
IND4	Metals/Minerals Processing	5	14	180.0%	9
IND5	High Technology	0	20	#DIV/0	20
IND6	Construction	30	145	383.3%	115
AGR1	Agriculture	5	17	240.0%	12
REL1	Religious	121	643	431.4%	522
GOV1	General Services	97	371	282.5%	274
GOV2	Emergency Response	1	0	-100.0%	-1
EDU1	Grade Schools	2	673	33550.0%	671
EDU2	Colleges/Universities	22	0	-100.0%	-22
Total		**163,646**	**159,296**	**-2.7%**	**-4,350**

overestimation in RES1 occupancy as seen in Table 4.5 above). The number of steel buildings is also overestimated by HAZUS™ even though the square footage of steel buildings in HAZUS™ is much lower than the real square footage of steel buildings in the City of Seattle (23.7 million sq ft as opposed to 59.8 million square ft). This points to larger building sizes of steel buildings and might account for the high density development in the downtown. A similar pattern is observed for the reinforced concrete building type. Reinforced masonry is overestimated but the variation in count of buildings is much larger than the variation in square foot information. Precast concrete is underestimated in HAZUS™ default values. A big variation is seen in unreinforced masonry which shows huge underestimation in HAZUS™, both in building count and in square footage. The data on mobile homes were not well documented in the local data to make the comparison of mobile homes very useful.

Table 4.6: Variation in Building Count by Building/Structure Type in City of Seattle

Building Type	**Default Data** (in 1000 sq ft)	**Local Data** (in 1000 sq ft)	**Total Diff.** (in 1000 sq ft)	**Percent Diff. in Sq Ft over Default**	**Default Bldg Count**	**Local Bldg Count**	**Total Diff.**	**Percent Diff. in Count**
Wood	336,792	316,840	-19,952	-5.9%	153,745	133,498	-20,247	-13.2%
Steel	23,729	59,821	36,092	152.1%	1,499	793	-706	-47.1%
Reinforced Concrete	27,747	58,420	30,673	110.5%	1,491	1,360	-131	-8.8%
Precast Concrete	18,536	50,246	31,710	171.1%	626	2,285	1,659	265.0%
Reinforced Masonry	33,236	32,880	-356	-1.1%	4,687	1,922	-2,765	-59.0%
Unreinforced Masonry	9,756	86,360	76,603	785.2%	534	19,285	18,751	3511.4 %
Mobile Homes	5,915	594	-5,321	-90.0%	1,033	130	-903	-87.4%
Total	**455,712**	**605,160**	**149,448**	**32.8%**	**163,615**	**159,273**	**-4,342**	**-2.7**

Therefore, the local data varies significantly from the default data for the City of Seattle at the level of the entire city. The variation is even larger when the total square footage for the city is analyzed by various occupancy classes. Interestingly, even though the total square footage is underestimated by HAZUS™ default data, the building count is overestimated (with much of the overestimation occurring in the RES1, single family residential occupancy class). The default data is also very different from the local data when analyzed by the type of structure information. In the next section, the spatial variation in building inventory over the various census tracts in the city are analyzed in greater details.

4.32 Variation at the Census Tract Level

The summary of distribution of variation over occupancy classes for the entire City of Seattle provides an incomplete picture. It is important to analyze the spatial distribution of the variation of data over the city in order to understand the implications of the use of HAZUS™ and other tools for local-level decision making. Map 4.2 provides a visual representation of the spatial variation in the total square footage for the City of Seattle. HAZUS™ overestimates the total area in 34% of the census tracts (42 census tracts) and underestimates the total area in the remaining 66% of the census tracts. The degree of underestimation is much larger ranging from .1% to 1385% (3000 square feet to 25.6 million square feet) whereas the overestimation ranges from 1% to 21% only (from 92 to 911 K sq ft).

The highest underestimation occurs in the case of University of Washington census tract. An effort was made to categorize the various census tracts into certain types such as Single Family Residential, Multifamily, Open Space, Mixed Use, Commercial, Institutional, Downtown, etc. based on the zoning dataset and aerial imagery provided by the City of Seattle. This helped analyze the patterns to see if the discrepancies (i.e. overestimation and underestimation) could be explained by the various types of census tracts. As would be expected, most of the HAZUS™ overestimation occurs in census tracts that are predominantly single family residential of higher densities (i.e. lot sizes of 5000 sq ft or 7200 sq ft) with little or no commercial or multifamily development or large amounts of open space such as parks. Significant underestimation of area is concentrated in downtown Seattle and its vicinity along with the census tract containing the University of Washington. HAZUS™ also does a poor job of estimating area for industrial census tracts.

The top twelve census tracts with the highest total underestimation (in real values) are highlighted in Map 4.2. These census tracts represent about 111 million square feet, which amounts to approximately 74% of the overall underestimation for the City of Seattle. Most mixed census tracts (comprising a mix of single family residential homes, multifamily homes, commercial area, and/or other uses such as institutional or commercial) are also underestimated largely by HAZUS™. Once the significantly large variations are removed from the data, the variation in square footage follows a fairly normal distribution. Therefore, a few large values in some census tracts skew this distribution significantly. It is important to note that the variation in the overestimation and underestimation is not completely and consistently explained by the type of census tract. Thus, two very similar census tract can each show underestimation and overestimation.

One of the other census tracts that show a large variation in square footage is the census tract with the University of Washington. The total area in this census tract is underestimated by 1384% (or 25.6 million square feet). As mentioned earlier, the assessment data contained no records for buildings on the University of Washington campus. Since the area and height of buildings in much of this census tract were estimated from other sources (i.e. building footprint data), this census tract is somewhat less reliable than other census tracts. Nevertheless, it is obvious that this census tract is grossly underestimated by HAZUS™ and much of this underestimation occurs in the education occupancy class (as high as 8551% over the default value).

Although the residential occupancy is fairly well estimated overall for the entire city, the default values overestimate the residential square footage in 88 census tracts and underestimate it in the remaining 36 census tracts (Map 4.3). The default overestimation occurs in most census tracts that have primarily higher density single-family residential development. There are some census tracts that deviate from this trend (such as the University of Washington census tract and some mixed census tracts that contain low-density commercial development). HAZUS™ also underestimates the residential square footage in most downtown census tracts or census tracts that have mixed development or industrial developments or ones with few housing units (such as the census tract with Discovery Park). The downtown census tracts have higher residential area because of high-rise apartments and also a concentration of hotels and transient lodging. While the underestimation in downtown census tracts is accountable to transient lodging, it is not clear why other census tracts show an underestimation or overestimation since the residential area is derived from demographic information from the US Census.

In direct contrast to the residential occupancy, the commercial occupancy is underestimated in 89 census tracts and overestimated in 35 census tracts (Map 4.3). The degree of underestimation is also much larger than the degree of overestimation. In absolute terms much of the underestimation is concentrated in the downtown census tracts and industrial and high-density commercial census tracts. In fact, the top 12 underestimated tracts account for 72 million square feet, which is 70.5% of the total underestimation of commercial occupancy in the city. Most of the overestimation occurs in census tracts that are primarily single family residential (SFR) or are SFR with open area and large parks. However, the reverse is not necessarily true – i.e. not all primarily SFR census tracts are overestimated in terms of commercial use area. Again, the patterns are not consistent based on the type of census tract.

According to local data, a third of the census tracts (43 census tracts) have no industrial buildings. However, the default data in HAZUS™ shows industrial square footage for these census tracts. The industrial square foot information is underestimated by HAZUS™ by more than half a million square feet in 7 census tracts (6 of which can be classified as Industrial). However, there are some census tracts that are industrial for which HAZUS™ overestimates the square footage. Interestingly, the extreme variation occurs in census tract with industrial uses – some are overestimated and others are underestimated (Map 4.4). There is little consistency in this and hence it is difficult to assess the source of such discrepancies.

The government occupancy is well estimated for most census tracts with the exception of few census tracts in the downtown (which have a concentration of governmental offices and facilities). Percentage difference of square feet information over default data is difficult to analyze for occupancy classes such as agriculture, religion, government and education because many of the census tracts have 0 values for default and the percentage change cannot be calculated.

The nine census tracts representing downtown Seattle are very different from the default values in HAZUS™ and show a much greater variation than the variation at the city scale. The total square footage for these census tracts as estimated by HAZUS™ is almost half of the actual square footage in Seattle based on local data. In other words, while the default square footage data in HAZUS™ is underestimated for the entire City of Seattle by 33%, for the nine downtown census tracts alone, the underestimation of total square footage is more than 98% (Table 4.6). This underestimation amounts to 47.4 million square feet of space and a large portion of it is concentrated in the commercial occupancy class – the square footage in the commercial occupancy is underestimated by 160% (about 43.1 million square feet of area). It is interesting to note that while residential occupancies are overestimated by HAZUS™ at the scale of the city, for the downtown census tracts, the residential area is underestimated. As mentioned before, this could be due to the concentration of high density housing downtown and other types of housing that are not well estimated by census data such as hotels and transient lodgings commonly found in downtowns of large cities.

The use of Dun and Bradstreet data for square footage information for other occupancies is also not a close reflection of the reality for the downtown census tracts as shown in Table 4.7. If the top 11 contributors to change in square foot in real terms are removed from the Seattle data, the percentage difference goes down from 33% to 10%. These census tracts include some of the downtown census tracts and other commercial and industrial tracts around downtown, and the University of Washington census tracts.

Table 4.7: Variation in Square Footage for Downtown Census Tracts in City of Seattle

Occupancy Class	**Default Data** (in thousand sq ft)	**Local Data** (in thousand sq ft)	**%Difference over Default**	**Total Difference** (in thousand sq ft)
Residential	16,932	19,527	15.3%	2,595
Commercial	26,894	70,002	160.3%	43,108
Industrial	1,544	1,696	9.8%	152
Agriculture	80	3	-96.3%	-77
Religion	946	539	-43.0%	-407
Government	839	3,544	322.4%	2,705
Education	793	139	-82.5%	-654
Total	**48,028**	**95,450**	**98.7%**	**47,422**

The variation in type of construction (or building type) is also largely concentrated in the downtown census tracts and the University of Washington census tract. Since the structure type information was not available in the local data for the City of Seattle for the University of Washington census tract, this census tract will not be analyzed closely for this information. Square footage of buildings with wood structure are overestimated by HAZUS™ in a majority of the downtown census tract but the square footage for steel, reinforced concrete, precast concrete and unreinforced masonry, is underestimated. However some of this underestimation is simply a function of the underestimation of square footage for the downtown census tracts and does not reflect truly a change in proportion of different types of buildings. There are also no apparent patterns of systematic changes in building types and the type of census tract that can explain the changes.

Thus in summary, the local data for the City of Seattle are very different from the default data in HAZUS™ across different types of census tracts. It is obvious that the data are significantly different for the downtown commercial census tracts and for special use census tracts such as ones containing the University of Washington. Although there are some variations for the residential census tracts, these variations are not as stark as the downtown commercial and special use census tracts. The downtown census tracts are the primary drivers of the difference in square footage information. In the next section, the change/difference in the damage estimates from HAZUS™ using local data and default data is discussed.

4.4 Results of Earthquake Scenarios: Default Data vs. Local Data

This section discusses the impact of the variation in the building inventory data between HAZUS™ defaults and local data on damage estimation from HAZUS™. To understand the variation in results, the HAZUS™ model was run for the same scenario using default data and building inventory data from local sources, i.e. parcel and tax assessment data. The purpose of this was to control for building data alone while keeping all other variables constant to understand the sensitivity of the model to better building data. However, it was not possible to fully overcome the limitations posed by the lack of structure information in the assessment data, particularly at the level of detail required by HAZUS™. The object of analysis was the loss attributable to building inventory data, which included direct economic loss due to capital stock loss and income loss. Also, the building inventory is a major determinant on other outputs such as shelter requirements, casualties, and other induced losses such as amount of debris and fires. All of these will also be analyzed.

Three scenarios were modeled based on three different magnitudes at the exact same location. Therefore, scenarios were run for a deterministic hazard from a Source

Event on the Seattle Fault Zone Northern Trace. This is a 70.87 km long, reverse slip fault with a 45 degree dip angle that runs just south of downtown Seattle and has a potential for a maximum 7.2 magnitude earthquake. Scenarios of magnitude 5.0, 6.0 and 7.0 on the Richter scale were modeled with the exact same epicenter and all other characteristics constant (Map 4.5). This allows the analysis of the sensitivity of the HAZUS™ model to local data at various magnitudes.

4.41 Damage Losses at the City Level

At the city level, HAZUS™ underestimates the total direct economic loss from buildings by more than $1 billion (69%) for a 5.0 magnitude earthquake, $7.9 billion (115%) for a 6.0 magnitude earthquake, and $15.5 billion (130%) for a 7.0 earthquake. The percentage difference in damage estimates increases with the increase in the magnitude of the earthquake. This is in contrast with the findings of Nordenson et al (1999) where analysis showed that the change in inventory was more sensitive at lower magnitude earthquakes. This may be attributed to the change in data in HAZUS™ since the time when Nordenson et al (1999) did their study or could be a result of the location of the modeled earthquakes vis-à-vis Seattle downtown and the dramatic effect of any earthquake in a densely populated area. It may also be a result of the composition of buildings types – i.e. structure type, age, height, etc.

More than 75% of the difference is contributed by building damage and content damage in all three scenarios. Although the percent difference in inventory loss and wage loss is very large (owing probably to the variation in the data in commercial and industrial occupancy classes), in real terms, they contribute to a smaller extent (particularly inventory loss) on the overall loss. Other factors such as relocation cost, income loss, and rental income loss also contribute to a smaller extent on the total loss. Figure 4.1 – Figure 4.3 shows the results of loss from different scenarios.

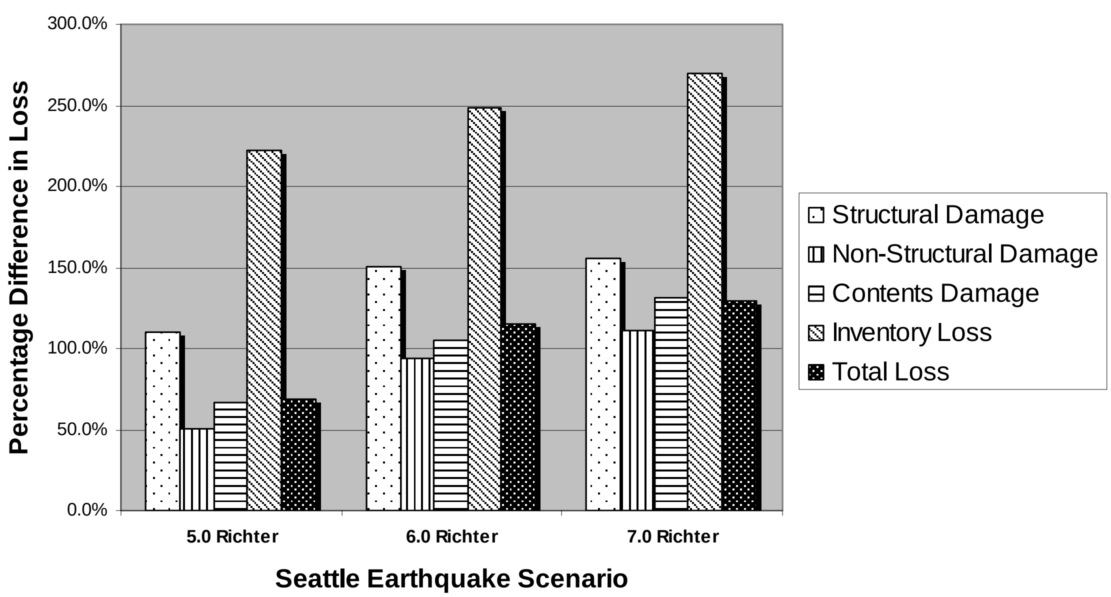


Figure 4.1: Percentage Difference in Loss Due to Building Inventory Change for 3 Earthquake Scenarios in Seattle

Another interesting finding is that the percentage of total loss that can be contributed to building damage (structural and non-structural) is lesser with real data than with default data. This is true for all three magnitude earthquake even though the amount of square footage has increased considerably (Figure 4.3). This may also be a result of the difference in variation of the mix of buildings by occupancy and structure types. The content loss as a percent of the total loss is the same with real data and default data for all three earthquakes. However, whereas the percentage of contribution of building damage loss increases with the increase in the magnitude of the earthquake, the percentage contribution of content loss decreases with the increase in the magnitude of the earthquake, i.e. building damage contributes a larger percentage to total loss for higher magnitude events and content damage contributes a smaller percentage to the total loss for higher magnitude events (with default data and with local data).

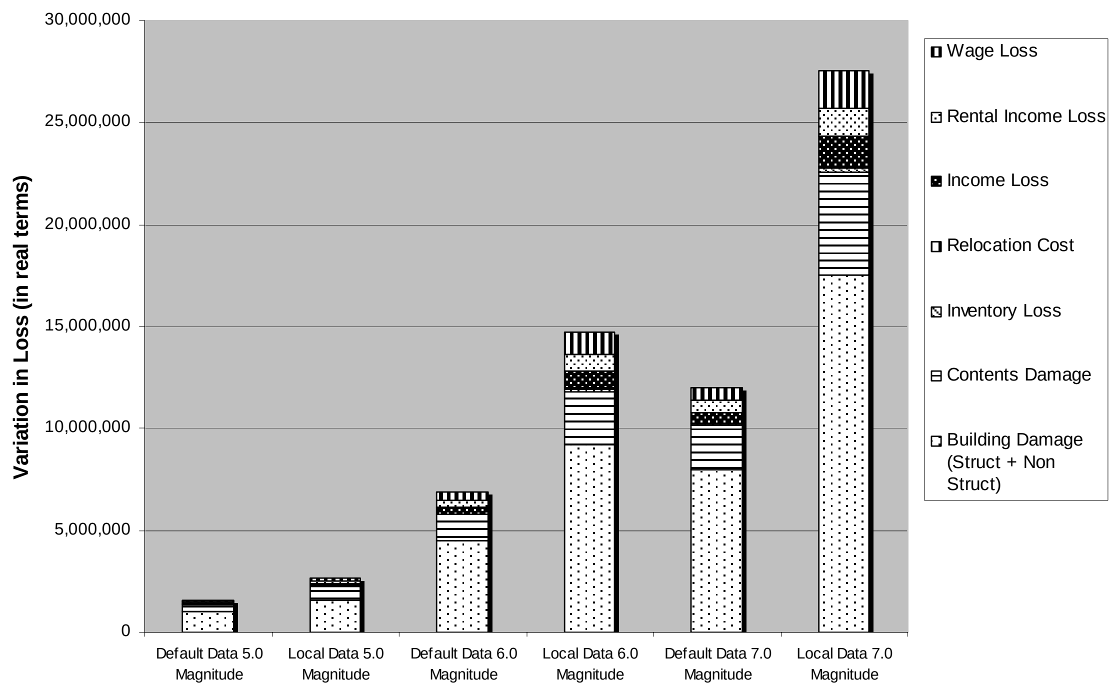


Figure 4.2: Variation in Loss Due to Building Inventory Change for 3 Earthquake Scenarios for Seattle (in real terms)

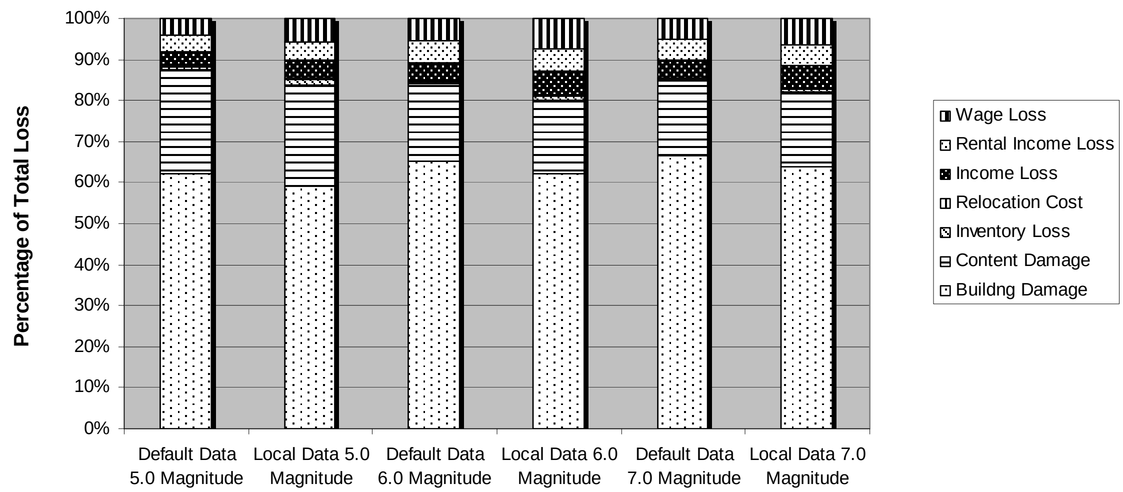


Figure 4.3: Variation in Loss Due to Building Inventory Change for 3 Earthquake Scenarios for Seattle (as a percentage of total loss)

Other losses such as shelter needs, casualties and induced damage such as debris are underestimated by HAZUS™ default data for all magnitude earthquakes (Table 4.8 – Table 4.10). The amount of underestimation increases as the magnitude of the earthquake increases. An interesting finding is that the number of ignitions in HAZUS™ seems to be largely unaffected by the input of better quality building inventory data in all the three earthquakes (Table 4.10). Interestingly, the population exposed is actually overestimated by HAZUS™ in the 5.0 and 7.0 event. Also, the value exposed is also overestimated by HAZUS for the 5.0 event. These variations in HAZUS™ are difficult to explain since no change is made to the demographic data or other data such as sites storing hazardous material, etc. Furthermore, it is unclear why the value exposed is overestimated in HAZUS™ for only the 5.0 event since value should clearly be derived from the value of buildings and contents which is significantly larger for the real data.

One explanation could be the spatial distribution of fires – if the fires affect census tracts that are overestimated by HAZUS™, the exposed value will also be overexposed.

Table 4.8: Variation in Shelter Needs for 3 Earthquake Scenarios in City of Seattle (assuming earthquake occurs at 2 am)

Event	Shelter Need	Local Data	Default Data	Difference	%Difference
Seattle Fault Northern Trace 5.0 Event	Number of Households Displaced	1,608	1,283	325	25.3%
	Number Needing Short-term Housing	430	342	88	25.7%
Seattle Fault Northern Trace 6.0 Event	Number of Households Displaced	21,557	13,697	7,860	57.4%
	Number Needing Short-term Housing	5,478	3,460	2,018	58.3%
Seattle Fault Northern Trace 7.0 Event	Number of Households Displaced	46,976	27,367	19,609	71.7%
	Number Needing Short-term Housing	11,699	6,803	4,896	72%

Table 4.9: Variation in Casualties for 3 Earthquake Scenarios in City of Seattle (assuming earthquake occurs at 2 am)

Event	Casualties	Local Data	Default Data	Difference	%Difference
Seattle Fault Northern Trace 5.0 Event	Severity 1	573	210	363	172.9%
	Severity 2	87	28	59	210.7%
	Severity 3	8	2	6	300.0%
	Severity 4	16	4	8	300.0%
Seattle Fault Northern Trace 6.0 Event	Severity 1	5,310	1,587	3,723	234.6%
	Severity 2	1,245	343	902	263.0%
	Severity 3	160	40	120	300.0%
	Severity 4	312	77	235	305.2%
Seattle Fault Northern Trace 7.0 Event	Severity 1	11,499	3,327	8,172	245.6%
	Severity 2	3,004	816	2,188	268.1%
	Severity 3	407	104	303	291.3%
	Severity 4	791	200	591	295.5%

Therefore, as shown in Table 4.9 by using HAZUS™ local data, emergency managers, planners, and local officials could be underestimating the needs for medical response by 20k people with severity 1 needs, 5K people with severity 2 needs, 938 people with severity 3, and 1830 people with severity 4 needs (for a magnitude 7 earthquake). These changes in result are a direct outcome of the change in building inventory. It is important to note that the results can change significantly by changing other parameters as well, which is beyond the scope of this research. Furthermore, Table 4.8 shows that based on default data alone, HAZUS™ underestimates the displaced households by 19K households and those needing short-term housing by almost 5K. The amount of debris generated could also be underestimated by over 3 million tons of wood and brick, and 6 million tons of concrete and steel (Table 4.10).

In summary, the City of Seattle case study shows that the use of default data in HAZUS™ can result in a large underestimation of direct losses from building damage as compared to the use of local data. This difference increases as the magnitude of the earthquake increases. Furthermore, there is a large variation in the estimation of shelter needs, debris generated, healthcare needs and number of fires and exposure of life and property which can have serious repercussions on planning and responding to a disaster. The next section analyzes the spatial variation in damage across the census tracts in the City.

Table 4.10: Variation in Debris and Fire for 3 Earthquake Scenarios in City of Seattle

(assuming earthquake occurs at 2 am)

Event	Loss Type	Local Data	Default Data	Difference	%Difference
	Debris				
Seattle Fault Northern Trace 5.0 Event	Brick Wood and Others	440	158	282	178.5%
	Concrete and Steel	501	242	259	107.0%
Seattle Fault Northern Trace 6.0 Event	Brick Wood and Others	2,475	812	1,663	204.8%
	Concrete and Steel	4,784	1,817	2,967	163.3%
Seattle Fault Northern Trace 7.0 Event	Brick Wood and Others	4,354	1,428	2,926	204.9%
	Concrete and Steel	9,360	3,302	6,058	183.5%
	Fires				
Seattle Fault Northern Trace 5.0 Event	Number of Ignitions	15	14	1	7.1%
	Population Exposed	591	782	191	-24.4%
	Value Exposed (in thousand $)	56,637	64,713	-8,076	-12.5%
Seattle Fault Northern Trace 6.0 Event	Number of Ignitions	58	55	3	5.5%
	Population Exposed	2,386	2,378	8	0.3%
	Value Exposed (in thousand $)	172,623	167,639	4,984	3.0%
Seattle Fault Northern Trace 5.0 Event	Number of Ignitions	85	83	2	2.4%
	Population Exposed	3,750	3,883	-133	-3.4%
	Value Exposed (in thousand $)	322,601	277,837	44,764	16.1%

4.42 Damage Variation at Census Tract Level

As discussed before, it is also important to analyze the spatial patterns of variation in loss over the different parts of the city when default data is used versus when local data is used. An interesting finding of this research is that, even though HAZUS™ underestimates the total square footage and building count in some census tracts, and overestimates it in other census tracts, the total building economic loss is underestimated by HAZUS™ for **all** the census tracts for all three magnitude earthquakes (with the exception of one downtown census tract that default data overestimates the total building economic loss for a magnitude 5 earthquake). Therefore, even though the square footage is overestimated by HAZUS™ in some census tracts, changes in other building characteristics in the local data leads to more damage even with lower square footage and hence the loss is underestimated by HAZUS™ default data.

The following analysis focuses only on the 7.0 magnitude event. Map 4.6 represents the spatial variation in direct economic loss due to building damage as a result of using local data for a magnitude 7.0 earthquake. In real terms, the variation is less for residential census tracts and increases with more mixed land use census tracts and downtown census tracts. In keeping with the findings with respect to spatial variation of building inventory, this analysis reveals that the census tracts in the downtown show large underestimation in loss due to building damage in real terms. This may be attributable to the larger change in local data from default data in these census tracts and to the fact that the epicenter of the modeled earthquake and the source fault are very close to the downtown. The total loss for the 9 Downtown census tracts discussed before is 156% more when local data is used. However, in percentage terms, the difference is not the highest in the downtown census tracts but is scattered across the City. The University of Washington shows the largest difference in total loss from building, both in real terms and in percent difference. Similar patterns are exhibited by the difference in building damage loss and loss due to damage in content in buildings.

4.5 Conclusions

The case study of City of Seattle proved to be an excellent one for many reasons. The City of Seattle not only had a good GIS program (with many GIS datasets easily available along with good metadata), but more important, the assessment data for the City were very robust. The data were reliable as was validated by the triangulation of many fields that tracked square footage. Furthermore, the City of Seattle data also had information about tax exempt properties that provided an accurate description of the religious, governmental and educational properties. This is often lacking in tax assessment data, making them less useful for damage assessment. However, the data lacked information on the University of Washington, but this was confined to one census tract and because this was identified, various steps were taken to estimate this. Other GIS datasets such as orthoimagery, building footprints, census tract boundaries and zoning proved to be very helpful in quality assurance and analysis.

The findings of this research clearly indicate that local level data, particularly parcel data with corresponding tax assessment data, are an invaluable source for building information in damage estimation and hazard assessment models. The research finds that even with a good dataset such as the one used here, there are limitations in the granularity of the data for the type of structure as required by HAZUS™. Although most assessment data do not carry the same granularity with respect to structure types that is required by HAZUS™, the advantages of having better square footage information over different occupancies alone can result in major differences in overall loss for the City as a whole. Furthermore, the improvement in data also changes distribution of losses attributable to different occupancy classes, and also to the spatial distribution of losses in the city.

An important challenge to using local data involves the availability of information on content value. While information on building value is available from tax assessment data, the value of content is not easily available from any local source. Therefore, this information is a more suitable candidate for estimates based on state or national averages. Furthermore, this information does not change too much across the Country. However, in the current version of HAZUS™, it is not possible to do this through the interface. Therefore, through the use of BIT tool, the square footage information can be updated to reflect the local data, but if the dollar exposure (building and content values) is not provided, this information remains at the default value and hence there is minimal change in losses and damage assessments. Therefore, it is virtually useless to improve the square footage without improving the exposure values through the use of local data. In the current research, some improvement in building and content value information was done to reflect at least the changes in the square feet information. Because there is no user interface provided in HAZUS™ to do this, it had to be done outside of HAZUS™ with help from the HAZUS™ development team.

This research finds that the default building inventory data in HAZUS™ definitely need improvement. The HAZUS™ manual should provide a better description of the source of the default data and how they are input into HAZUS™. This would inform the decision-makers on the reliability of the model and inform them of appropriately using the model. While the variation in education, government and religion occupancy is somewhat expected (since the Dun and Bradstreet data do not provide good estimates for these classes), the large variation in the commercial and industrial occupancies are unexpected and not easy to explain. The variation in building inventory is less in residential census tracts and more in mixed, commercial, industrial and educational census tracts. Particularly the downtown shows stark differences from the default values in HAZUS™. By simply improving the data for the downtown and its vicinity, and the University of Washington census tract, a lot of improvements can be affected for the entire City of Seattle.

This research also indicates that a smaller change in inventory can lead to a lot larger variation in the total loss. This variation increases with increase in magnitude of the earthquake. The impact of change is not confined to economic loss alone, but is spread to other output results that can have a significant impact on the management of the disaster. For example, the results shown in Table 4.7 indicate that by using default data, local officials could severely underestimate the need for hospital rooms, medical services, morgues, shelters, equipment for clearing debris, and so on. Furthermore, the need for personnel to respond to the disaster (such as inspect buildings, coordinate shelter needs, etc.) can be seriously underestimated when default data is used. This can result in poor preparedness and longer recovery. Finally, if HAZUS™ is used in the response stage of a disaster to request assistance (both operational and financial) from higher levels of government such as state and federal governments, the requested aid may be much smaller than the impact of the disaster.

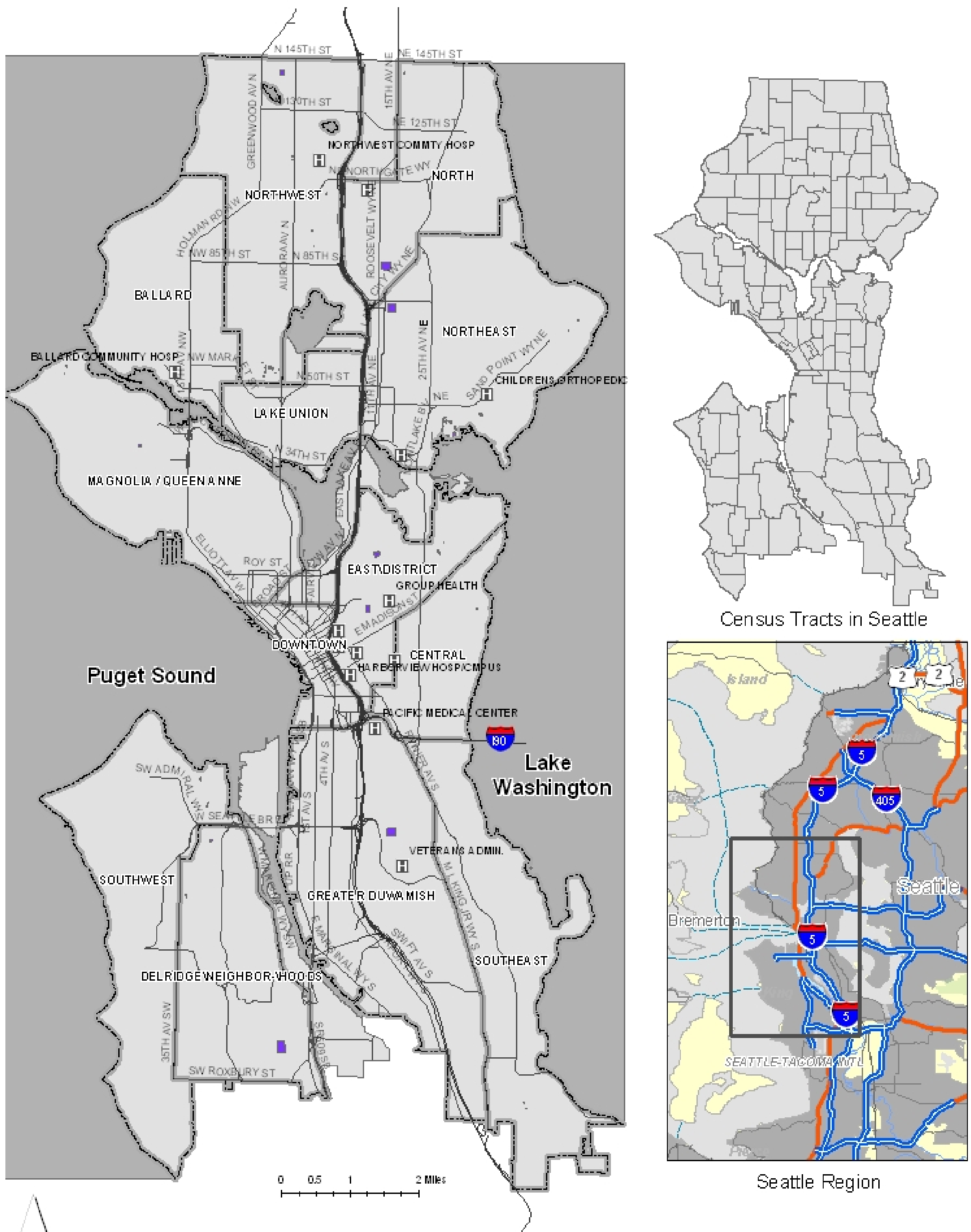
It may be argued that the HAZUS™ tool is meant for regional assessments (where overestimation in one area can be compensated by underestimation elsewhere). However, the findings of this research show that the variation, even at the city level implies that inappropriate distribution of aid and logistical resources will occur when decisions are based on the results of the default data in the HAZUS™ model. It also means that poor decisions will be made regarding mitigation measures and where further investment needs to be made to minimize the loss of life and property (such as retrofitting buildings, etc.).

Even at a regional level, the scenarios can show some misleading patterns if large cities are part of the region. For example, the State of Washington can make some flawed decisions by using scenarios in HAZUS™ with default data to compare loss estimates over several jurisdictions to help prioritize resources on a larger scale. Therefore, national level estimates or other large-scale aggregate data for building inventory can severely restrict the use of tools such as HAZUS™ for local-level decision-making.

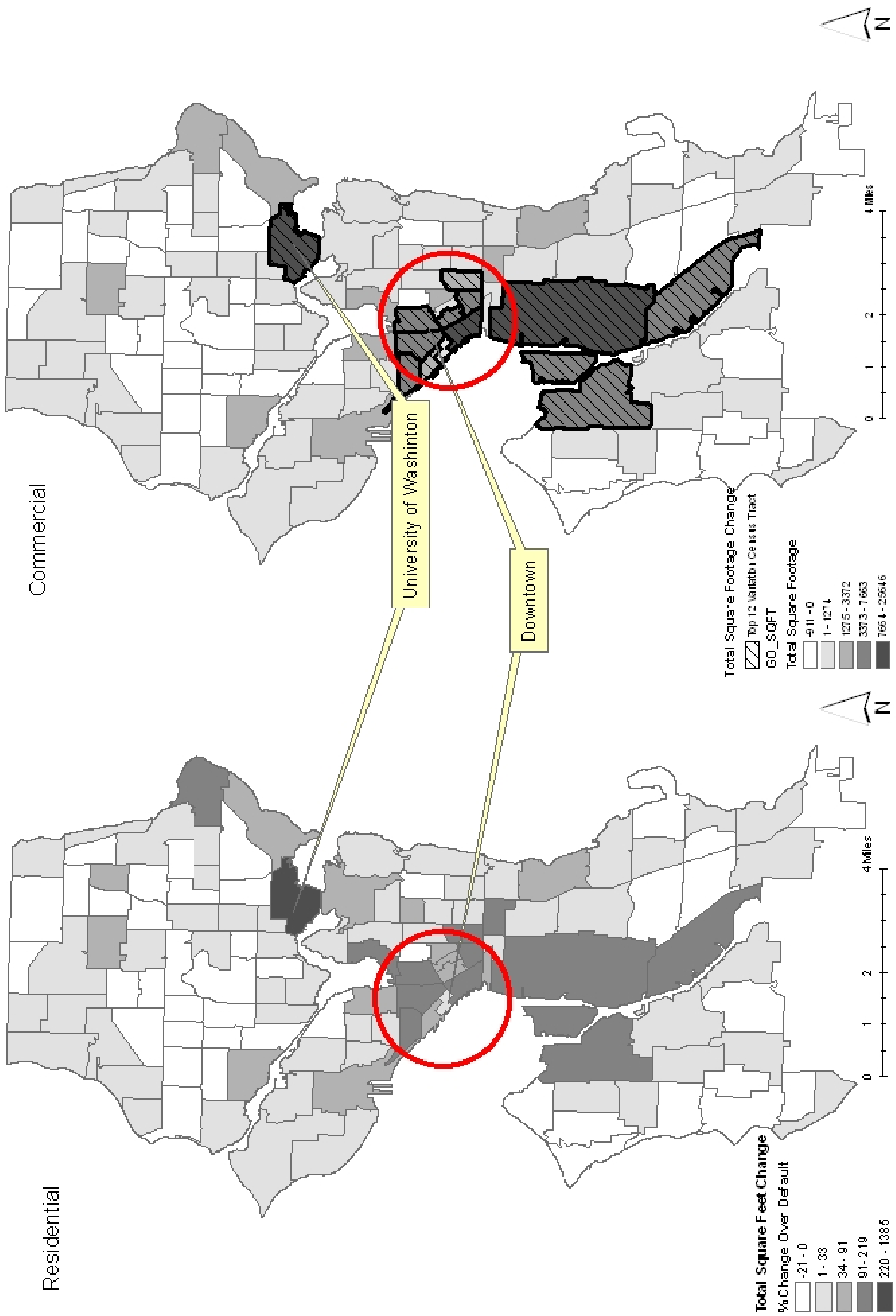
At the scale of King's County or the City of Seattle, officials would have to be extremely cautious in using HAZUS™ with default building inventory data to make choices about mitigation measures for various parts of the city and run what-if scenarios. At the very least, it is crucial that data for the downtown and surrounding census tracts or special census tracts such as the one with University of Washington or with highly specific occupancies be examined very carefully before making any local decisions about allocation of limited resources for different mitigation measures and their impact.

Finally, it is not easy to assimilate local data into the HAZUS™ models, necessitating the need for some expertise with databases and GIS. Even with good data and a well-structured dataset with reliable values, data manipulation needed significant expertise in GIS, databases, and HAZUS™. Hence, data preparation should be undertaken by a team knowledgeable about all three aspects discussed above. Furthermore, some knowledge of the assessment data and planning data is useful.

Therefore, in order to use HAZUS™ with local level data, an emergency manager would need the expertise and help of people in different departments. It is therefore useful to assemble the data before the event rather than in the aftermath of a disaster. Once the data are improved and assembled, they can be a significant resource for preparedness, exercises and planning.

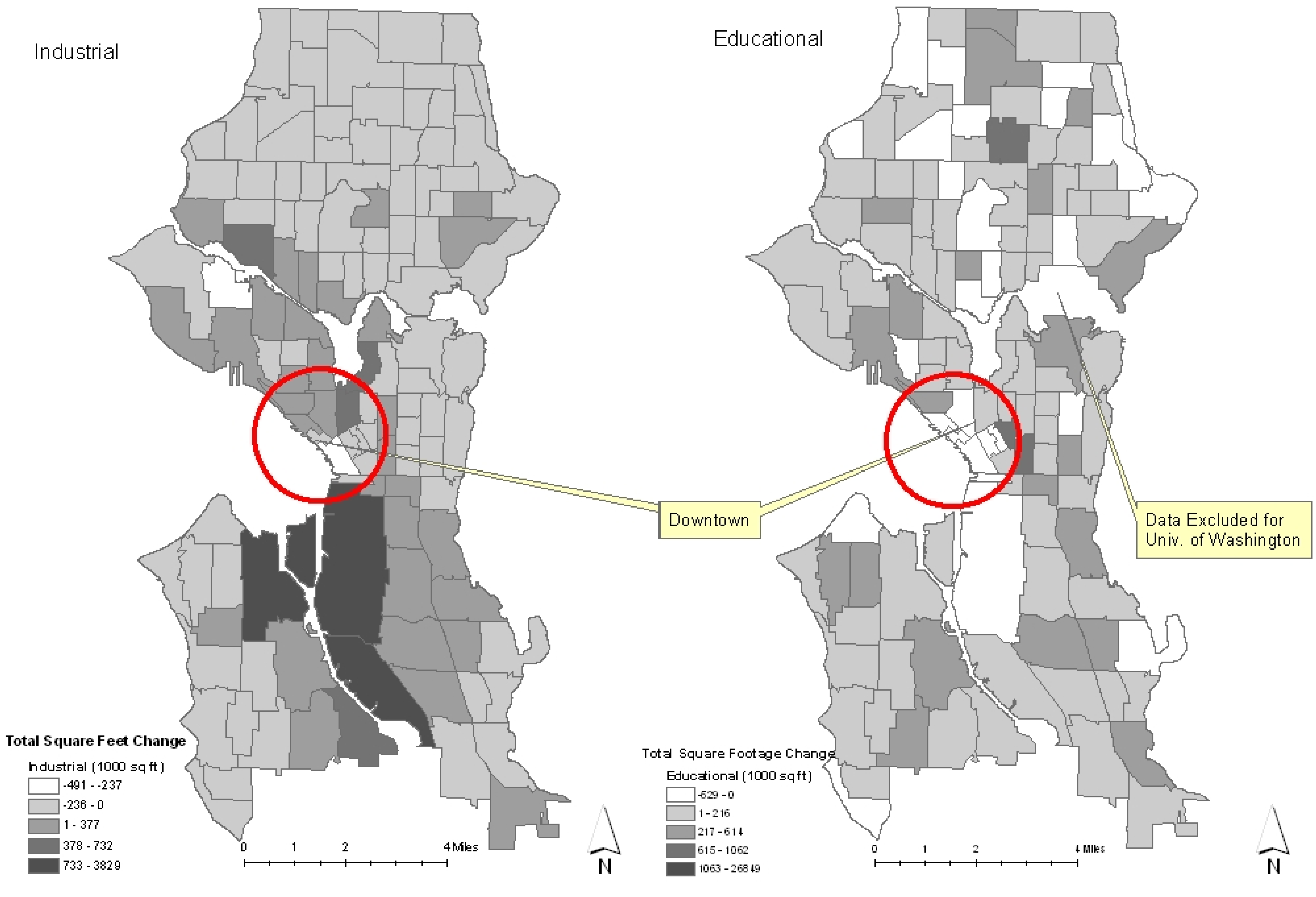


Map 4.1: City of Seattle, WA

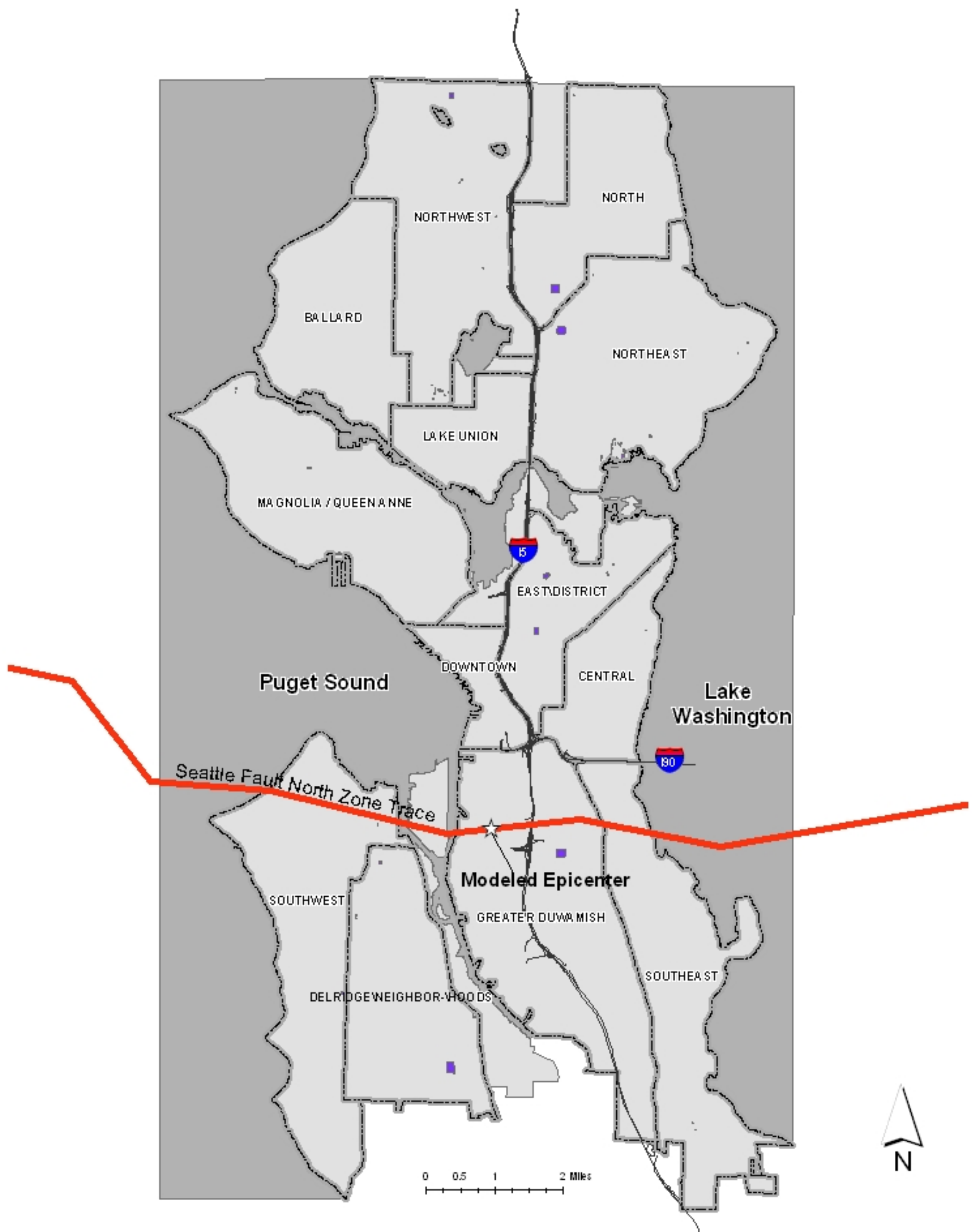


Map 4.2: Variation in Total Square Footage: HAZUS Default vs. Local Data City of Seattle, WA

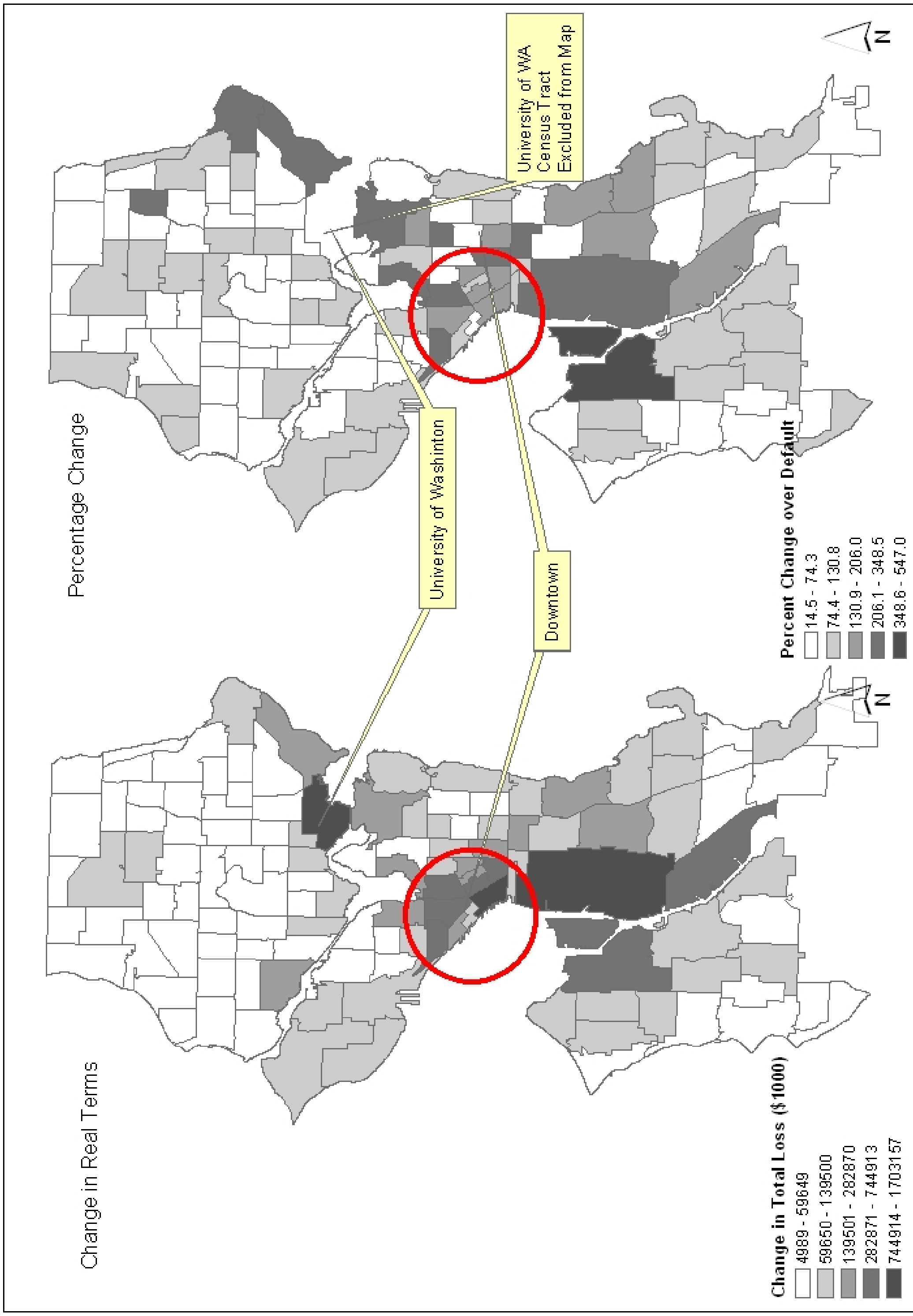
Map 4.3: Variation in Total Residential and Commercial Square Footage: HAZUS Default vs. Local Data
City of Seattle, WA



Map 4.4: Variation in Total Industrial and Educational Square Footage: HAZUS Default vs. Local Data
City of Seattle, WA



Map 4.5: Location of Seattle Fault Zone North Trace and Epicenter of the Modeled Earthquake



Map 4.6: Variation in Total Loss, HAZUS Default vs. Local Data
City of Seattle, WA

Chapter 5: City of Long Beach Case Study

5.0 Introduction

The City of Long Beach was chosen as a case study because it represented a "typical" medium sized city that had a decent GIS division and many GIS datasets (although not as advanced as Seattle). The City also expressed it's willingness to share data for this research and there is an imminent threat of earthquakes in the City (since the Newport-Inglewood fault runs through the City). This chapter will present the findings from the Long Beach case study. It is structured similar to the previous chapter on the case study of the City of Seattle. In this section, a general introduction of the City is provided. This is followed by an overview of GIS in the City and the GIS data provided by the City for this research. A detailed discussion of preparing these data for use in HAZUS™ follows and focuses on data availability, completeness and accuracy and on suitability for input into HAZUS™. The next section discusses the variation of local-level data from default data in HAZUS™ - the section explores the variation for the city as a whole as well as for the various census tracts in the City of Long Beach. The results from running various scenarios of earthquakes will then be analyzed – again looking at the variation of results between using local data and default data in HAZUS™ at the scale of the city and at the census tract level. Finally, the chapter will summarize the findings from this case study.

The City of Long Beach is located in the Greater Los Angeles Metropolitan Region about 22 miles south of downtown Los Angeles. The city was incorporated in 1897 and has a population of 461,522 (US Census 2000). The population of Long Beach has been steadily increasing – the population grew from 358,879 in 1970 to 361,355 in 1980 to 429,433 in 1990 and 461,522 in 2000. According to the 2000 US Census, Long Beach is the most ethnically diverse city in the United States, comprising 45.16% White, 14.87% African American, 0.84% Native American, 12.05% Asian, 1.21% Pacific Islander, 20.61% from other races, and 5.27% from two or more races.

The city comprises an area of 50.4 square miles and density of 9,157 in 2000. It is ranked 34th in population in the United States and 5th in the State of California. The City surrounds another jurisdiction called Signal Hills which is 2 square miles in area approximately (Map 5.1). The City of Long Beach provides many services to Signal Hills although it is a separate jurisdiction. Long Beach is a hotbed for earthquake activity. In the period 1812-1979, there have been about 24 earthquakes ranging in magnitude from 5.1 to 6.9 on the Richter scale in a 35-mile radius of the City. Hence, earthquakes are not a remote possibility in the City of Long Beach but an imminent threat.

The City of Long Beach is the busiest port on the West coast and has a mix of industries (some of the biggest employers in the City include the Long Beach Unified School District, Boeing, City of Long Beach, California State University, Long Beach Memorial Medical Center, VA Medical Center, St. Mary Medical Center, Verizon, etc). Corporations such as Epson America, SCAN Health Plan, (a Medicare non-profit HMO) and Molina Health Care Inc (a Medicaid management healthcare program) are headquartered in Long Beach. Subsidiaries of Toyota make car parts in Long Beach and Polar Air Cargo, an international cargo airline is based in Long Beach (http://en.wikipedia.org/wiki/Long_Beach,_California).

5.1 GIS Description and Data Quality

The City of Long Beach has an Enterprise GIS program in the GIS Central Management Division housed in the Department of Technology Services. This GIS Central Management Division has 4 GIS professionals lead by a GIS Project Manager who was the respondent for the survey in this research and also provided the data for the in-depth case study. Organizationally, it is a typical enterprise implementation. Most core GIS activities take place in the GIS Central Management Division and there are other GIS professionals or advanced users in the various departments (4 in Water, 1 in Gas, 2 in Public Works, 2 in Planning and Building). There are several users of GIS data and services spread over various departments in the city administration.

The City's GIS Central Management Division is responsible for creating/procuring and maintaining many key enterprise datasets including parcels, building footprints, street centerline, ortho-imagery, digital elevation model, schools, hospitals, fire stations, etc. The Los Angeles County creates the tax assessment data and only limited fields are made available to the City for use at a charge. Although the City of Long Beach has a good GIS program, unlike the City of Seattle, there are no proper policies for data sharing and no prepackaged CDs containing data for the city. For this research, the GIS Project Manager at the GIS Central Management Division provided all the required GIS datasets. These datasets included: parcels, building footprints, city boundary, census tracts, location of schools, hospitals, fire stations, police stations, earthquake zones and liquefaction potential zones, and ortho imagery.

The parcel dataset consisted of 88,264 parcels which included parcels that had multiple owners (or condos). There were no core attributes attached to the parcel database. However, an APN (a 10 digit assessor's parcel number) was assigned to each parcel through which assessment data could be linked or joined to the parcel data. APN numbers whose last three digits ended with 500, 501, 502, 503 or so on tracked parcels with multiple owners in the parcel data. A corresponding table that tracked all the multiple parcel numbers for each of these parcels was provided and this table was linked to the assessor's parcel attribute table. The parcel data set is updated on a monthly basis by the City of Long Beach and was provided in December 2003 for use for this research.

The building footprint data were compiled by the City of Long Beach in 1985. Since then, no major updates were undertaken until 2003. In 2003, the GIS Central Management Division updated some of the large buildings. The building footprint data did not carry any meaningful attributes with them (besides the polygon area and perimeter). The height/elevation of the footprint polygon was not captured in the data (even though the data was compiled through photogrammetric techniques). Because of the lack of attribute data and the lack of reliability of this dataset (in terms of updates), it was originally deemed to be of limited use but was later used extensively.

The City also provided high altitude color orthoimagery (coarse resolution with 3 m pixels). Although more recent imagery was available (from 1999), this was in a format that was too large to be provided for this research without a lot of work on the part of the City. Since orthoimagery was used only for verification purposes in this research this was not considered a major setback. Furthermore, where necessary, imagery from public sources such as Google (www.google.com) was used.

The tax assessment data were not easily acquired for the City of Long Beach. The City was unable to provide these critical data, which are created and maintained by the

Los Angeles County Tax Assessor's Office. The City had limited access to these data through a software application that allowed them to view the tax assessment data record by record and with limited attributes. Some cumbersome process/script was available to do exports of the data in another format such as dbf, or txt file, but this had to be done by postal zip codes and the city lacked the personnel to do this. They wanted direct acquisition of the data from the County for this research. This was not easy since Los Angeles County charged 2 cents/parcel for tax assessment data and this charge was not waived for anyone (even for research purposes). The cost of getting the data directly from the County was more than $2,000. However, the data were available through various resellers and these were used for this research. The data used were from a company called DataQuick and dated to June 2001. The data were available through an application and had to be downloaded by zip code and compiled for the entire city by querying for every zip code in the city. The data carried an APN number through which it could be linked to the parcel data. The data were checked randomly for reliability and consistency with approximately 100 records against data available through an Internet website (Los Angeles County Office of Assessor, http://assessormap.co.la.ca.us/mapping/viewer.asp). The randomly selected data were found to be in conformity with the data provided by the reseller and hence this third party source was considered to be reliable for the purposes of this research.

In terms of completeness, about 99% of the parcel data had corresponding data in the tax assessment tables. The remaining 1% missing data may be attributable to data entry errors and the difference in time periods between the GIS parcel data and the tax assessment data. As in the case of City of Seattle, although the assessment data carried many fields (describing the parcel, its ownership, value of land, tax status, and building characteristics), very few were actually populated. However, unlike the City of Seattle, only one single table was used to track all the data (whether it was a residential property, commercial property, or condos). This may have been because the data were acquired from a third party and they may have collapsed multiple tables into one table. Furthermore, attribute data for single parcels were often tracked in multiple records in the parcel data, with only one of these records being complete or current.

Table 5.1 lists the key fields needed by HAZUS™ and the percentage of data that were populated. As can be seen from Table 5.1, data on use of parcel, the square footage of living area, and the year built were very well populated and varied from 95% - 99% of the data. An analysis revealed that 4% of all the parcels were vacant and hence were unlikely to carry any of the above information. However, as can be seen on Table 5.1, data on the height of the building, and the type of structure were very limited and posed a major challenge. Various assumptions, discussed below, were used to improve the data as much as possible before inputting into HAZUS™ and the improvement in data is also shown in Table 5.1.

5.2 Data Preparation for HAZUS™

Unlike the City of Seattle where assessment data were modeled with the individual building in mind, the data for the City of Long Beach were tracked with very little consideration to the individual building and the parcel was the smallest unit at which the data could be meaningfully used. For this research, the parcel was considered the smallest unit and the individual buildings in each parcel were not modeled separately as there was no way of allocating the square footage, year built, use etc. to each building. Since HAZUS™ estimates the number of buildings based on area of the building and the average area based on the use of the building, this was not a major handicap.

As mentioned above, the assessment data carried various fields of data but with different degrees of completeness. While the data for use, area, and year built were fairly complete, the data for height of the building and the type of construction (or structural system) of the building were incomplete. It is important to note that just because a field is populated with value, it does not mean that the data is good or meaningful. There may be data entry errors, or classification inconsistencies that can render the data difficult to use. These will be discussed below about each field along with the methods for improving the data or supplementing the data for each of the above categories required. It is important to note that the assessment data for City of Long Beach had multiple records for many parcels. This was particularly the case for parcels with multiple owners. All values in the multiple records were similar with only one or two field values different. Thus for each parcel, the most pertinent information was derived from different records for that parcel. For example, the maximum square footage, minimum year built, maximum number of stories was used for each unique parcel number. For string fields, data was sorted for the field and both first and last values were compared for fields such as type of construction, condition, quality, IRIS landuse. In the case where values of these were the same, they were used as is. For the purpose of this research, the parcels with multiple ownership records (condos) were treated separately from the parcels with single ownership and various rules were used to select the appropriate record as will be discussed under. Once the calculation for area, height and use were completed, the two types of parcels (parcels with single ownership and parcels with multiple ownerships) were merged together in a single table and joined to the parcel data for the whole city.

The following is a discussion of the local Long Beach data for the fields required by HAZUS™.

Use/Type of Occupancy

The use of parcel was tracked in the field called "Iris_land". Although this data was fairly complete (more than 99% as shown in Table 5.1), there were various discrepancies in how the use of parcel was recorded in the data. For the City of Long Beach, there were 82 different uses ranging from single family residential to airport to cemetery. While there was only one use for single ownership parcels, some of the multiple ownership parcels had multiple uses. In this case, the most common or large use was considered the use of the parcel (without allocating areas to different uses). Some of the use descriptions were not very meaningful – for example, condominium is a type of ownership, not a use. The Long Beach data tracked condominium as a use - the use condominium can refer to either a residential use or commercial use. This was not possible to determine through a generic use such as "condominium". Likewise, many parcels had a use of PUD (which stands for Planned Unit Development, and are usually a mix of commercial, residential, light industrial, open space, and trails etc). PUDs indicate more a zoning type than a use of land. There was no way to break PUDs into occupancies such as residential, commercial, etc. All PUDs were therefore assigned to residential occupancy since a review of PUD parcels indicated largely residential use.

Similarly, “utilities” as a use does not specify the use by specific kinds of utilities such as electricity, gas, communication, etc. Another example is that of use “Public Service” which contained public lands, governmental and administrative buildings and even public housing (which are completely different uses). However, such inconsistencies were very difficult to correct and were left as is for the most part – only the glaring mistakes and errors were corrected. Given the poor quality of the data, other GIS data were used to supplement or correct existing data. For example, all parcels that had a school on them (based on Schools data provided by the city) were calculated as Education land use (some of these parcels had to be selected manually). All parcels that had a hospital on them were coded hospital/nursing home, and those that had fire stations and police stations were also coded as government emergency services. Finally, all parcels that had “*Housing Authority*” as their owner were coded public housing. Other uses were left as is. The use could be further refined through the field that recorded owner names – however this is a very labor-intensive task and beyond the scope of this dissertation.

The City of Long Beach or the Los Angeles County did not collect information on government owned parcels or parcels owned by tax exempt agencies such as schools, etc. This is unlike the case of City of Seattle but very much a standard practice in most assessment data. For such landuses, some of the above data were improved through other means, though not the most accurate, as is discussed below. For parcels with multiple ownerships, the most common and widespread use for each parcel was used to determine the use of the parcel.

Some local uses could not be mapped to HAZUS™ uses. For example, 350 records that had occupancy information in the local data could not be mapped – most of these were uses such as utilities, cemetery, airport, truck terminal, TV facility, easement, waste disposal, truck terminal, etc. While many of these are empty parcels, some (100 + records) had building footprint on them, amounting to 6 million square feet of space. These records were discarded by HAZUS™ when processed by the BIT tool.

Type of Construction/Type of Structure

As in the case of City of Seattle, the information available to determine the type of structural system, and building material was very limited in the assessment data and the classifications were nowhere as detailed as the 36 classes required by HAZUS™. However, unlike the City of Seattle where the data were sufficiently complete for the type of construction, in the City of Long Beach, the field with information about the type of construction/structure was very incomplete. Only 16% of the data had information on the type of construction and there was little else to supplement that information. Another field tracked the frame code and this was also populated sparsely. Furthermore, the frame code was often populated where the construction was not even frame. Appendix D, Table 1 and Table 2 show the various types of construction information and frame codes available from the Los Angeles County Assessor's Office.

The field tracking frame code was combined with information on the type of construction. Therefore, where there was no information on the type of construction, but there was information on the frame code, the frame code data was used to supplement the type of construction information. Otherwise the two were combined together and later classified into HAZUS™ structure types. By combining these two fields, the construction information was improved from 16% to 17.5%. Since 99% of the structures in parcels with land use single family residential (SFR) were frame construction, it was fair to assume the remaining parcels with land use SFR with no construction information to be frame structure. This significantly improved the percentage of data populated from 17.5% to 84%. However, by doing this, an inherent bias was introduced in the data where the single family residential had better information on construction than other land uses. Particularly the type of construction for many of the important high rise buildings was still missing through this methodology. However, since no other source of information was available, this could not be improved any further. For the remaining 16% of the data, it was decided that HAZUS™ could use default structure types based on the height of the structure, use, and the year the structure was built.

Area/Square Footage

Square footage information was well recorded for 95% of the parcels in the City, and was used as the source for area information. For parcels with multiple ownerships, the living area of each owner was recorded separately and was aggregated to calculate the total square footage of that parcel. For the parcels with no area information, the area of the building footprint was used as the total area of the building for that parcel (since for parcels with no square footage information, the number of stories was also missing). This may have underestimated the area since many of the buildings probably had multiple stories. However, it does significantly reduce the underestimation that may have occurred without doing this. Furthermore, since many of the footprints themselves could be larger than the living space (to account for canopies, staircases, multiple car attached garages etc), it was fair to assume that this step was a rational one.

Approximately 26 million square feet of space was calculated by using building footprints for parcels that had use information. Much of this space was concentrated in the government (6.2 million sq ft), education (8.9 million sq ft) occupancy classes, and under the utilities (1.5 million sq ft) and airport (3.9 million sq ft) uses. By supplementing the area information as such, the data were improved by almost 1.5% as can be seen in Table 5.1. There is also a possibility of overestimation in some cases by using the assumptions as some building footprints may represent sheds or temporary shelters, oil tanks, etc. In the absence of anything in the building footprint dataset to distinguish between buildings and non-buildings, there is no way to remove these extraneous structures.

Table 5.1: Improvement of Assessment Data through Supplementing in City of Long Beach

Field Description	Records populated (out of 88017)	% populated	Records populated by supplementing	% records populated by supplement-ing
Use of Parcel	87,123	99%	87,130	99%
Area/ Living Area	83,696	95%	84,872	96.4%
Age/Year Built	83,697	95%	83,697	95%
Height/No. of Stories	63,982	73%	83,059	94%
Type of Construction	13,812	16%	64,248	73%

Age/Year Built

There were no supplemental data available to improve this field. Since 95% of the data was already populated, this was considered suitable enough for this research.

Height/Number of Stories

Information about height/number of stories needed significant improvement since only 73% of the data had height information as shown in Table 5.1. Various steps were used to calculate height/number of stories. For all data that had height/number of stories information in the assessment data, this was used without any changes. For multiple ownership parcels, the highest number of stories was used. For records that had no information about the number of stories, but had information on the total living area, the number of stories was calculated by dividing the total living area by the area of the building footprint in that parcel. All footprints of less than 600 sq. feet were discarded to eliminate small sheds, detached garages, etc. For parcels with no information on living area, the height was assumed to be 1 storey. This method had some errors due to the different vintage of GIS data and assessment data (the GIS analysis caused some building centroid to fall in a different parcel, etc). All height information that was calculated in the above manner that was over 8 stories was reviewed case by case and corrected based on aerial imagery and other sources on the Internet. For records that didn’t have height information or area information in the assessment data, the height calculated based on the building footprint was obviously 1 storey. Since HAZUS™ lumps height into 3 main categories (>8 stories, 4-7 stories, and 1-3 stories), for the most part, it was assumed that for a majority of the data that the height was calculated, the information led to classification in the correct category for HAZUS™ purposes.

Census Tract

Although there was a field in the assessment data to track the census tract of every parcel, this data was also not complete and it was more suitable to assign census tracts based on the census tract layer provided by the City of Long Beach.

Building and Content Exposure

Exposure data for the City of Long Beach was not compiled because of the research design discussed in detail in Chapter 4. For the City of Long Beach, it was obviously not easy to get this information as the data had a lot of issues and various assumptions were used to interpolate and populate square footage and other attribute data. However, such interpolation was not possible for value data. As expected there was also no information in the assessment data about content exposure. Therefore, the exposure per square feet for every census tract was calculated for both building and content exposure from HAZUS™ default data to reflect the updated areas as was done for the City of Seattle. Similar issues to Seattle were encountered for this (a detailed discussion in provided in the previous chapter). Other fields such as condition and quality were left as is and were mapped to corresponding fields by using the Building Inventory Tool (BIT) in HAZUS™. It is important to note that even though a lot of fields were populated, not all values could be mapped to HAZUS™ classifications.

Therefore, in conclusion, for a city with sketchy assessment information, it is not easy to transform the local data to HAZUS™. Various GIS datasets were useful in making assumptions and improving the data. However, this embedded a level of uncertainty that is difficult to quantify. It was particularly difficult to map the information on the type of building, since there were very few sources of this data other than assessment data. On the other hand, there are hardly any sources for these data for the content exposure at the local level. HAZUS™ does not provide any tools to update the content and building exposure value based on improvements in other data (particularly square footage data) based on local data. In the absence of improved exposure information, it is virtually useless to improve other data using local data. However, like the City of Seattle some back door methods were used to incorporate the improvements in exposure values into HAZUS™ for this research.

In the next section, the variation of default data from local data will be analyzed, first for the City as a whole and then at the level of the census tracts. This will be followed by an analysis of change in damage estimation for the same scenarios when default data and local data are used.

5.3 Building Inventory Data Variation – Default vs. Local Data

This section will analyze the difference in default data in HAZUS™ as compared to the local data. The differences will be first analyzed at the city level and then analyzed at the census tract level to understand the spatial variation across different parts of the city.

This will help identify strategies to improve the data in HAZUS™ and to use the model appropriately.

5.31 Variation at the City Level

The variation of building inventory data for Long Beach is presented in Table 5.2. The local data for the City of Long Beach compares well with HAZUS™ default data with respect to the total square footage in the city. However, even though the data compares well, it is to be noted that the real data is underestimated in itself because of poor quality data and the method used to interpolate some missing data as discussed before. Furthermore, the BIT tool discards data that has missing values in some fields (if use/occupancy is not available, the data is discarded). Therefore, although 270 million square feet of data was input into BIT for Long Beach, the output from BIT reduces it to 252 million square feet. Thus, the default data is off from the original input data by 8.4% but after being processed by BIT in HAZUS™, records that have unknown use/occupancy information are discarded and this makes the data off by only 2.3%.

As mentioned before, the 270 million sq ft is in itself an underestimation - 36 million sq ft of space was calculated by using the building footprint area and with the assumption of 1 storey structure. Thus, for buildings that were more than 1 storey, there was a significant underestimation. If area was assumed to be two stories and all the unknown uses were mapped to some use, this would lead to the default data being off by about 23%. However, given the poor quality of the local data and the assumptions made to calculate some of the missing data, it is difficult to say what is closer to reality in the case of Long Beach.

Even though the overall square footage is well estimated by HAZUS™ default data for the City of Long Beach, there is significant variation in the various occupancy classes as can be seen in Table 5.2. The square footage for the residential occupancy class is overestimated by HAZUS™ defaults for the City of Long Beach. The only other occupancy class that is overestimated by HAZUS™ is the agriculture occupancy. All other occupancies are underestimated by HAZUS™ and the percentage difference over default is very high for government and education occupancy classes. However, in real terms, the higher variations are seen in the residential and commercial occupancy classes.

Table 5.2: Variation in Square Footage by General Occupancy Classes in City of Long Beach (in thousand square feet)

Occupancy	Default Data	Local Data	Total Difference	Percent Difference over Default
Residential	215,773	173,109	-42,664	-19.8%
Commercial	26,929	49,952	23,022	85.5%
Industrial	4,325	15,533	11,207	259.1%
Agriculture	94	44	-50	-52.7%
Religion	991	2,618	1,627	164.2%
Government	308	6,925	6,617	2,147.1%
Education	1,167	7,088	5,920	506.9%
Total	**249,590**	**255,270**	**5,680**	**2.3%**

The overestimation for the residential occupancy class is very high in real terms and offsets the underestimation in many other classes so that the data appears very accurate overall in terms of total square footage in Long Beach. Table 5.3 shows that a high degree of the overestimation of the residential occupancy class is concentrated in single family homes. However, the square footage for duplexes and triplex/quads

(RES3A and RES 3B) are underestimated and the square footage for multi-dwellings (with the exception of dwellings with 20-49 units) are overestimated by HAZUS™. These findings might be a reflection of the demographic make-up of the City of Long Beach. Temporary lodgings are also underestimated by HAZUS™ and this might reflect the greater number of hotels and temporary lodgings that tend to be concentrated in downtowns and large cities. The local data did not yield itself to be categorized into Institutional dormitory and hence the overestimation by HAZUS™ for this occupancy class.

The commercial occupancy class is also underestimated in HAZUS™ and much of this underestimation is concentrated in COM1 occupancy class (retail trade) and COM4 (professional/technical services). A part of the underestimation in HAZUS™ is also contributed by the fact that the default data in HAZUS™ does not contain any data for parking garages. There is also a lot more wholesale trade in Long Beach than is estimated by HAZUS™ and this might be a reflection of an old city with a lot of warehousing and old industrial uses. The square footage for banks is overestimated by HAZUS™. Industrial occupancy as a whole is underestimated by HAZUS™ as seen in Table 5.2 and the amount of square footage for industrial occupancy is less than half of what is in the local data. This underestimation is spread throughout all the specific occupancy classes within the industrial general occupancy class including heavy industries, light industries, high tech, and foods/drugs/chemicals. The default data overestimate the square footage for construction and metals/minerals.

Other occupancies that are underestimated by HAZUS™ are the religious, government and education occupancy classes. It is interesting to note that two-thirds of

Table 5.3: Variation in Square Footage by Specific Occupancy Classes in City of Long Beach (in thousand square feet)

HAZUS Specific Occupancy	Description	Default Data	Local Data	Total Difference	Percent Difference over Default
RES1	Single Family Dwelling	132,543	99,849	-32,694	-24.7%
RES2	Manufactured Housing	2,283	1,268	-1,016	-44.5%
RES3A	Duplex	11,599	14,291	2,692	23.2%
RES3B	Triplex / Quads	11,757	14,208	2,451	20.8%
RES3C	Multi-dwellings (5 to 9 units)	15,632	11,206	-4,426	-28.3%
RES3D	Multi-dwellings (10 to 19 units)	15,475	12,070	-3,406	-22.0%
RES3E	Multi-dwellings (20 to 49 units)	8,560	9,890	1,329	15.5%
RES3F	Multi-dwellings (50+ units)	9,944	7,161	-2,784	-28.0%
RES4	Temporary Lodging	1,292	2,692	1,399	108.2%
RES5	Institutional Dormitory	6,309	21	-6,289	-99.7%
RES6	Nursing Home	375	454	79	20.9%
COM1	Retail	6,503	15,949	9,446	145.3%
COM2	Wholesale Trade	4,216	6,400	2,184	51.8%
COM3	Personal and Repair Services	2,879	2,781	-98	-3.4%
COM4	Professional/Technical Services	7,372	11,371	3,999	54.2%
COM5	Banks	393	53	-340	-86.5%
COM6	Hospital	2,421	3,012	592	24.4%
COM7	Medical Office/Clinic	1,374	33	-1,341	-97.6%
COM8	Entertainment & Recreation	1,740	2,474	734	42.2%
COM9	Theaters	32	71	40	125.3%
COM10	Parking	0	7,808	7,808	#DIV/0
IND1	Heavy	1,482	5,803	4,321	291.6%
IND2	Light	1,728	7,669	5,941	343.7%
IND3	Food/Drugs/Chemicals	281	825	545	193.9%
IND4	Metals/Minerals Processing	203	322	119	58.5%
IND5	High Technology	11	35	24	215.3%
IND6	Construction	620	878	258	41.6%
AGR1	Agriculture	95	45	-50	-52.7%
REL1	Religious	991	2,618	1,627	164.2%
GOV1	General Services	280	6,925	6,645	2369.9%
GOV2	Emergency Response	28	0.0	-28	-100.0%
EDU1	Grade Schools	731	4,173	3,442	471.0%
EDU2	Colleges/Universities	437	2,915	2,478	567.1%
Total		**249,590**	**255,270**	**5,680**	**2.3%**

the parcels that were assigned the education occupancy did not contain any square footage or height information. For these buildings the square footage was calculated based on the building footprint and the height was also estimated to be 1 storey.

Therefore, for this occupancy, the local data is itself underestimated. According to the default data there is less than half a million square feet of space for the college/university occupancy class as opposed to 2.9 million square feet per the local data. The Long Beach campus of California State University alone comprises 1.86 million square feet of space (California State University 2000).

There are many colleges and universities in Long Beach – California State University, Long Beach City College, Brooks College, Keller Graduate School of Management, Nova Institute of Health Technology, Travel and Trade Career Institute, American Institute of Health Sciences, John Wesley International Barber and Beauty College, and Pacific Coast University School of Law. Therefore, the local data seems more representative of the reality and the default data is far from it. The breakdown in specific occupancies for government is not available from local data and hence it is not possible to analyze the specific occupancy variations for this class.

While the variation in square footage is small, the variation in building count in Long Beach is much higher. The building count is **overestimated** in HAZUS™ (even though the building square footage is underestimated). Much of the overestimation in the count of buildings is concentrated in the residential occupancy class. For all the other occupancies, the count of buildings is underestimated by a large amount. Within the residential general occupancy, the building count overestimation is largely concentrated in single family homes (Table 5.5) and is proportional with the overestimation of the square footage for this occupancy class. However, multifamily buildings (i.e. occupancies RES3A through RES3F) are underestimated by HAZUS™ by more than 7,500 buildings or by 75% which is much higher than the underestimation of the square footage for these buildings. This is because of higher underestimation of duplexes and triplexes which are fairly common in cities.

Table 5.4: Variation in Building Count for General Occupancy Classes in City of Long Beach

HAZUS General Occupancy Class	Default Data	Local Data	Total Difference	Percent Difference over Default
Residential	95,833	79,253	-16,580	-17.3%
Commercial	1,083	5,170	4,087	377.4%
Industrial	105	2,033	1,928	1836.2%
Agriculture	0	13	13	#DIV/0
Religion	38	222	184	484.2%
Government	28	521	493	1760.7%
Education	10	122	112	1120.0%
Total	**97,097**	**87,334**	**-9,763**	**-10.1%**

The count of buildings in the commercial occupancy class is also underestimated by HAZUS™ and some of it is contributed by the parking structures that are missing from the Dun and Bradstreet data. However, a significant underestimation is seen in the retail occupancy (COM1). The default data shows only 9 buildings in the City of Long Beach that are used for retail trade whereas, the local data shows almost 2900 buildings dedicated to retail trade. Likewise, other commercial occupancies are also underestimated in HAZUS™ with the exception of banks which are overestimated. The local data does not record the number of banks and the square footage associated with them very well. The number of buildings in all specific occupancies under the industrial general occupancy is also underestimated by HAZUS™ even though some of them were overestimated in terms of square footage.

Schools and colleges were not distinguished in the local data – instead, they were assigned the Education occupancy and HAZUS™ was allowed to split the data between grade schools and universities. As mentioned before, the square footage for the Education occupancy class was missing for a large part in the assessment data and was

Table 5.5: Variation in Building Count for Specific Occupancy Classes in City of Long Beach

HAZUS Specific Occupancy Class	Description	HAZUS Default Data	Local Data	Total Difference	Percent Difference over Default
RES1	Single Family Dwelling	82,832	61,119	-21,713	-26.2%
RES2	Manuf. Housing	2,144	112	-2,032	-94.8%
RES3A	Duplex	3,784	7,666	3,882	102.6%
RES3B	Triplex / Quads	3,795	5,009	1,214	32.0%
RES3C	Multi-dwellings (5 to 9 units)	1,777	2,188	411	23.1%
RES3D	Multi-dwellings (10 to 19 units)	1,074	1,814	740	68.9%
RES3E	Multi-dwellings (20 to 49 units)	68	930	862	1,267.6%
RES3F	Multi-dwellings (50+ units)	98	206	108	110.2%
RES4	Temporary Lodging	7	160	153	2,185.7%
RES5	Institutional Dormitory	239	17	-222	-92.9%
RES6	Nursing Home	15	32	17	113.3%
COM1	Retail	9	2,106	2,097	23,300.0%
COM2	Wholesale Trade	114	288	174	152.6%
COM3	Personal and Repair Services	246	625	379	154.1%
COM4	Professional/Technical Services	65	817	752	1,156.9%
COM5	Banks	91	5	-86	-94.5%
COM6	Hospital	44	102	58	131.8%
COM7	Medical Office/Clinic	173	22	-151	-87.3%
COM8	Entertainment & Recreation	340	495	155	45.6%
COM9	Theaters	1	3	2	200.0%
COM10	Parking	0	707	707	#DIV/0
IND1	Heavy	44	348	304	690.9%
IND2	Light	50	914	864	1,728.0%
IND3	Food/Drugs/Chemicals	3	71	68	2,266.7%
IND4	Metals/Minerals Processing	4	23	19	475.0%
IND5	High Technology	0	577	577	#DIV/0
IND6	Construction	4	100	96	2,400.0%
AGR1	Agriculture	0	13	13	#DIV/0
REL1	Religious	38	222	184	484.2%
GOV1	General Services	25	521	496	1,984.0%
GOV2	Emergency Response	3	0	-3	-100.0%
EDU1	Grade Schools	2	106	104	5,200.0%
EDU2	Colleges/Universities	8	16	8	100.0%
Total		**97,097**	**87,334**	**-9,763**	**-10.1%**

calculated based on various assumptions which underestimates the amount of real square footage for this occupancy class. The number of schools and universities in HAZUS™ default is grossly underestimated both in terms of square footage and in terms of building count. Particularly noteworthy is the number of schools estimated by HAZUS™ - the default data show 2 grade schools in the entire city whereas the local data shows 106 schools. Interestingly enough, a separate dataset in HAZUS™ compiles data for schools from a national-level database and this source documents 124 schools in Long Beach.

As discussed in the previous section, the structure/building type information in the local data was neither complete nor reliable. Hence, given the great deal of uncertainty in the local data, an in-depth analysis of this is not worthwhile. The dollar exposure (although not completely reflective of local data since average default values were used as discussed before), also show wide variation in the various occupancies. Table 5.6 shows the changes in dollar exposure for building and content. Thus, overall the square footage is underestimated by HAZUS™ by 2.3%, the total dollar exposure is also underestimated by 2%.

Table 5.6: Variation in Dollar Exposure for General Occupancy Classes in City of Long Beach

HAZUS General Occcupancy Class	**Default Bldg Exposure** (in 1000 $)	**Local Bldg Exposure** (in 1000 $)	**Default Content Exposure** (in 1000 $)	**Local Content Exposure** (in 1000 $)	**Default Total Exposure** (in million $)	**Local Total Exposure** (in million $)	**Percent Diff. in Total Exposure over Default**
Residential	22,261,963	17,640,196	11,135,673	8,824,001	33,397,636	26,464,197	-21%
Commercial	2,739,142	3,987,313	3,024,541	4,221,976	5,763,683	8,209,289	42%
Industrial	339,021	1,186,604	487,887	1,743,395	826,908	2,929,999	254%
Agriculture	6,341	3,072	6,341	3,072	12,682	6,144	-52%
Religion	122,100	330,665	122,100	330,665	244,200	661,330	171%
Govt.	31,440	671,995	33,484	671,995	64,924	1,343,990	1970%
Education	127,351	775,243	154,413	953,785	281,764	1,729,028	514%
Total	**25,627,358**	**24,595,088**	**14,964,439**	**16,748,889**	**40,591,797**	**41,343,977**	**2%**

In summary, the square footage at the City level in Long Beach is fairly well represented by HAZUS™ defaults as compared to local data. However, the assessment data for the City of Long Beach are of poor quality and various assumptions were made to improve the data. A few minor changes in assumptions could lead to larger variation in the local data from HAZUS™ defaults as discussed earlier. Furthermore, even though the square footage overall for the whole city is not very different, the variation in individual occupancy classes (both in square footage and in building count) is quite large. This variation in the various occupancy classes ultimately results in little change in total dollar exposure. In the next section, the spatial variation in building inventory over the various census tracts in the City of Long Beach are analyzed in greater details.

5.32 Variation at the Census Tract Level

The spatial variation in total square footage per occupancy class shows some interesting patterns. HAZUS™ overestimates the total area in 61 (58%) census tracts and underestimates the total area in 44 (42%) census tracts. However, the degree of underestimation is higher even though it is in fewer census tracts. The census tract with the largest variation is the one containing the airport. In this census tract, HAZUS™ significantly underestimates the square footage for the total area – this underestimation would be even larger given the fact that 13 parcels in this census tract (with almost 5 million square foot of space) did not have use/occupancy information from the assessment data and were discarded by HAZUS™. The model overestimates in census tracts that are predominantly residential (with a mix of housing types). It underestimates in census tracts with special uses such as downtown, and tracts peripheral to the downtown with high density residential, tracts with mixed use development, industrial uses, and tracts containing California State University, US Naval hospital and El Dorado

Park. There are certain exceptions to this general finding. The spatial variation of total area in the City of Long Beach is provided in Map 5.2. Unlike the City of Seattle, the downtown in the City of Long Beach comprises fewer census tracts and is not as dense and hence does not show the same concentration of underestimation.

Residential occupancy is overestimated for 91 census tracts and underestimated for 13 tracts and is the same for the census tract that comprises the El Dorado Park. The residential occupancy is overestimated in all census tracts that are primarily residential census tracts. In most of these census tracts, commercial and other occupancies are underestimated. The underestimation of residential occupancy is occurring in special use census tracts such as the airport and El Dorado Park census tract, mixed use tracts, and the census tracts with high density residential developments (Map 5.3).

Commercial is overestimated in only 16 census tracts and is underestimated in the remaining 89 tracts. The overestimation of commercial is occurring in census tracts that are predominantly one use such as residential, industrial, etc. or used for a specific purpose such as El Dorado Park, California State University and US Naval Hospital as shown in Map 5.3. There is very little commercial activity in these tracts. In other census tracts the commercial square footage is underestimated with the majority of the commercial underestimation occurring in the downtown and mixed use census tracts. This may point to default data being better for larger buildings and establishments used for commercial activity than for reporting small businesses and other mom and pop establishments. Furthermore, it may point to the way in which Dun and Bradstreet data are being processed. However, although some generalizations may be reached, as found in City of Seattle, there are no clear and consistent patterns of overestimation and underestimation by the type of census tracts in Long Beach either.

Industrial is overestimated in 68 census tracts and underestimated in 37 census tracts. The underestimation is occurring in census tracts that have industrial activity occurring in them (i.e. industrial census tracts or residential industrial census tracts or other tracts that have warehouses and other industrial activity interspersed between residential). The degree of underestimation is also higher even though fewer census tracts show underestimation (Map 5.4). Again, this may point to the inaccuracies in the Dun and Bradstreet data or in the way these data are being used in HAZUS™.

The education occupancy is generally well estimated in HAZUS ™ spatially. There are a few exceptions such as the census tract with California State University, Long Beach Community College, and a few other census tracts that have larger public high schools and private schools where square footage is underestimated. Approximately 67% of the parcels with education or public school use did not have any square footage information in the assessment data. The area for these were estimated based on building footprint and since no height information was available for these parcels, the height was assumed to be 1 floor. Hence there may be some underestimation even in local data sources due to the lack of assessment data for this occupancy class. Only the Long Beach Airport census tract shows a large overestimation and it is not clear what the source of this is since there are no large educational establishments in this census tract (Map 5.4).

Similar to the parcels used for education purpose, for those used for public service (largely dedicated to government occupancy class), the assessment data only documented the use of the parcel, but no other information regarding the square footage and height.

Square footage was estimated from the building footprint and the height was assumed to be 1 storey since no other height information was available. Of the 141 parcels with use of “Public Service” that actually had buildings on them, only 5 had square feet information from the assessment data. The rest were calculated based on building footprints.

These data indicate that for both the education occupancy and government occupancy, the local data also provides limited reliability. Government occupancy is well estimated for most census tracts (since both the local data and default data show no facilities in this occupancy for a large number of census tracts). The governmental occupancy is underestimated in HAZUS™ for 54 census tracts and the largest of this occurs in the tract containing the Airport (which is coded in the local data as public service use) as shown in Map 5.5. The underestimation would be a lot larger if better assessment data were available for this occupancy class for the Airport census tract as discussed earlier. Other census tracts where the government occupancy is largely underestimated by HAZUS™ include census tracts that contain the US Naval Hospital, public buildings like the City Hall, post office and large parks etc.

Map 5.5 shows that square footage information for the religious occupancy class is fairly well estimated for all census tracts and the variation in overestimation and underestimation are small. However, more census tracts are underestimated by HAZUS™ and less are overestimated. It is to be noted here that although parcels that were used for education and public purposes did not contain square footage information, parcels used for churches and religious purposes were populated with information on square footage. The height was not well documented in the HAZUS™ data but it was safer to assume 1 storey for religious occupancy than for school or government use.

Since the structure information was missing for a majority of the parcels in the assessment data and much of this data was inferred, a comparison of this spatially is unlikely to yield any significant conclusions and hence is not attempted in this section. Therefore, this section reveals that even though the overall total square footage for the City of Long Beach is well estimated by HAZUS™ at the level of the city, there are significant variations when analyzed spatially at the census tract level across the city. However, this variation is not consistent and cannot be fully explained by the type of census tract. In general HAZUS™ default data tend to overestimate for census tracts that are predominantly residential census tracts and severely underestimate square footage for census tracts that are special use tracts such as downtown tracts, or tracts containing large amounts of particular uses such as airports, parks, or universities. These census tracts are particularly vulnerable to bad decisions if default data are used. In the next section, the variation in earthquake damage and losses from using default data and local data will be analyzed in greater detail.

5.4 Results of Earthquake Scenarios: Default Data vs. Local Data

This section discusses the impact of the above variation in the building inventory data between HAZUS™ default data and local data on damage estimates from HAZUS™ for the City of Long Beach. To understand the variation in results, the HAZUS™ model was run for various scenarios using HAZUS™ default building inventory data and building inventory data from local sources, i.e. parcel and tax assessment data. It is important to reiterate that information on building type is very important in assessing the level of damage and hence loss. However, this information was incomplete and unreliable for the assessment data for the City of Long Beach. Hence it was not possible to fully overcome the limitations posed by the lack of structure information. The data were prepared to match the HAZUS™ building types as well as possible and was input into HAZUS™ with the help of the Building Inventory Tool (BIT). Parameters in HAZUS™ were allowed to further break the structure type information into more specific classes. Therefore, building inventory data were the only data that were changed in the model. The purpose of this was to control for building data alone while keeping all other variables constant to understand the sensitivity of the model to better building data. The object of analysis was the loss attributable to building inventory data, which includes direct economic loss due to capital stock loss and income loss. Also, the building inventory is a major contributing factor on other post-disaster needs such as shelter requirements, casualties, and other induced losses such as debris and fires, which were also analyzed.

Three scenarios were modeled based on three different magnitudes at the exact same location. The scenarios were run for a deterministic hazard from a source event on the Newport-Inglewood Fault with an epicenter (latitude 33.7775 and longitude -118.132) located very close to a historical earthquake of magnitude 5.4 that occurred in 1933. The Newport-Inglewood Fault is a 65.74 mile long, strike slip fault with a 90 degree dip angle that runs through the City of Long Beach and has a potential for a 7.1 magnitude earthquake. Scenarios of magnitude 5.0, 6.0 and 7.0 on the Richter scale were modeled with the exact same epicenter and all other characteristics constant. This allowed the analysis of the sensitivity of the HAZUS model to local data at various magnitudes (Map 5.6).

5.41 Damage Losses at the City Level

At the city level, HAZUS™ shows different results based on the magnitude of earthquake for the City of Long Beach. For a 5.0 magnitude earthquake, HAZUS™ local data results in lesser total direct economic loss from buildings than that produced using default level data by more than $128 million (-17%). However, for an earthquake of magnitude 6.0, the losses are larger with local data by $456 million (17%) and for an earthquake of magnitude 7.0, the losses are larger with local data by $837 million (16%). Therefore, the percentage change in damage is not consistent with the magnitude of the earthquake.

The non-structural damage is more with HAZUS™ default data for the 5.0 magnitude earthquake and is largely responsible for the larger total loss and also larger building damage loss with default data for a 5.0 earthquake scenario. The percentage difference in loss from using local data versus default data is much smaller for the City of Long Beach for all the loss categories as compared to the City of Seattle. The percentage difference is higher for inventory loss – however, in real terms and as a percentage of total loss, inventory loss contributes only a small amount. Figure 5.1 – Figure 5.3 shows the results of loss from different scenarios.

The percentage of total loss that can be contributed to building damage (structural and non-structural) is lesser with real data than with default data for all three magnitude earthquakes and the difference is higher for a 5.0 magnitude earthquake. For a magnitude 5.0 earthquake, the lower building damage losses from default data are offset by larger losses due to content damage (Figure 5.3).

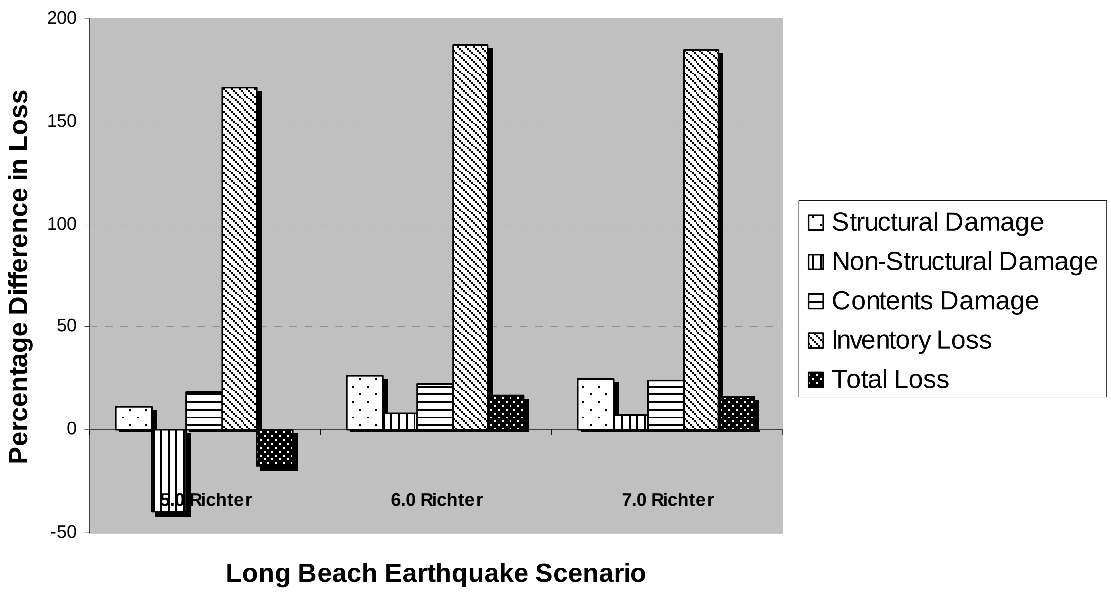


Figure 5.1: Percentage Difference in Loss Due to Building Inventory Change for 3 Earthquake Scenarios in Long Beach

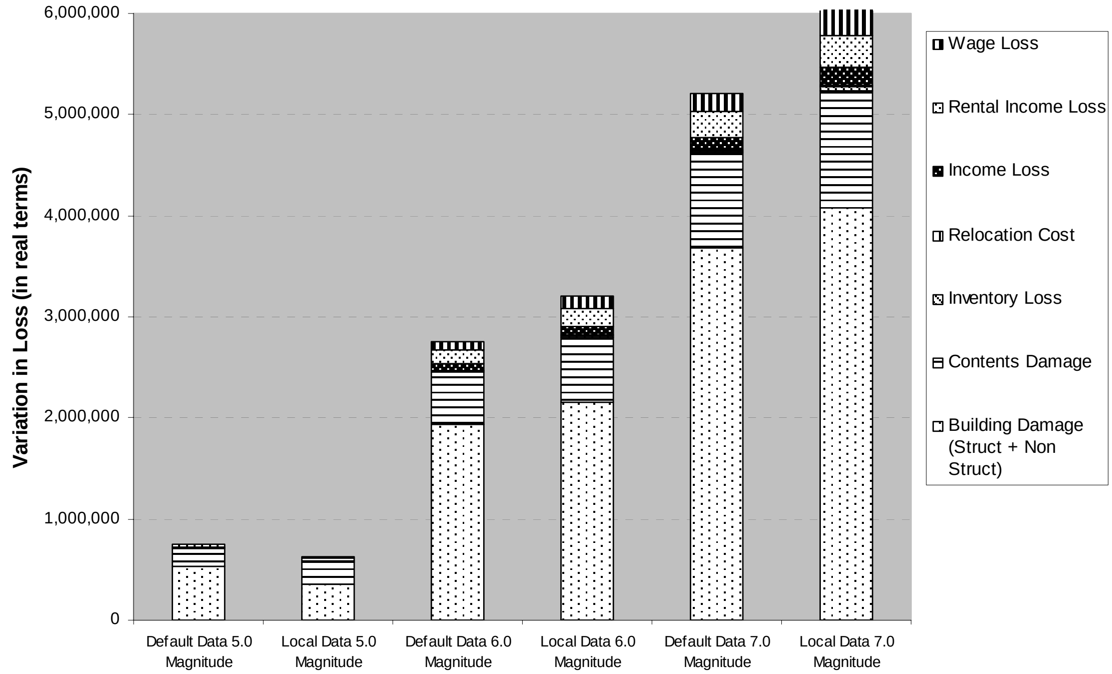


Figure 5.2: Variations in Loss Due to Building Inventory Change for 3 Earthquake Scenarios in Long Beach (in real terms)

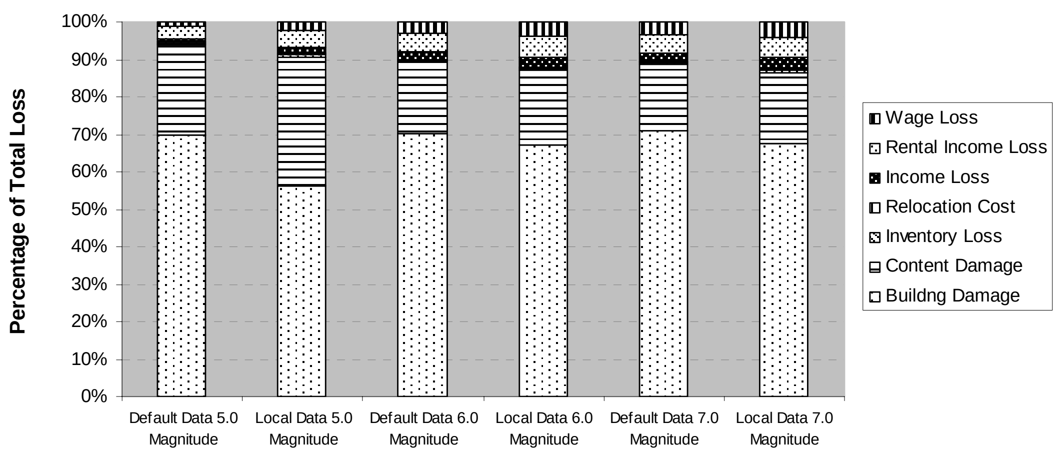


Figure 5.3: Variations in Loss Due to Building Inventory Change for 3 Earthquake Scenarios in Long Beach (as a percentage of total loss)

It is interesting to note that while the total loss is more when using default data for a 5.0 magnitude earthquake, the shelter needs and the number of people displaced is less with default data for an earthquake of magnitude 5.0. For all the other magnitudes, loss is less but shelter needs and number of displaced people are more with HAZUS™ default data (Table 5.7). The percentage difference for these losses is much higher than the total loss. In this case also there is also no consistent trend that can be seen.

Table 5.7: Variation in Shelter Needs for 3 Earthquake Scenarios in City of Long Beach
(assuming earthquake at 2 am)

Event	Shelter Need	Local Data	Default Data	Difference	%Difference
Newport Inglewood 5.0 Event	Number of Households Displaced	665	622	43	6.9%
	Number Needing Short-term Housing	180	169	11	6.5%
Newport Inglewood 6.0 Event	Number of Households Displaced	5,461	5,984	-523	-8.7%
	Number Needing Short-term Housing	1,574	1,737	-163	-9.4%
Newport Inglewood 7.0 Event	Number of Households Displaced	12,423	13,512	-1,089	-8.1%
	Number Needing Short-term Housing	3,595	3,939	-344	-8.7%

The number of casualties increases with the increase in magnitude but as can be seen in Table 5.8, there are no trends that can be inferred about the percentage difference at the three magnitudes. Table 5.8 shows the casualties for an event at 2 a.m. Similar trend is seen for events at other times that can be modeled in HAZUS™ (i.e. 2 p.m. and 5 p.m.). The number of ignitions, and the population and property exposed to fires is also lower with real data than with local data for the 5.0 and 7.0 earthquakes (Table 5.9). A similar trend was also seen in Seattle and may point to some algorithm bugs.

Table 5.8: Variation in Casualties for 3 Earthquake Scenarios in City of Long Beach
(assuming earthquake occurs at 2 am)

Event	Casualties	Local Data	Default Data	Difference	%Difference
Newport Inglewood 5.0 Event	Severity 1	362	131	231	176.3%
	Severity 2	54	14	40	285.7%
	Severity 3	5	1	4	400.0%
	Severity 4	10	1	9	900.0%
Newport Inglewood 6.0 Event	Severity 1	2,923	827	2,096	253.4%
	Severity 2	686	152	509	351.3%
	Severity 3	92	15	81	513.3%
	Severity 4	179	29	155	517.2%
Newport Inglewood 7.0 Event	Severity 1	6,564	1,979	4,585	231.7%
	Severity 2	1,743	448	1,295	289.1%
	Severity 3	250	53	197	371.7%
	Severity 4	488	101	387	383.2%

Table 5.9: Variation in Debris and Fire for 3 Earthquake Scenarios in City of Long Beach

Event	Loss Type	Local Data	Default Data	Difference	% Difference
	Debris				
Newport Inglewood 5.0 Event	Brick Wood and Others	No results from HAZUS™	64	Missing value	Missing value
	Concrete and Steel	No results from HAZUS™	57	Missing value	Missing value
Newport Inglewood 6.0 Event	Brick Wood and Others	483	299	184	61.5%
	Concrete and Steel	719	524	195	37.2%
Newport Inglewood 7.0 Event	Brick Wood and Others	905	600	305	50.8%
	Concrete and Steel	1,624	1,149	475	41.3%
	Fires				
Newport Inglewood 5.0 Event	Number of Ignitions	20	23	-3	-13.0%
	Population Exposed	237	542	-305	-56.3%
	Value Exposed (in thousand $)	15,889	38,367	-22,478	-58.6%
Newport Inglewood 6.0 Event	Number of Ignitions	31	20	11	55.0%
	Population Exposed	568	421	147	34.9%
	Value Exposed (in thousand $)	35,423	30,972	4,451	14.4%
Newport Inglewood 7.0 Event	Number of Ignitions	26	37	-11	-29.7%
	Population Exposed	286	1,523	-1,237	-81.2%
	Value Exposed (in thousand $)	16,052	78,665	-62,613	-79.6%

Thus by using HAZUS™, it is apparent that planners may underestimate health care and debris removal needs after a large earthquake. However, depending upon the magnitude of the earthquake other needs such as shelter requirements, and fire suppression may be overestimated or underestimated depending upon the magnitude of earthquake.

In summary, the data in Long Beach provided more limitations in truly understanding the impact on loss. However, even improving the square footage and occupancy information minimally resulted in a much larger percentage difference in loss estimates. In the case of Long Beach, the loss estimates were more for a 5.0 magnitude earthquake and less for a 6.0 and 7.0 magnitude earthquake with default data as compared to local data. Also a small improvement in square footage resulted in much larger discrepancies in displaced households, shelter needs, number of casualties, amount of debris and number of estimated fires along with the exposure on life and property.

The next section looks at the spatial patterns of loss differences and trends related to the type of census tract.

5.42 Damage Variation at Census Tract Level

The spatial variation of total loss is analyzed at the census tract level for the 7.0 magnitude earthquake only. In the City of Long Beach, the total loss is variable across the different census tracts – in some cases the losses are more with local data and in other cases, the losses are more with default data. For the most part, the overestimation and underestimation of losses follow the pattern of overestimation and underestimation of square footage as shown in Map 5.7. This may be due to lack of improvement in building structure information. Losses are more with default data for 48 census tracts and less with default data for 57 census tracts. For 47 of the above 48 census tracts above, the total square footage of the default data is more than the local square footage. A majority of these census tracts (87%) are also primarily residential. Out of the 57 census tracts where losses are more with real data, in 43 census tracts the square footage in the local data is also more and in 14 census tracts, the square footage in the local data was less than the default data. The 43 census tracts with losses and square footage greater with local data are mostly non-residential census tracts (i.e. downtown tracts, industrial tracts, tracts with airport, university, naval hospital, or tracts with mixed uses). The 14 census tracts that have larger default square footage but smaller default losses are primarily residential census tracts scattered throughout the city. No significant clustering is occurring near the epicenter to explain these discrepancies.

Thus, like the building inventory square footage, losses are overestimated by default data for residential census tracts but underestimated for downtown census tracts or tracts with a single use such as university, park, airport, etc. The patterns are similar when analyzed at the level of the building loss or content loss. One of the reasons why the loss discrepancy follows a similar pattern to the square footage discrepancy might be the fact the there was not much improvement in the building structure information (or in the occupancy matrices) due to the lack of availability of this information.

5.5 Conclusions

The selection of the City of Long Beach as a case study proved to be a useful one for many reasons. The City of Long Beach represented a city with a typical GIS implementation which had many GIS datasets available but with varying degrees of completeness and with inadequate metadata. This made it difficult to use some of the data. Furthermore, it provided various GIS data layers but could not provide the assessment data (created and maintained by Los Angeles County) because the City had limited access to it. Los Angeles County was not open to sharing the data. This can be a significant impediment to the use of local level data in models such as HAZUS™. For this research, data from a 3rd party reseller were used but this presented difficulties since such 3rd party resellers repackage the data in their own data formats and doing bulk data downloads were not easy.

Furthermore, the assessment data were itself very poor – i.e. they were incomplete, did not contain square footage and building characteristic information for tax-exempt properties, and did not contain reliable information on building type. For tax exempt properties and for missing information on other properties, the building characteristics information had to be inferred from various other GIS data (primarily building footprint layer) which were again not well collected or maintained. Furthermore, these sources of data only helped infer information on height and square footage and did not overcome the need for information such as building type, year built, building and content exposure, etc. Other commercial proprietary data are available for parts of the City but this was beyond the scope of this research. These data are often not available for the entire city and hence can assist to improve data only for a limited part of the city.

Information about content exposure was also not available through any local sources. Though building value was available in the assessment data, given the poor quality of the data and the large percentage of data missing, it was not used for this research. The average building and content value per square feet for each occupancy class based on default data was multiplied by the improved square footage and input into HAZUS™. This operation had to be accomplished outside of the software interface since the interface provided no capabilities for this.

The data reality for the City of Long Beach made the use of local data very difficult. A significant amount of time and effort was needed to prepare the data for input into HAZUS™. Various assumptions had to be made. Therefore, even with this investment of time and effort, it is not clear how much uncertainty could be reduced. As discussed, changing some of the assumptions above could significantly change the difference in default and local data. However, the improvement in square footage alone resulted in large differences, particularly for the breakup of the total square footage into different occupancy classes for the entire city. Therefore, while the overall data for City of Long Beach was improved only by 2%, the individual occupancy classes in the City of Long Beach showed much larger differences, both in real terms and in percentage terms.

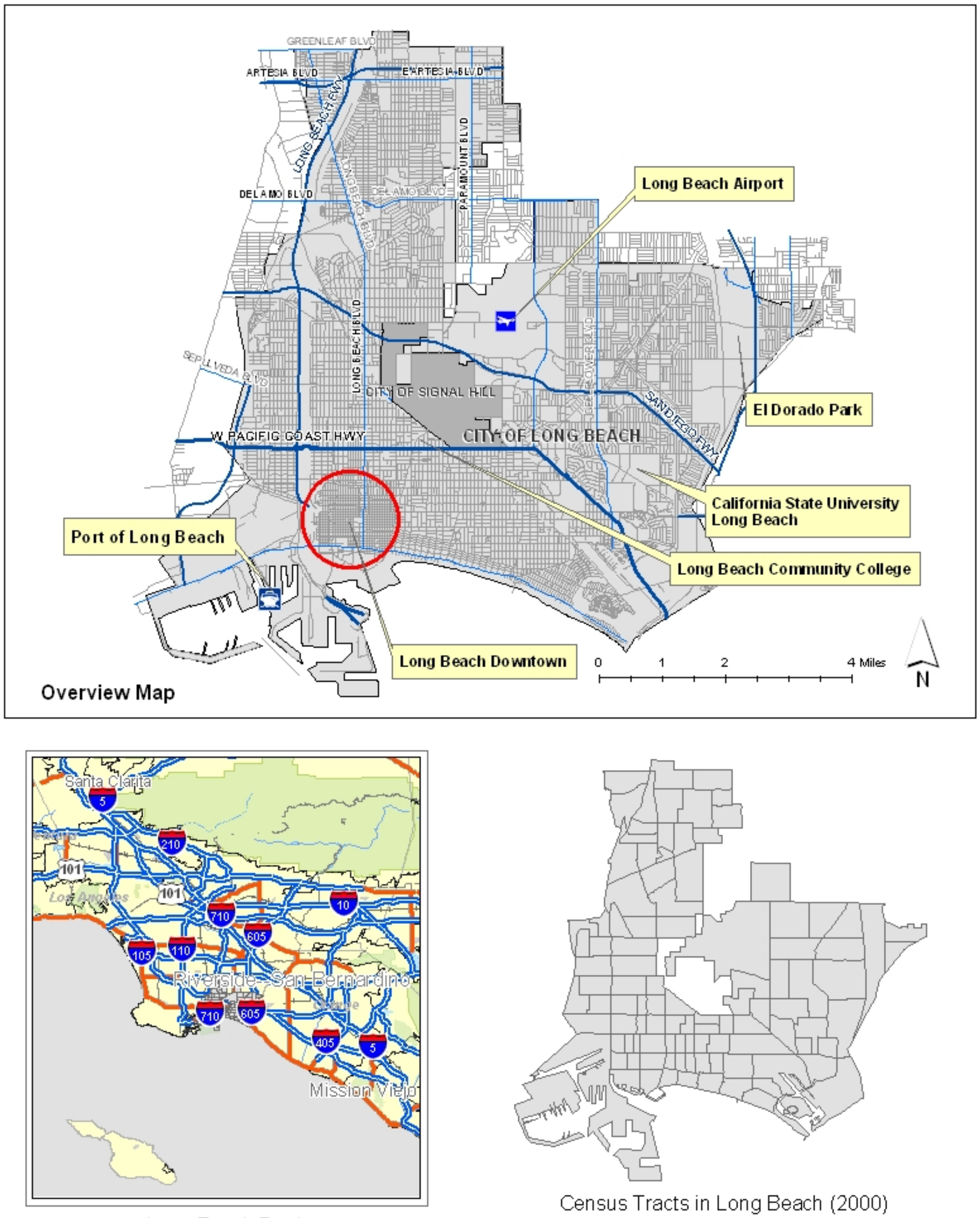
There was also a great deal of spatial variation based on the type of census tracts. While census tracts that were primarily residential census tracts were overestimated by HAZUS™ default data, non-residential tracts such as downtown and surrounding census tracts, mixed use census tracts, and primarily single use census tracts such as those containing airports, university, etc. were largely underestimated by HAZUS™. These census tracts were also more different both in real terms and in percentage terms.

The difference in building inventory significantly impacted the overall damage estimates (for both direct and induced losses). A 2% change in inventory led to a variation of difference in total loss ranging from a decrease of 22% for a 5.0 event to an increase of 11% for a 7.0 event. Significant variations were also seen in the indirect losses such as amount of debris, shelter needs, displaced households, and casualties. These have a tremendous impact in the use of HAZUS™ for any phase of the disaster management cycle - preparedness, response, recovery and mitigation.

The spatial variations in losses followed a pattern similar to that observed in the difference in building inventory between default and local data. Therefore, losses were more different for downtown and surrounding census tracts, mixed use census tracts, and primarily single use census tracts such as those containing airports, university, etc. The losses for these census tracts were less with HAZUS™ default data whereas the losses for residential census tracts were more. While some general trends existed, it is important to note that the trends cannot be completely described by the type of census tracts. Thus, two census tracts with similar land use patterns sometimes showed very dissimilar results with respect to change in losses based on local data versus default data.

The Long Beach case study also shows that there is no consistent pattern of change in the loss estimates based on the magnitude of the earthquake (i.e. the loss does not increase or decrease consistently with the increase or decrease in the earthquake magnitude). This could be partially attributed to the lack of improvement of the structure type information in the local data resulting in only minimal improvements to the occupancy to structure type matrices.

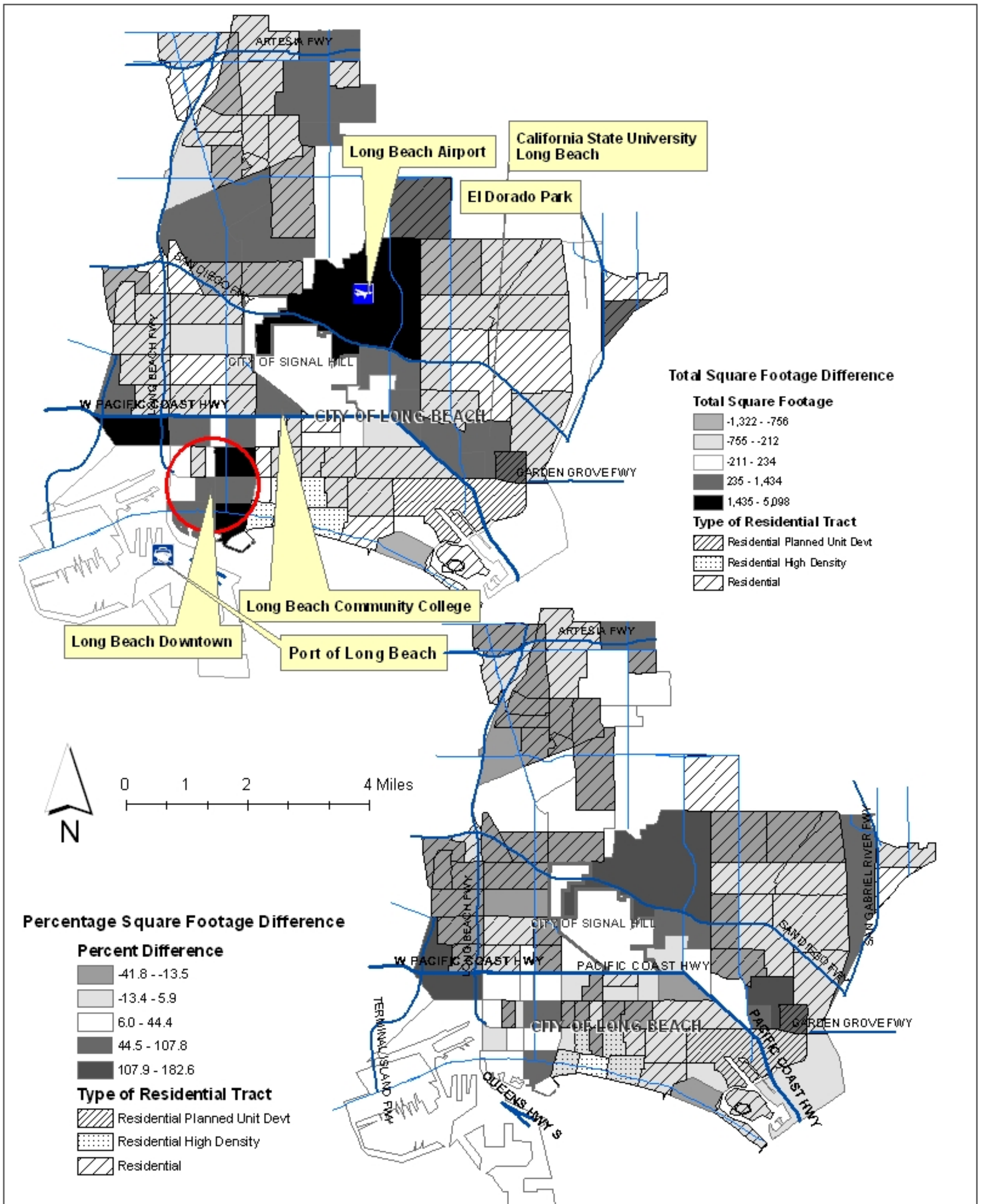
Thus, it is clear from the Long Beach case study that local data may not be the panacea for the use of integrated models for disaster damage assessment where local data are not reliable themselves. There are several challenges associated with using local data and even local data need significant improvement (through field surveys, proprietary building data or other estimation techniques). However, this improvement needs to be done in the planning and preparedness phase of the disaster. Without the improvement of data in this phase, if the City of Long Beach experienced a large earthquake and decided to use HAZUS™ in the response and recovery phase, it would have very little option but to use default data. This is also because local data are not easily available to the city. Furthermore the improvement in data involves a significant investment in time and effort. If the local data cannot be improved due to lack of expertise, resources, or other reasons, this research's findings can help decision-makers understand some of the uncertainties associated with using default data. Since the risk of making poor decisions is large through the use of default data, the risk can be reduced by at least understanding the level of differences.



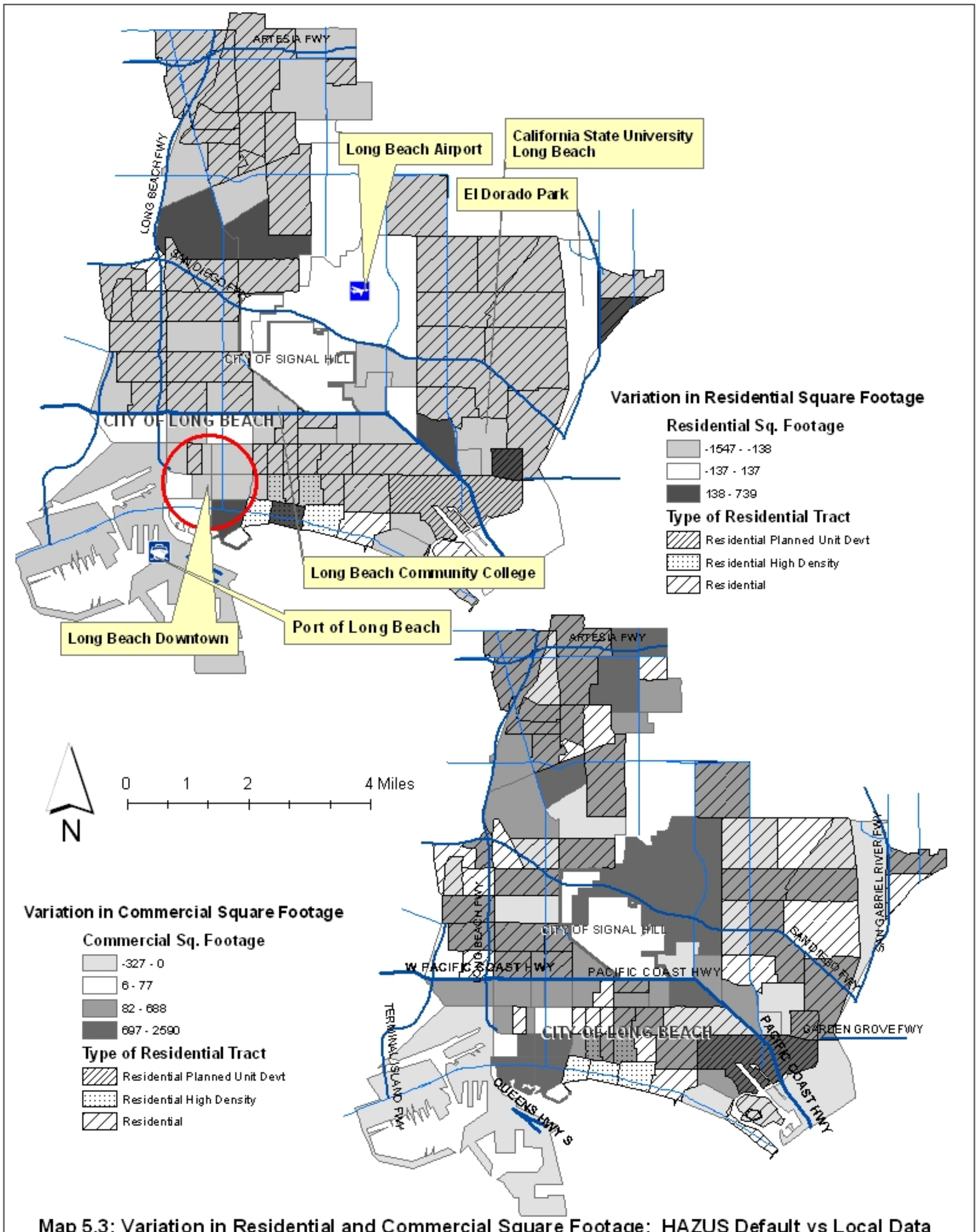
Long Beach Region

Census Tracts in Long Beach (2000)

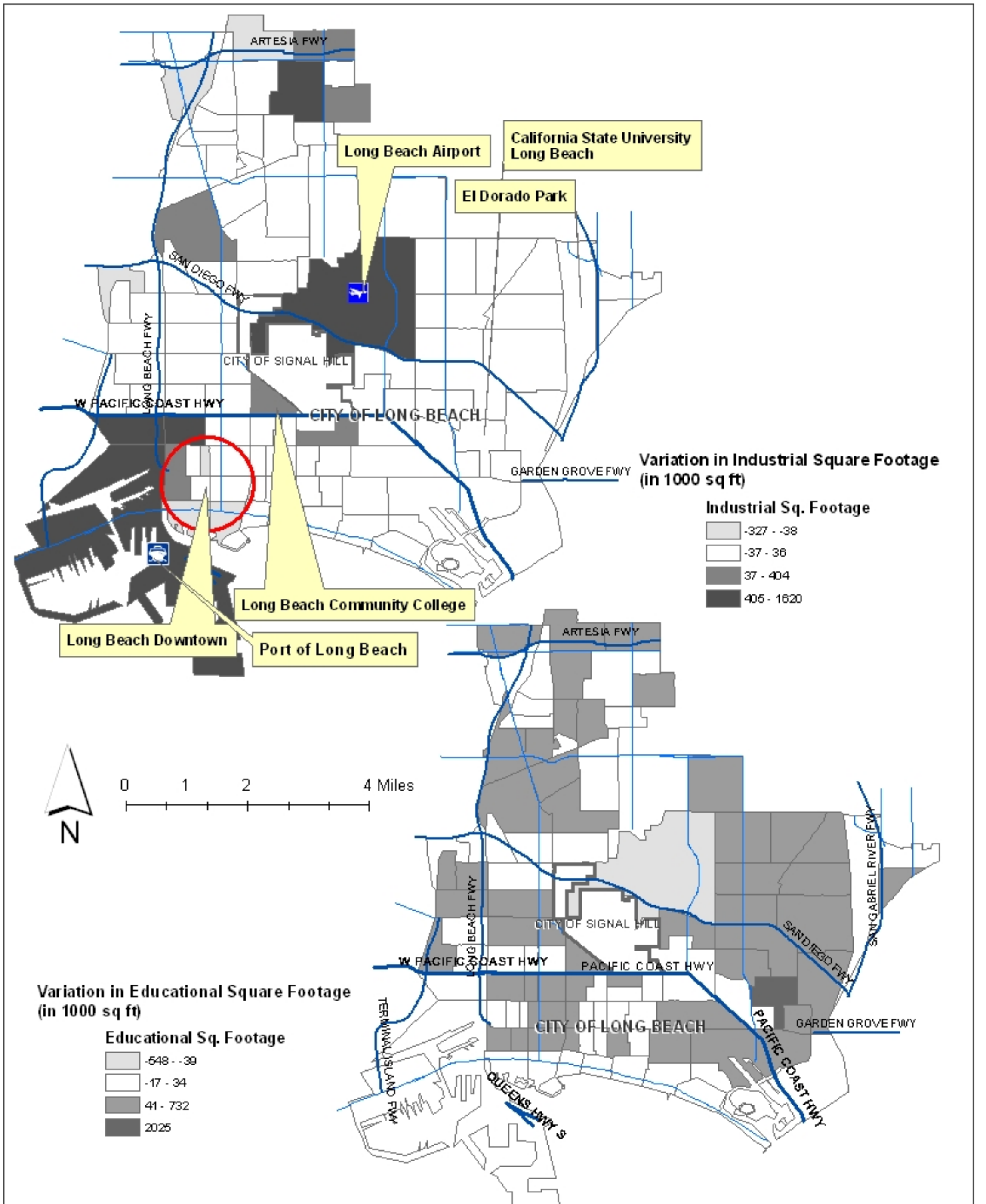
Map 5.1: City of Long Beach



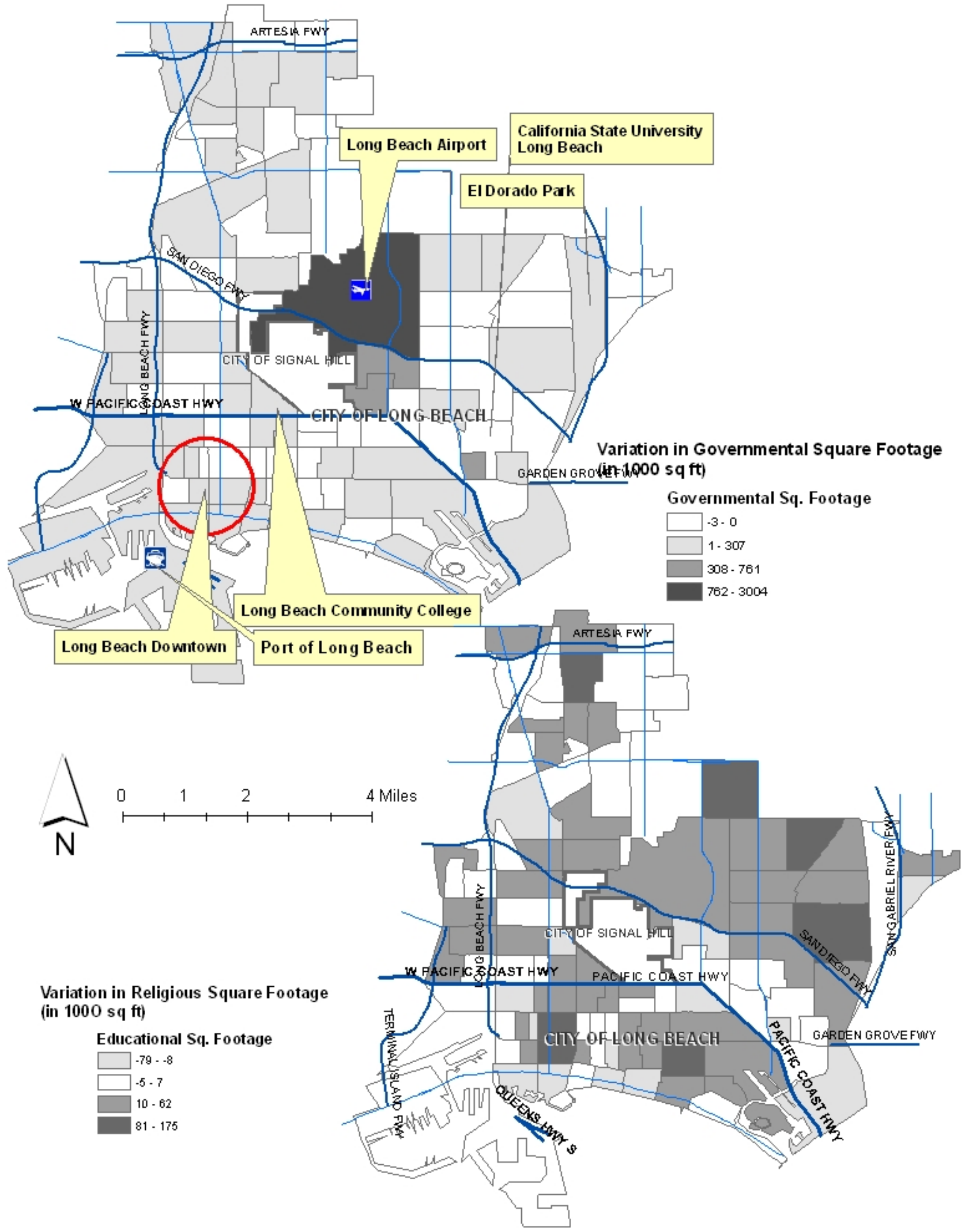
Map 5.2: Variation in Total Square Footage: HAZUS Default vs Local Data
City of Long Beach, CA



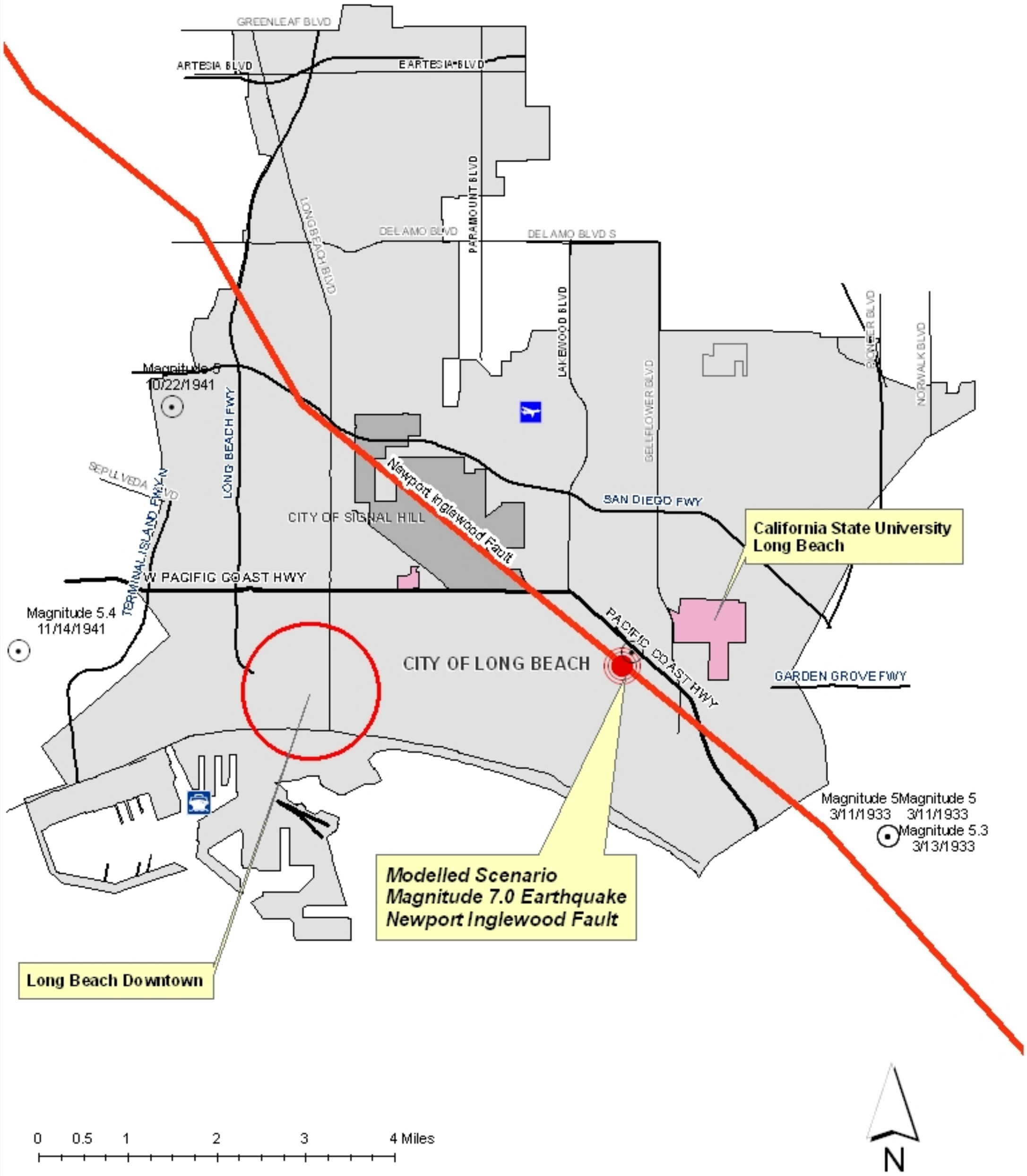
Map 5.3: Variation in Residential and Commercial Square Footage: HAZUS Default vs Local Data
City of Long Beach, CA



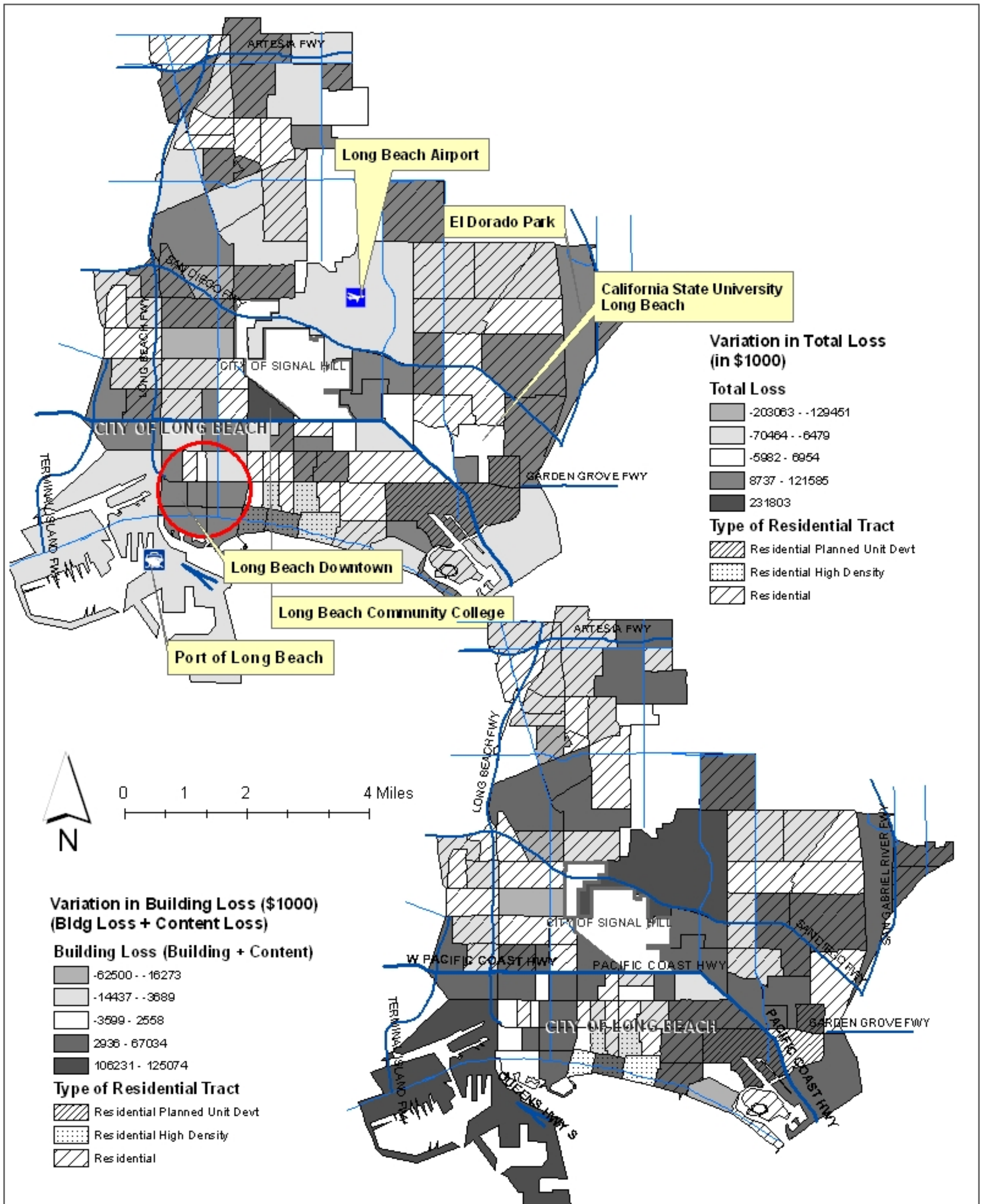
Map 5.4: Variation in Industrial and Education Square Footage: HAZUS Default vs Local Data
City of Long Beach, CA



Map 5.5: Variation in Governmental and Religious Square Footage: HAZUS Default vs Local Data
City of Long Beach, CA



Map 5.6: Scenario Earthquake and Historical Earthquakes
City of Long Beach, CA



Map 5.7: Variation in Loss Due to a 7.0 Earthquake: HAZUS Default vs Local Data
City of Long Beach, CA

Chapter 6: Comparison of Findings from Case Studies

6.0 Introduction

The central premise of this dissertation is discussed in Chapter 2. Integrated models are useful for public policy purposes, and their application for disaster management purposes can help decision-makers make more rational decision in an application area already wrought with uncertainty. However, integrated models are very data hungry and there is inherent uncertainty embedded in these models. Lack of good data (particularly about the built environment) is a large contributor of uncertainty in most integrated models. Assumptions about the built environment are made from data available through various sources and at various levels of aggregation making these models useful only at certain scales.

The HAZUS™ model is an integrated model representing the state-of-the-art in damage estimation and assessment. This model is being widely used at various levels of governance. The HAZUS™ model comes with many default datasets that are collected from various sources. The building inventory data, which is a large contributor of loss of life and property, is assumed in HAZUS™ on the basis of various datasets available nationwide such as census data and Dun and Bradstreet data. However, sources of data exist at the local level that can improve the building inventory and subsequently improve the output from the HAZUS™ model.

This dissertation set out to investigate 3 broad questions:

1) Given the data requirements of hazard assessment models for disaster damage estimation, what is the state of GIS in large cities for sustained use of integrated models at the local level?

2) How do default estimates for building inventory in HAZUS™ compare with local data for building inventory?

3) How sensitive is the HAZUS™ damage estimation model to improvements in building inventory data from local sources?

Question 2 and 3 above aimed to understand the need for local data and in light of the data reality (assessed through question 1 above), the ultimate goal was to address the needs for improvement of data in integrated models such as HAZUS™ and to understand the appropriate use of these models. In order to understand and answer the above questions, a survey was undertaken to assess the availability of local data in 19 large cities. Based on this survey, two cities (Seattle and Long Beach) were chosen for further case studies to better understand issues related to inputting local data into HAZUS™ and the sensitivity of the HAZUS™ earthquake model to local data. The City of Seattle represented a city with high investment in GIS and very good data at the local level whereas the City of Long Beach represented a city with a modest GIS implementation and modest data at the local level.

This chapter discusses the findings for each of the above questions. Each question will be discussed in a separate section (Section 6.2 – 6.4). In order to understand each question, inferences will be drawn based on the findings of both the survey and the case studies. The policy implications of the findings and need for future research will be discussed in the concluding chapter (Chapter 7).

6.1 Availability of Local Data

Based on the survey of 19 cities, the research clearly shows that GIS is widely diffused at the local government level of large cities. Most core layers such as parcels, tax assessment information (which includes a wide range of information very crucial to disaster management and damage assessment), street centerline, topography, and orthoimagery are available for a majority of the cities. However, information on building footprints, utilities, and critical facilities is still not as widely available and efforts need to be made to acquire these datasets for large cities since they can be very useful for disaster management purposes.

This research shows that although theoretically the datasets are available, in reality this does not always mean that these datasets are easily usable for the purpose of disaster damage assessment. While datasets such as parcels and road centerline are updated regularly, other datasets such as building footprints and orthoimagery are rarely updated and maintained. This research finds that although 12 cities out of 19 have building footprint data for the entire jurisdiction, and 3 cities have footprint information for a small part of the city, only 4 cities actually maintain these data on a regular basis. This is particularly important given the fact that for 9 cities, these data were collected in the nineties and for 3 cities the data were collected in the eighties. Furthermore, the building footprint data rarely carry any intelligence (i.e. building properties, use, etc). Therefore, they provide only limited usability for disaster damage assessment. The building footprint data are a crucial dataset for managing a disaster, assessing damage, recovering and mitigating from a disaster. In the aftermath of a disaster, the recovery effort needs to track the damage and condition of each and every building and a detailed GIS inventory can be very useful. The building footprint data proved to be very useful in the case of City of Long Beach since they were used to supplement the poor quality assessment data with respect to square footage, height, etc. However, since these data were not updated regularly for the City, they provided limited value and were not able to reduce uncertainty significantly. The City of Long Beach building footprint data also had no attribute information such as height, etc. When compiled photogrammetrically, this information is usually captured and if this was available for the City of Long Beach, it would have allowed much better calculation of number of floors and consequently the square footage.

The tax assessment data are important and provide critical information regarding the use of parcels, property values, and information about the building characteristics such as height, age, square footage, building material and type of construction, etc. This is perhaps the most comprehensive dataset that provides information about the built environment that is available for most large cities nationwide. The interview/survey as a method of investigation had various limitations for this research. While it informed about what datasets were available, it did not inform well on the quality of the data and the issues associated with actually using them in a model. The case studies were useful in understanding the latter. Therefore, although the survey indicated that 16 out of the 19 surveyed cities have all tax assessment information linked to the parcels, and the rest had partial data linked, the case studies revealed that the quality of assessment data can be variable. Thus, for both City of Seattle and City of Long Beach, the tax assessment data carried many fields of information but not all of them were fully populated or even partially populated. Likewise, even for the fields that were populated, the quality of information was variable and often not easy to map to HAZUS™ fields and values. Furthermore, not all parcels had a match with corresponding assessment data because the assessment information was often collected by an agency that was different from the one that created the parcel data and no rigorous quality control measures were in place. Both the case study cities showed that many parcels were missing corresponding assessment data. While some of this can be due to different vintages of parcel data and assessment data, the degree of mismatches warrants more stringent quality control and accuracy standards.

The two case study cities also revealed that tax assessment data are not always complete. Many assessors do not collect information on tax-exempt properties such as schools and educational institutions, publicly-owned parcels, and religious use parcels. This can seriously impact the use of assessment data for the purposes of disaster management. Thus, in the City of Seattle, data on tax-exempt properties (with the exception of the University of Washington properties) were tracked meticulously. However, for the City of Long Beach information about these properties was not collected and hence any damage assessment required a lot of assumptions, leading to more uncertainty in the results.

The assessment data are usually collected at the parcel level. However, often there are multiple buildings in a parcel or multiple owners within a building. Appropriate modeling of tax assessment data should capture such information in a way that data can be further broken up. This was done in the data for City of Seattle. However, similar breakdown was not observed for the City of Long Beach making it more difficult to prepare for HAZUS™. As mentioned above, the availability of data by building are important since damage occurs by buildings and not by parcel and it is important to track the data at this granularity, particularly for large parcels with multiple buildings such as university campuses, hospital complexes, public housing complexes, trailer parks, etc.

The case studies also show that just because assessment data were available and complete, it did not imply that they were easy to use in the HAZUS™ model. HAZUS™ requires many classifications of occupancy and building type. While the occupancy (i.e. use) could be matched to appropriate occupancy classes in HAZUS™ without too much difficulty, the building type classifications needed by HAZUS™ (as discussed in Chapter 1 and 2) were not easily available at the local level. While the City of Seattle had fairly complete information on the type of structure, there were only a few types of buildings and there was no way to match them to the 36 specific HAZUS™ building types. For the City of Long Beach, not only did the domain of values not match HAZUS™ building type, the field that carried this information was only populated for 16% of the records.

Another piece of information required by HAZUS™ is information on building exposure and content exposure. While building exposure information is available from the tax assessment data, there are very few sources for content exposure as contents (other than some kinds of personal property) are not taxed by local assessors. The building exposure in assessment data is tracked through the assessed value field, which is not a correct reflection of the market value. Some assessors maintain a multiplier field that that can be used to calculate market value from assessed value. However, the use of assessment data can lead to a large underestimation in building exposure in cases where the assessment data do not carry information on tax-exempt properties. The content exposure is very difficult to get from local sources. In the absence of these data (i.e. content and building exposure) at the local level, it is virtually useless to improve the data in HAZUS™ through local sources.

This research also shows that organizationally, many cities have advanced GIS programs with enterprise implementations (centralized data collection and GIS management). Enterprise GIS is seen in 13 of the 19 cities. This is important since models such as HAZUS™ need integration of data from many different government departments. An enterprise implementation implies data standardization, common platforms, and easy access to data. However the utilities data were not always part of this enterprise and can pose significant challenges in integrating with the rest of the data. For many cities, the tax assessment data are also not part of this enterprise and hence it is not easy to acquire or use for hazard assessment.

Both the City of Seattle and the City of Long Beach had enterprise GIS implementations but the tax assessment data were created and maintained by their corresponding counties (King County and Los Angeles County). While the City of Seattle had easy access to tax assessment data from King County, and provided it in easy to use format for this research, the City of Long Beach did not. The City of Long Beach had access to the tax assessment data record by record through software provided by the County. They could also do some data exports by zip codes in a cumbersome process. The Los Angeles County was not interested in sharing the data either (without a large charge) and hence it was difficult to access this dataset. While it is understandable that the County sells these data, the lack of easy access to these data to municipal agencies can be a serious obstacle for damage assessment, particularly in the aftermath of a disaster. Since accessing and preparing local data for one jurisdiction alone can be so cumbersome, the task of accessing and preparing data for multiple jurisdictions for a regional analysis requires a lot of effort. Inter-governmental data sharing policies should be clearly defined in planning and preparedness stage of the disaster management cycle. An enterprise GIS implementation can help in establishing these policies by providing single points of contacts rather than establishing contracts with many different departments in a single organization. Furthermore, applications that allow easy access and manipulation of data are just as important as the availability of data. In the case of Long Beach, although ultimately the data were available to the City, they were not easy to access, or distribute and posed a serious obstacle for their use.

Good metadata is also needed for analysis and data-sharing and results of the survey show that only 5 cities out of the 19 cities have metadata that complies with any kind of national standard. Most cities have “home-grown” metadata. In the case of City of Seattle, the data had very detailed metadata and hence they were much easier to use. This was not the case for the City of Long Beach. The purpose of using local level data is to reduce uncertainty from models. However, using data with no information on accuracy, lineage, creation methodology, etc. can actually add to the uncertainty rather than reduce it and therefore, the need for metadata cannot be underestimated.

This research also shows that although many cities have advanced GIS programs, the use of GIS for disaster management has been limited. The use of GIS has largely been in the response and recovery phase (i.e. after a disaster has struck). Although some examples of use of GIS for planning, simulations, and proactive analysis to understand the likely impacts of hazards exist, such applications are few and restricted to a handful of cities. Furthermore, only respondents from 6 out of the 19 cities had ever heard of HAZUS™ and none of them had ever used HAZUS™. The lack of use of GIS (and tools such as HAZUS™) in preparedness and planning phase could be because of the low priority given to planning for unlikely catastrophic events or the lack of any imminent threat from such events for some of the surveyed cities. It could also point to the low degree of interaction between disaster managers and GIS experts – since the use of HAZUS™ requires considerable GIS expertise, it is unlikely that the tool is being used in the surveyed cities without the knowledge of GIS staff (particularly if local data is to be integrated).

In summary, this dissertation finds that there is a proliferation of core GIS data at the local level for use in damage assessment models. A dataset that can be very useful and not very widely available or well-maintained is the building footprint data. Data on utilities and critical facilities are also lacking. However, one key dataset that needs significant improvement, in terms of completeness, standardization, and data modeling is the tax and assessment data. Although there is a proliferation of data, the quality of data is more important and cannot be easily assessed through surveys. The case studies reveal that local data, particularly information on type of building is incomplete, inaccurate and not easy to use in HAZUS™. One piece of information that is hard to find at the local level is the content exposure (and to some extent the building exposure), the lack of which makes it virtually useless to improve data from local sources in HAZUS™. Finally, in order to use GIS and tax assessment data at the local level, it is critical to have good data-sharing policies and metadata in place before a disaster. For analysis at the regional level, it is also important to have some data standardization.

The next section discusses the variation of building inventory based on local data from the building inventory based on default data in HAZUS™. This will be followed by a discussion of the sensitivity of losses from HAZUS™ based on local level data vs. default data.

6.2 Variation of Data in HAZUS™: Local vs. Default

In the previous section, this research established that although a lot of data are available at the local level for disaster management and particularly damage estimation, the tax assessment data need improvement to be useful for disaster management purposes. However, even when the tax data are complete and accurate, they are difficult to input into HAZUS™ since the level and domain of classifications at the local level for occupancy and building types are not per HAZUS™ classifications. Furthermore, values for dollar exposures are difficult to attain at the local level (particularly content exposure). This section deals with the variation of local level data from default data for building inventory in HAZUS™. To understand this, two case studies were undertaken for the City of Seattle and the City of Long Beach. The purpose was to analyze the variation at various levels, for the city as a whole and for the smallest level of analysis possible in HAZUS™ (i.e. census tracts). The variation in total square footage and in various occupancy classes is assessed both for general occupancies (such as residential, commercial, etc) and specific occupancies (such as single family residential, multi-family, retail, wholesale, etc.).

The tax assessment data (along with the corresponding parcel data) were the most significant source of data for this research but were supplemented by other GIS datasets (such as building footprints, orthoimagery, etc.) to improve the data and to validate the local data where needed. It is important to note that the quality of data were variable for the two cities. The tax assessment data for the City of Seattle were modeled better to take into consideration the assessment by buildings rather than parcels (since there can be more than 1 buildings in a parcel). Furthermore, the Seattle data were much more extensive, complete, and reliable. The data for the City of Seattle also encompassed assessment of properties that were tax-exempt, such as educational, governmental and religious uses (with the exception of University of Washington). On the other hand the tax assessment data for the City of Long Beach proved to be very sketchy, incomplete and unreliable. The assessment data for the City of Long Beach did not have any information for tax-exempt properties. Hence a lot of assumptions had to be made to infer missing data in the case of City of Long Beach. Therefore, while comparing the two cities, it is important to keep in mind that the results from the City of Seattle are more reliable than the results from the City of Long Beach. For both cities, however, the local data classifications did not conform well to HAZUS™ classifications for occupancy and building type and it was not easy to conflate the local classifications to HAZUS™. This was discussed in the previous section.

This research finds that at the level of the city, HAZUS™ default data underestimate the building inventory square footage for both the cities researched. The degree of underestimation for Seattle (33%) is much larger than the underestimation for Long Beach (2%). The local data for the City of Long Beach are itself underestimated since the assessment data did not contain information on parcels used for educational, religious and governmental uses. Therefore, various assumptions were made to calculate some basic information such as square footage, height etc. For all buildings in parcels that didn't have any height and square footage information, the area of the building footprint was used as the square footage and the height was assumed to be one storey. Furthermore, because of the quality of data, a lot of records were eliminated when input into HAZUS™. Therefore, if HAZUS™ did not eliminate any records for the City of Long Beach, the total default square footage would be underestimated by 8%. Also, for data missing assessment information (particularly height and square footage information), if the number of stories were assumed to be 2 stories rather than 1 storey (i.e. 36 million square feet of data), and all unknown uses were mapped to some valid HAZUS™, the default data would be underestimated in HAZUS™ by 23%. These variations in assumptions as discussed above are not unreasonable, but unfortunately cannot be validated.

Both cities show a significant variation at the city-level when the total square footage is broken down into different occupancy classes. Figure 6.1 shows the variation in percent difference for the general occupancy classes for both cities. The residential and agriculture occupancy is overestimated by HAZUS™ default data for both cities. All other occupancy classes (i.e. commercial, industrial, religion, government and education) are underestimated by HAZUS™. Even though the percentage difference for the residential occupancy class is not too high for both cities, it translates to a very high absolute value and has a large impact on the overall difference at the city-level. For example, if residential occupancy is removed from both cities, the total square footage is underestimated by 172% percent for the City of Seattle and 143% percent for the City of Long Beach rather than 33% and 2% respectively. The difference in the commercial occupancy is more in the City of Seattle, perhaps because of better data and also because the City of Seattle has a larger and denser downtown than the City of Long Beach. The difference in industrial occupancy is much more for the City of Long Beach than in the City of Seattle, reflecting the different types of cities. However, it is important to note that given the uncertainties in the Long Beach data, it is not possible to draw too much comparative inferences (or even conclusions on why differences exist). However, it is instructive to note some trends between both the cities with greater emphasis on reliability of the Seattle data.

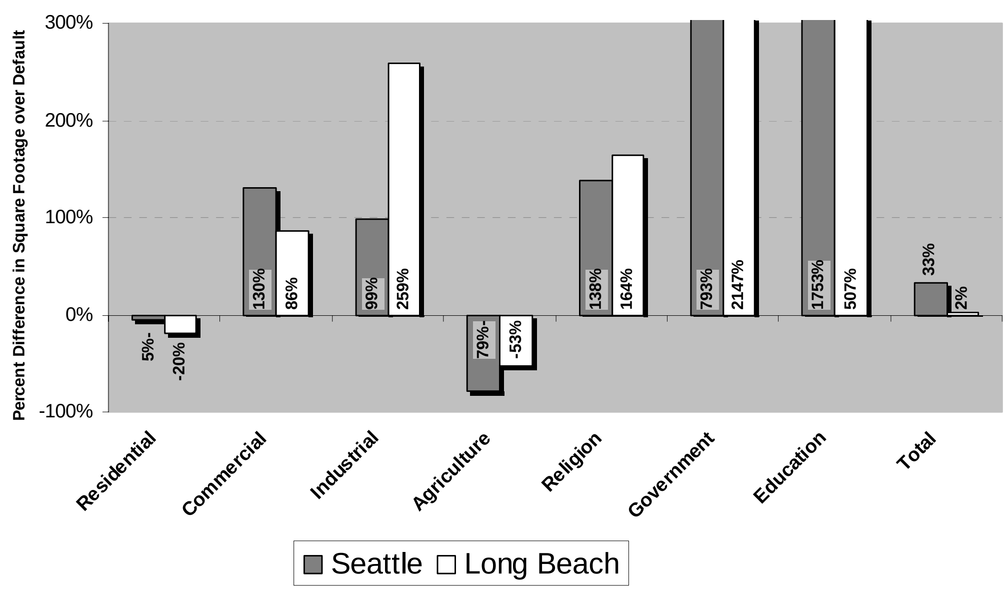


Figure 6.1: Variation in Percent Difference of General Occupancy Classes

Within each of the general occupancy classes, much difference exists in both cities when the general occupancy classes are broken into specific occupancies. Figure 6.2 shows the variation for various categories for the residential occupancy class. Single family residential occupancy, mobile homes and institutional dormitories are overestimated by HAZUS™ in both cities. Mobile homes and institutional dormitories may be overestimated because of poor recording of data in the local sources – for example, the institutional dormitories that may exist in educational facilities are often classified under the education occupancy. Both cities show more triplex and quads than is shown by HAZUS™ defaults. On the other hand, the trend is in contradiction for duplex and apartments in the two cities (i.e. one city shows an overestimation and another shows an underestimation for the two occupancy classes). It is important to note that

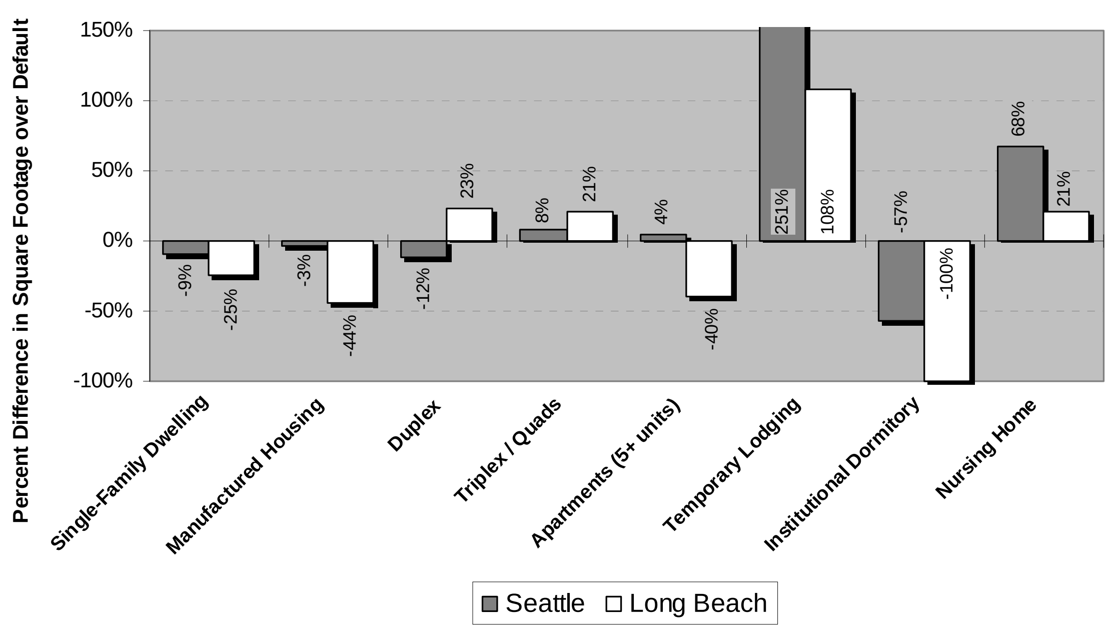


Figure 6.2: Variation in Percent Difference of Residential Specific Occupancy Classes

temporary lodgings (hotels, motels, etc) and nursing homes are largely underestimated by HAZUS™ default data.

Figure 6.3 shows the variation in difference for various categories comprising the commercial occupancy. All the sub-categories of the commercial occupancy are underestimated by HAZUS™ for both cities with the exception of personal/repair services and banks which are overestimated in both cities.

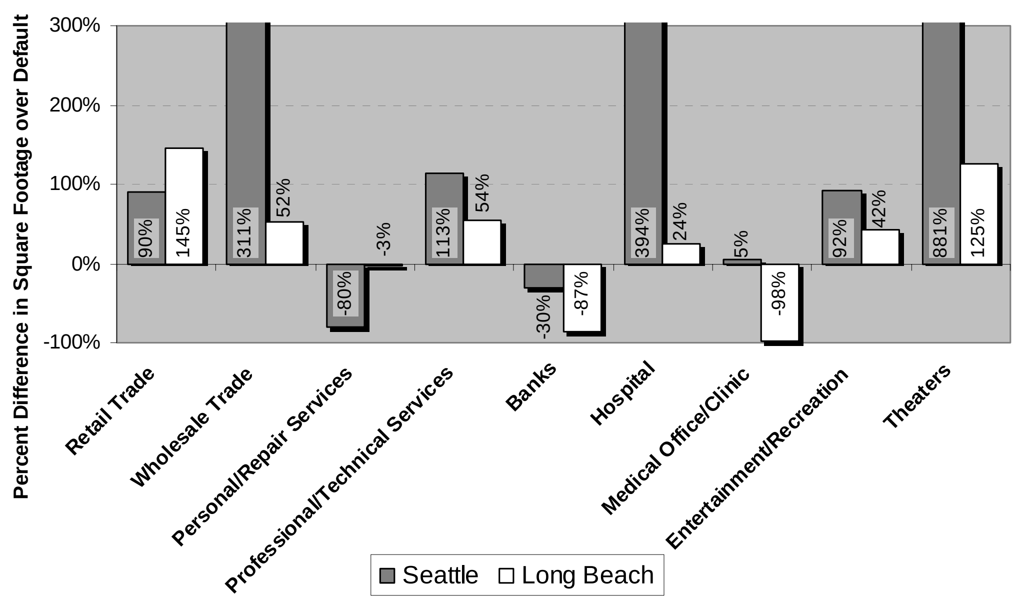


Figure 6.3: Variation in Percent Difference of Commercial Specific Occupancy Classes

Medical offices and clinics are overestimated for the City of Long Beach and this could be an indication of poor local data classifications for this class. For both cities the underestimation is high (both in real terms and in percentage terms) for retail trade, wholesale trade, professional and technical services, hospitals, entertainment and recreation, and theaters. The underestimation of these is more for the City of Seattle than for the City of Long Beach. This may point to the difference in the nature of the two cities (the City of Seattle being a large regional city with a prominent downtown whereas the City of Long Beach being a city that is in the larger metropolitan region of Los Angeles). Furthermore, as discussed before, the quality of local data for Long Beach could also contribute to the differences. Since parking garages are missing in the HAZUS™ default inventory data, this is not analyzed here.

Figure 6.4 shows the variation of difference for the industrial occupancy class. For the City of Long Beach, all the sub-categories of the industrial occupancy class are underestimated whereas, for the City of Seattle, foods, drugs and chemicals, and metals and mineral processing and construction are overestimated by HAZUS™.

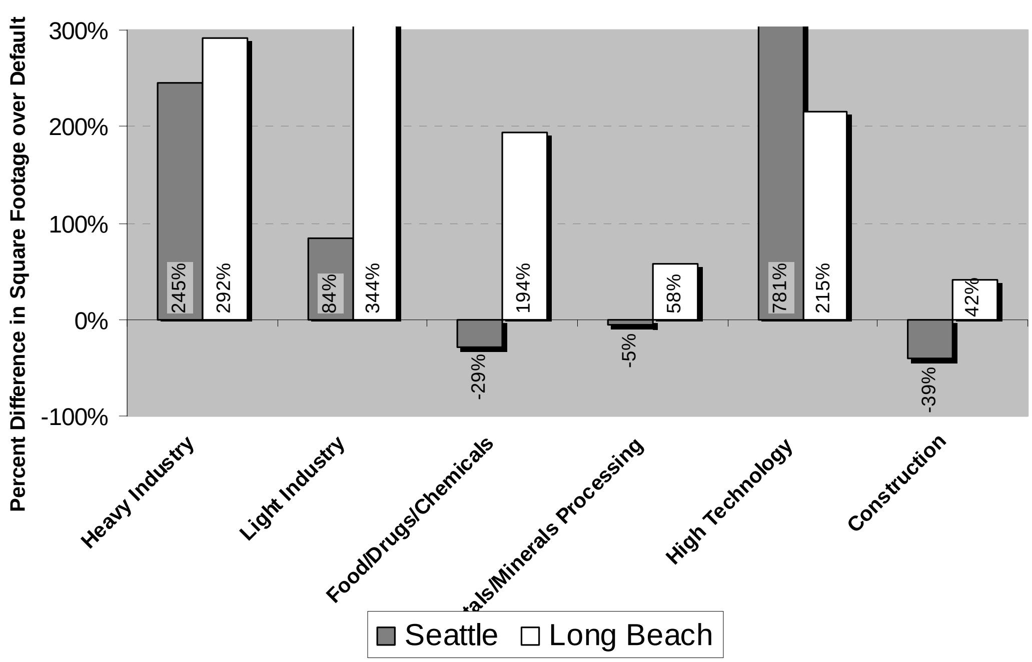


Figure 6.4: Variation in Percent Difference of Industrial Specific Occupancy Classes

Although not validated, one of the reasons for this overestimation could be that Dun and Bradstreet data may be recording all the square footage for some companies at their corporate headquarters and hence show larger square footage in Seattle even though the square footage is not physically located in Seattle. The data for breakups for government and education occupancy classes are either not available for one or both cities or not reliable enough to analyze at this degree of granularity.

Since the building count for the HAZUS™ default is based on the square footage divided by the average building size for that particular occupancy, the square foot variation is somewhat indicative of the variation in building counts. However, the average size of buildings in different occupancy classes can themselves be different from the reality for many different occupancy classes, as shown in Appendix E, Table 1. A comparison of the average sizes for the various occupancy classes shows that the single family residential house is fairly well assumed for HAZUS™. The average size of a mobile home in HAZUS™ is about 1100 sq ft but based on the data from the two cities, the average size of a mobile home is 4567 square feet in Seattle and 11318 square feet in Long Beach. These results indicate that for both of the cities, the number of mobile homes is underestimated by local data. In reality, in both cities, large parcels are designated for use as a mobile home park with many units in them. Therefore the local data reflects the total square footage of all the units in the mobile home park and not of individual units as in the case of default data. The size of duplexes is also overestimated by HAZUS™ while triplexes/quads are well assumed by HAZUS™.

An interesting observation is the size of hotels. The average size of hotels according to HAZUS™ default data is 177K square feet for the City of Seattle, and 185K square feet for the City of Long Beach. However, the local data shows the average size to be 46K and 17K respectively. This finding may show that the Dun and Bradstreet data captures more information on very large facilities and the smaller facilities are not recorded in this dataset. Furthermore, the City of Seattle may have more large hotels than the City of Long Beach which is quite intuitive. However, most cities also have an abundance of small hotels and motels and hence the average size is a lot lower.

Similar is the case with buildings used for retail trade. As per HAZUS™ default data, the average size of a retail establishment is 339K square foot for the City of Seattle and 723K square feet for the City of Long Beach. However, the average size of a retail building is 10K square feet and 7.5 K square feet for the two cities respectively. In terms of count, according to the HAZUS™ default data, the City of Seattle has only 47 retail establishments and the City of Long Beach has only 9 retail buildings, but in reality the local data shows 2987 and 2106 retail establishments in the two cities respectively. The average size of a wholesale trade establishment is also lower than HAZUS™ defaults.

Buildings classified as professional and technical services are also smaller in number in both cities in the HAZUS™ default data as compared with local data and average size of these buildings is far smaller than the HAZUS™ defaults. However, there are far less banks in the two cities in comparison with the HAZUS™ and this discrepancy is difficult to explain when compared to other uses. The size of a heavy industrial facility is half the assumed average size in HAZUS™ and this is the same for both cities. However, the light industry building in the City of Seattle is half that of the light industry building in the City of Long Beach. The size of religious building in both cities is also less than half the default size and this is in conformity in both cities.

There are also far more educational establishments than shown in the HAZUS™ default data in both cities. The building inventory data for HAZUS™ shows only 2 grade schools in Seattle and 2 grade schools in Long Beach and both these schools being very large buildings. However, the local data shows there are far more schools but they are much smaller in size. In fact, HAZUS™ also comes with additional inventory of essential facilities (which includes schools, medical care facilities and emergency response facilities). This inventory has 186 schools in the City of Seattle and 124 schools in the City of Long Beach. Therefore, it is clear that the default counts and the square footage are significantly different from the local data.

Since the City of Long Beach did not have good structure type information and much of the data were inferred, the analysis of the structure type information is not undertaken here. Furthermore, because no content exposure information was available for both cities, all exposure information for the two cities was simply improved to reflect better square footage data and hence also not suitable for analyzing in great details.

So far, this dissertation has analyzed the variation in local data with respect to the default data in HAZUS™ for the city as a whole and for the various occupancy classes within the entire city. Next, it looks at the spatial variation across the two cities to see if there are any spatial trends in the two cities based on census tracts. This dissertation finds that spatial trends are a little more apparent in the City of Seattle. In Seattle, HAZUS™ overestimates the square footage in less census tracts (34%) and underestimates the square footage for more census tracts (66%). In the City of Long Beach, a reverse trend exists – with more census tracts overestimated (58%) and less underestimated (42%). However, in both the cities, the degree and range of underestimation is much larger than the overestimation.

In general, in both cities, the residential occupancy is better estimated for the entire city, as opposed to commercial and industrial occupancies in most census tracts. Also, tracts that are primarily residential in character are usually better estimated than tracts that are largely non-residential single-use (such as commercial, industrial, educational, etc) or mixed use. Special use census tracts such as those comprising universities, airports, large parks or recreational facilities also show large discrepancies. Furthermore, for residential tracts, commercial, industrial, and educational uses are largely underestimated. However, for both cities, the patterns of overestimation and underestimation cannot be fully explained by the type of census tracts. In other words, two very similar census tracts can show reverse trends (one is overestimated and the other is underestimated). A better understanding of the source data could explain many of the discrepancies but such an understanding of the source data is not provided in the HAZUS™ manual.

In the City of Seattle, there is a large concentration of underestimation in the downtown census tracts and the census tract comprising the University of Washington. The top 12 census tracts in Seattle represent 111 million square feet of underestimation which accounts for 74% of the overall underestimation in Seattle. Although the downtown and surrounding census tracts in Long Beach also show an underestimation, this is not as stark as in Seattle. This might reflect the difference in the character of the downtown in these two cities.

In summary, it is obvious that the default data in HAZUS™ are a poor reflection of the reality. Better local level data are available for the City of Seattle and reveal that default data are very different from the local data when one compares the total square footage for the entire city. The City of Long Beach shows less variation in the total square footage for the entire city - the various assumptions used to improve the local data for the City of Long Beach make these data less reliable. However, for both cities, there is a great deal of variation in data at the various general and specific occupancy classes. There is also significant spatial variation in the local data from default data. These variations are smaller in residential census tracts but are very large in commercial, industrial, and special use census tracts (such as downtowns and their periphery, airports, universities, large hospitals, etc.). However, there is no consistent pattern of difference that can be explained by the type of census tract alone. It is also clear from this research that the quality of local data is variable and various assumptions can be made to improve the data. But this can also add to uncertainty and in these cases the local tax assessment data may not be the most appropriate source for improving HAZUS™ default data.

The next section focuses on the variation in damage estimates and output results from HAZUS™ based on local data and default data.

6.3 Variation in Damage Estimates from HAZUS™: Local vs. Default

For both cities, local data were input into HAZUS™ using the Building Inventory Tool (BIT). Three scenarios were modeled for each city at the same epicenter with three different magnitudes (5.0, 6.0 and 7.0 on the Richter scale) using default data and local data. The events in Seattle were on the Seattle Fault Zone North Trace with an epicenter near to the downtown (Map 4.5). The events in Long Beach were modeled on the Newport Inglewood Fault, with an epicenter closer to the airport (Map 5.6). Therefore for each city, 6 scenario events were modeled and losses (both direct and induced) were analyzed for the entire city as well as for the individual census tracts that comprised the city.

In Seattle, the percentage difference in loss between local data and HAZUS™ default data increased with the increase in magnitude of the earthquake. However, the same was not seen for the City of Long Beach. The degree of difference was a lot greater for the City of Seattle than for the City of Long Beach. This could be a reflection of the larger difference in inventory for the City of Seattle than the City of Long Beach. Thus, in Seattle a 33% difference in total square footage resulted in a 130% difference in the total loss, whereas in Long Beach, a 2.3% difference in total square footage resulted in 11% difference in total damage for the for a 7.0 magnitude event. It is to be noted that damage losses could vary based on the location of the scenario earthquake – for example, for the City of Seattle, the modeled epicenter was close to a dense downtown, whereas for the City of Long Beach the scenario epicenter was in a relatively less dense area. Hence the location of the epicenter could contribute to some of the above difference observed.

In Long Beach, the non-structural damage loss was more when using default data for the 5.0 event and this contributed largely to the greater losses for this magnitude with default data. It is not clear why the non-structural damage loss was greater for default data for a 5.0 event than for a 6.0 event. For all three scenarios, the loss due to content damage was more with HAZUS™ default. While in the City of Seattle, the percentage of total loss that could be contributed to building damage (structural and non-structural) was more with local data for all three magnitude earthquakes, this trend was not seen in the City of Long Beach. In Long Beach, for the 5.0 magnitude earthquake, the local data contributed less to total loss than the default data – for all other magnitudes, the percentage was the same for real and default data. In Seattle, the building type information was much more reliable and complete and hence the occupancy matrices were modified far more than that for the City of Long Beach which had very limited and unreliable building structure information. For both cities, the content damage as a percentage of total loss decreased with the increase in earthquake magnitude, which is not unusual since the building damage is more with larger magnitude earthquakes.

Other losses analyzed included shelter needs, casualties, amount of debris generated, and fires. For the City of Seattle, the trend was uniformly the same as building losses – i.e. the percentage difference increased with the increase in the magnitude of the earthquake and the local data always lead to more losses in all instances (except with the value exposed to fires for a 5.0 event) . However, the same was not the case for the City of Long Beach. The patterns of damage in Long Beach were much more erratic and in some cases the losses were more with local data and in other cases, less with local data. The number of casualties was always higher with local data in both cities for all three magnitudes. The discrepancies in the Long Beach findings may be partly attributable to the lack of structure type information for the City of Long Beach. The results for the City of Seattle show large discrepancies in numbers and have serious implications for planning and preparedness if HAZUS™ is used with default data as will be discussed in the next chapter (Chapter 7).

The two cities also exhibited differences in trends regarding the spatial distribution of losses. For both cities, patterns were only analyzed for the 7.0 magnitude event. For the City of Seattle, even though some census tracts showed overestimation and other census tracts showed underestimation in building inventory square footage, the damage losses for all census tracts were more with local data for all three magnitudes of earthquake. In real terms, the variation was less for residential census tracts and more for census tracts with mixed land use and very large for the downtown census tracts. In percentage terms, the change was not highest in the downtown census tracts but was scattered in census tracts throughout the city. The difference in losses for the City of Long Beach followed a similar pattern to the difference in building inventory for the City. Thus, for the most part, census tracts with higher square footage with local data showed higher losses with local data and vice versa. However, as with the City of Seattle, the difference in losses were less for residential census tracts and more with other single land use (commercial, industrial, university, etc) and mixed land use census tracts.

In summary, the improvement in building inventory from local sources such as tax assessment data can have a significant impact on the loss estimates from HAZUS™. Even marginal changes in data can lead to much larger changes in loss estimates as was seen in the case of City of Long Beach. In the case of City of Seattle, the larger difference in building inventory yield much larger difference in loss estimates and the trends are much more consistent – i.e. loss estimate difference with local data vs. default data increased with the increase in earthquake magnitude. The City of Seattle also showed significant differences in induced losses such as shelter needs, debris removal, casualties, and fires, most of which increased with local data and with increase in the magnitude of the earthquake. The trends were not so clear for the City of Long Beach. The lack of structure type information could be a factor leading to these inconsistencies in Long Beach. Therefore, improvement in building structure type is an important component of the sensitivity of the HAZUS™ to improvement in local data. Spatially, the City of Seattle showed higher losses for all census tracts whether the building square footage is overestimated or underestimated by the default data. In the City of Long Beach, the differences in loss estimates followed a pattern similar to the change in square footage. However, in both cities, losses for residential census tracts were usually better estimated (overestimated by a smaller extent) than other single use and mixed census tracts (which are usually underestimated by a larger degree).

In the next chapter, the policy implications of the above findings are discussed followed by a discussion of the areas for future research.

Chapter 7: Conclusions and Policy Implications

7.0 Introduction

So far this research has established that although data are available at the local level, there is a dearth of data that can assist in damage assessment. Particularly, local data on building inventory are lacking for information on the type of building, and exposure value (both building and content exposure). Many of these building characteristics are often not available from local sources and even when available, these data can be unreliable and incomplete. For example, in many cases, information on tax-exempt properties (e.g. public properties, religious properties, and educational properties) is not available. Other GIS datasets available at the local level (i.e. building footprint) can assist minimally in supplementing information available from the tax assessment data. But unless these datasets are properly developed and the right attributes collected at the time of digitization, they can offer limited usage and can even compound the uncertainty to the model.

However, where good data are available (as in the case of City of Seattle), this research shows that the default data in the HAZUS™ model are very different from the ground reality in the case of large cities. This difference can be seen at all levels: at the overall level of the city, when the various occupancy levels are broken down, and also spatially across the city (at the level of the various census tracts that constitute the city). In general, residential occupancy is one that is overestimated by HAZUS™ and even though it is overestimated by a smaller percentage, this amounts to a large value in real terms. This overestimation of residential square footage in default data undermines the level of underestimation of square footage in other occupancy classes. The data in HAZUS™ are very poorly estimated for commercial, residential, industrial, government and religious occupancy classes. There is also significant variation across the city at the census tract level. In general, census tracts that are primarily residential are better estimated than census tracts that are primarily commercial, industrial, mixed uses or other single use such as university, airport, hospital, etc. The downtown shows a large concentration of underestimation. Even a small difference in building inventory data can result in a much larger difference in damage estimates. Thus, a 33% underestimation in total square footage in Seattle results in 130% difference in damage loss estimates for a 7.0 magnitude earthquake. There is also significant variation in other direct and indirect losses such as shelter needs, casualties, amount of debris and fires.

In the case of City of Long Beach where good data are not available and even though the overall difference in building inventory square footage citywide is not high, there is a significant variation in the breakup of the total square footage into various occupancy classes. As in the case of City of Seattle, the City of Long Beach also shows that the residential square footage is overestimated by default data by a small percentage but this leads to a large square footage in real terms and offsets some of the large underestimation in all other occupancy classes. Primarily, there is significant underestimation in other occupancy classes such as commercial, industrial, education, and government. When good data are not available (as in the case of Long Beach), various assumptions about the data need to be made. The results are mixed and it is not clear how much uncertainty can be removed by using local level data. However, even slight improvement of data, can lead to larger differences in damage estimates. Thus a 2% increase in the inventory for the entire City of Long Beach resulted in 21% decrease in losses for a 5.0 magnitude earthquake and an 11% increase in loss for a 7.0 magnitude earthquake. The difference is also substantial for induced losses (shelter needs, debris, fires and casualties).

The findings of this research have various policy implications which are discussed in this section. First the discussion focuses on the appropriate use of models with and without local level data. Second, it looks at issues related to improving local data. Third, the chapter discusses improvements in the HAZUS™ model based on the findings of this research. Finally, there is a discussion on future areas of research.

7.1 Appropriate Use of Damage Estimation Models at Local-Level

Based on the findings of this research, the most fundamental question that arises is regarding the appropriate use of damage assessment models. Given the uncertainty that comes with default data, what is the utility of the HAZUS™ model at the local level? As discussed before, damage assessment models can be very useful for the disaster managers and planners for disaster preparedness, response, recovery and mitigation. Decision makers have very little basis to make decisions without such models. However, decisions based on these models at a local-level can be risky if default building inventory data are used since the default data are not reflective of the ground reality in the case of large cities, as shown by this research. This research also shows that for models to be useful for local decision-making, local-level data should be used for building inventory, particularly where good data are available at the local level. This is a very important step in reducing the uncertainty from the model.

Where good local level data are not available, a significant amount of time and effort is needed to improve the data, without really decreasing the uncertainty and changing the results significantly. In such cases, is the use of HAZUS™ warranted at all? The dissertation concludes that default data should only be used in large cities where all other sources of data have been exhausted. Even small improvements in the square footage can be beneficial. However, it is equally important to understand and assess local data thoroughly and be aware of uncertainties that get introduced due to local data irregularities. Local data are particularly needed for non-residential census tracts (downtown census tracts, mixed use tracts, and other special use census tracts such as census tracts with university, large mall, airport, hospital, etc). Such data can be collected through field surveys or other commercial sources and will be discussed more in details in the next section. If no other data exist at the local level, default data in HAZUS™ may be used, but only with a lot of caution and with a clear understanding of the fact that the results may show a great deal of underestimation, particularly for non-residential occupancies and for non-residential census tracts. These changes in inventory can have a much larger impact on the loss estimates for total losses as well as for other induced losses such as casualties, debris, shelter needs, and fires.

It is also appropriate at this time to revisit the discussion on inductive versus deductive models (Alexander 2000) as discussed in Chapter 2. When an inductive model such as HAZUS™ makes “spuriously precise” prediction of damage, based on poor default data, it can mask the risks and uncertainties involved with making decisions based on these models. Therefore, inductive models with poor data can be very dangerous for the decision-makers and particularly when the risks involved are not somehow quantified or illustrated with some case studies. The use of these models with poor data can be suitable for deductive modeling where the results are used to understand and establish relationships and causalities. Hence the use of models such as HAZUS™ with default data may also be somewhat suitable for preparedness and exercise purposes. Even for this, caution should be taken to rely solely on HAZUS™ default data without any understanding of the sensitivity of the model and the degree of variation of results. However, these models should definitely not be used with default data for damage estimation after a real event as the results can lead to some serious discrepancies and decisions that can have dangerous implications.

The uncertainty inherent in modeling a complex event (such as an earthquake or hurricane) may warrant that more attention be paid to stochastic models rather than using deterministic models. Stochastic models allow more suitable modeling of the randomness by allowing random variations of one or more parameters of the model and analyzing the probability distribution of outcomes from many simulations. Stochastic modeling has its origin in physics and is now being used in life sciences, social science and actuarial sciences. While stochastic modeling provides some obvious advantages to applications involving uncertainty, the efficacy of stochastic models to incorporate randomness is a large topic and beyond the scope of this dissertation.

It may be argued that the HAZUS™ model is meant more for regional damage assessment, so that overestimations in some parts are offset by underestimations in other parts of the region. If the purpose of HAZUS™ is solely regional assessment, again there is virtually no reason to have the “spuriously precise” (Alexander 2000) results that are provided by HAZUS™. Furthermore, it raises the question - what constitutes “regional”? Is “region” defined solely by the number of jurisdictions and geographical extent, or also by the population, built environment, exposure, and economic diversity? A large city (such as the ones studied here) can have a population, economic mix, and exposure very similar or even more complex than a multi-jurisdictional region comprising many counties. Even for a regional assessment, it is clear that where large cities are part of a larger region, the use of default data is so different from reality that its use can result in very poor decisions. For example, if the State of Washington uses HAZUS™ to allocate funding for retrofitting buildings in the greater Seattle region which include the City of Seattle, it is obvious that the City’s share would be underestimated and some of the suburban areas may be overestimated. Likewise, the planning for the number of casualties, debris, shelter needs, etc. may be based on erroneous data. All of the above data are critical for use in disaster management and particularly to support local level decision-making and the needs of the first responders. Inappropriate use of the tools can lead to poor decisions (which in the case of disasters can mean the difference between life and death) or simply the confidence in these tools can be so eroded that they will no longer be part of a disaster manager’s toolbox.

The next section will discuss in details the policy implications related to improving local data.

7.2 Improving Local Data

This research shows that improving the default data for building inventory in HAZUS™ with local data sources has a large impact on the change in damage estimates from the HAZUS™ model. Hence the sensitivity of the HAZUS™ model to local building data is well established by now. Therefore, as more and more advanced scientific models are developed, it is also crucial to improve the data that are input into the models. Any model development and improvement should occur with an eye on the strengths and weaknesses of locally available data since default data are seriously lacking in large urban areas.

As mentioned before, a critical dataset for damage assessment is parcel data and the corresponding tax assessment data. Although these data are available for most large cities now, the utility of these data for damage assessment is largely dependent on the quality of the data. Historically, tax assessment data have not been used for purposes other than assessment and often these data are not complete or collected consistently and stored in a proper database structure, restricting their use for other applications (Spencer 2003; Wiggins 1999; Eichenbaum et al 1993). Furthermore every local government has its own ways of collecting these data and there is very little consistency in the use of codes and data structure. Thus, when disasters occur and data from various agencies are needed, it is difficult to combine these data and use them appropriately. In the information rich world that we live in, it is important to recognize and acknowledge that the wealth of information that is available in the assessment datasets can and should be used for purposes other than tax and assessment. Only when this fact is well recognized, will there be an effort to improve these data.

There are various ways in which these data can be improved. First, there is no existing standard data model established for these data. Setting up a data model for assessment data can be the first step in standardizing assessment data by specifying fields and domains of values for those fields. Therefore, the field that captures building type information can be the same name and can have a domain of values that are used consistently. An advisory group comprising experts in tax and assessment and representatives from other departments such as planning, disaster management and other fields that consume this information should be formed to establish some minimum data requirements that should be collected along with some quality control measures. This group should also establish some basic attributes that should be collected for tax-exempt properties including educational, religious and governmental institutions. The data model can be established for the nation and can be altered or modified to reflect local situations. The data model can also be established at the state level, regional level or even at the County level, if the higher level standards are not conforming to the local needs. The federal and state government can provide incentives (in form of grants and cost-sharing) to migrate local data and legacy systems to established models, particularly for large urban regions. Although this is a lofty goal, data models exist for many GIS datasets such as parcels, road centerlines, etc. and increasingly being used at the local level. Furthermore, recent disasters have made both the assessment community and the disaster management community realize the value of tax assessment data for disaster management purposes.

Apart from the assessment data, efforts should be made by large cities to improve information on buildings. Many cities such as the City of are already making a concerted effort to improve their data for large buildings, including collecting building plans and information on entrance to buildings, fire escape routes, etc (Roberts 2007; City of Chicago 2002). As discussed above, other datasets can also be used to improve local data. The permit data (which are not available for many cities in digital format, particularly for older buildings) should be further explored to see how they can be leveraged to improve the assessment data for better building characteristics. Likewise, commercially available datasets can also be used to supplement local data. Commercial data sources include fire insurance maps from Sanborn (commonly called Sanborn Maps) which are now available as 3D models called Citysets® for many large cities, real-estate datasets such as CoStar, and also other sources such as SIMmetry City Models from EarthData. These commercially available datasets contain a wealth of information on high-value buildings including information on square footage, use, number of stories, occupants, and photos of the buildings. However, there are limitations to these sources too – primarily the fact that they cover a small extent of a limited number of cities. Furthermore, many do not contain much more improved attribute information, particularly information on the type of structure. Such commercial datasets should be tapped, particularly for census tracts that HAZUS™ has poor default data such as census tracts in the downtown, mixed use census tracts and those containing universities, airports, etc. Local governments should explore mutually beneficial relationships with these companies to get affordable access to these data for emergency management.

Building footprint data should also be collected and maintained on a regular basis for large cities and if possible some of the above information should be linked to this dataset. The collection of these data should be at the local level and given their needs for disaster management, the development of these data can be promoted by federal and state agencies through incentives such as grants and cost-sharing. However, the needs have to be local and this will ensure that the data are maintained, updated and verified for currency and accuracy. It is also evident from this research that most cities have fairly advanced GIS programs and hence are well prepared to gather such data. Where cities do not have a good GIS program, assistance from the State and Federal sources can provide the impetus in establishing such programs so that they are sustainable in the long term.

Another important aspect of using data is the issue of sharing data across jurisdictional boundaries and the establishment of policies and standards that will promote such data sharing. This research points to the lack of regional and inter-local agreements for local data use for disaster management. Such agreements have to be put in place before the occurrence of a disaster and should include agreements for data sharing with private utility companies in the aftermath of a disaster. Again, federal and state agencies can provide financial and other incentives for data sharing, standardization for easy regional integration and regional assessment with improved data.

Finally, in order to effectively use HAZUS™ in the aftermath of a disaster (i.e. in the disaster response phase) it is important that local data be prepared well before the onset of a hazardous event. It is obvious from this research that input of data for one jurisdiction is difficult enough. To do so for a region comprising many jurisdictions requires much more preparation as different local governments use different data classifications, schema, and each of the datasets yield different levels of uncertainty. Some regional efforts may be needed to get the data ready for use in the aftermath of a disaster and FEMA and other regional agencies can play a large role. In fact, the input of local data should be integrated into preparedness plans. If HAZUS™ is to be used effectively, getting data prepared for input into HAZUS™ should be part of the preparedness plan for FEMA, state governments, and local governments.

Thus, any federal development of tools and models, or acquisition of data must take into consideration the programs that exist at the local level and the presence of some key datasets at most local levels government of large cities. In the absence of this, tools and models will not serve the needs of the local emergency managers and will lead to early rejection rather than diffusion and adoption. While federal dollars are used to improve models, some funding should also be available to improve local data through incentives such as grants and cost sharing or as seed money to fund data acquisition.

The next section discusses the improvements needed in HAZUS™ based on the findings of this research.

7.3 Improvements in HAZUS™

HAZUS™ can be an important tool in the toolkit of a disaster manager. It provides disaster managers with the capabilities of running scenarios and analyzing the impacts of various types of events. The integration of many disparate models, scalability and flexibility of the tool are some of its greatest strengths. It is being used not only in the preparedness phase but also in the disaster response and recovery phase for many real events. However, this research points to the need for many improvements in HAZUS™, most of them focused on issues of data and issues of usability. The need for improvement here focuses only on the building inventory data and the usability of

HAZUS™ (i.e. issues explored through this dissertation). Improvements may well be warranted in other areas, including the strength of assumptions, improvements in other data and models used (such as the modeling of the physical hazard), etc. which were beyond the scope of this research and will not be discussed here.

It is apparent that one of the biggest drawbacks of HAZUS™ is the data for building inventory. The building inventory data is not a good reflection of the reality, particularly for large cities. While there may be limited data sources nationwide to get the type of building information needed to undertake damage assessments, it is important to check the quality of the Dun and Bradstreet data and provide a good discussion to the users about the strengths and weaknesses of these data. It is also important to provide the users with a thorough description of the nature of the source data, how it is input into HAZUS™ and what limitations might be attributable to the processing of the data. The metadata also does not provide information on how these data are gathered or input into HAZUS™ or what quality control measures are undertaken for the Dun and Bradstreet data. There is also no discussion of what types of facilities are likely to be under-represented in the data and what might be over-represented. A proper understanding of these shortcomings will help decision-makers understand the uncertainty better, use the tool more appropriately and will allow them to better assess the need for integration of local data, acquisition of new data or incorporation of third party data sources.

From this research, it is apparent that the Dun and Bradstreet data are not very good for smaller businesses and establishments. Retail and most other commercial specific occupancies are grossly underestimated and this reveals that only large retail establishments are available through the Dun and Bradstreet data. It is also not suitable for industrial, educational and governmental occupancies, all of which are grossly underestimated. Some of these occupancies can be improved by using other datasets already available in HAZUS™. For example, there is a separate inventory for schools and universities. This inventory shows 124 schools in Long Beach whereas the HAZUS™ default data shows only 10 educational buildings and the local data shows 122 educational buildings. This inventory of schools also contains information on number of students, square footage, etc., although not all of these fields are populated. These data should be leveraged to improve the building inventory.

At the very least, there should be tools provided to the users to leverage these datasets if they were interested in improving the data or supplementing local data. Furthermore, alternatives to the Dun and Bradstreet data should also be explored. This may include calculation of square footage based on land use and zoning maps available locally. Particularly if the data are as bad as this research finds, the methodology should also be simplified. Thus, the scalable product's complexity should be scaled based on the inputs into the model so that the users are made aware of the uncertainty in the results. For example, when default data are used, the results can also be provided in ranges rather than as precise numbers (Alexander 2000).

The issue of lack of content exposure at the local level also needs to be given some thought in next version of HAZUS™. It is clear that in new versions of HAZUS™, for local data to be input, building and content exposure are required from the local data. If these are not available, the software keeps the default exposure values and does not improve it based on the improvement in the square footage. Therefore, all the effort involved in inputting better square footage and building characteristic information is a complete waste since the dollar exposure is not updated and hence loss estimates are not significantly impacted. At the local level, assessment data carries information on assessed values, which are only indicative of market values of building exposure, but not equal to market values. However, if the assessment data do not carry any information on tax-exempt properties, this means that the exposure value of buildings with education, government and religious occupancies will be underestimated. The content exposure data is virtually impossible to get at the local level and therefore, this fact needs to be recognized and accepted. This will ensure that HAZUS™ devises a methodology to improve the content and building exposure data using default per square foot exposure values (that can be changed by the user to account for local conditions) to reflect improvements in locally-available data (particularly square footage). If this is not done, it is virtually useless to improve the data in HAZUS™ from local sources in cases where both building and content exposure data are not available locally.

As versions of GIS software changes, new versions of HAZUS™ need to be released. While a lot of effort is put into these new releases (which does not really improve the underlying science or the data behind these models), there should be some effort to improve the data as well. Software releases of HAZUS™ should also be managed better. Particularly better quality control is needed for not only the software interface but also the default data. It is apparent that the default data in HAZUS™ are continuously changing without much testing. Default data in a new version can be significantly different from the old version and it is not discussed why these discrepancies occur. The HAZUS™ team should have some sample datasets for various types of regions that have good local level data (urban, suburban, rural, etc.) for testing purposes when new data are incorporated into the model. Furthermore, based on these tests, users should be provided with some results on how far the data vary from local data based on these tests. This will also help in informing users of the degree of uncertainty.

Finally, as far as usability is concerned, HAZUS™ can do a lot to improve its graphical user interface. Without going into the litany of things that need significant improvements, this research found many features of the software that need improvement. Primarily, it should allow users to run-what if scenarios a lot easier and save them within one region. As it stands now, each scenario has to be built as a separate region. Each region, the size of Long Beach and Seattle are about half a gigabyte in size and hence quite resource intensive. Furthermore, no two regions can be open at the same time. There are no tools available to analyze difference between scenarios other than reports or exporting data into software such as Excel to run some statistics. Hence, it is very cumbersome to analyze the differences between two scenarios, or assess the impact of various mitigation policies. Finally, the building inventory tool (BIT) needs significant improvements in usability and if this is not accomplished, this can turn out to be a significant obstacle in the use of local data in HAZUS™.

7.4 Limitations and Future Research

Based on the findings of this research, there are many areas that need further exploration and are discussed in this section.

First, this research only looked at variations of default data from local data and discrepancies in local data for large cities. Similar analysis can be undertaken for other types of areas as well. Thus, data can be analyzed for smaller cities, suburban areas, and rural areas (possibly at the county scale) and in various parts of the country. Furthermore, this research only looked at data needs for earthquake analysis and the losses from earthquakes only. Similar research can be undertaken to take into account data needed for other disasters such as hurricanes and floods (both modules are now available in HAZUS™) for a variety of region types.

Second, with only two case studies, there are limited possibilities for drawing conclusions about trends and generalizations. The research design for this research comprised of two case cities - the City of Seattle represented a city that had good local data whereas the City of Long Beach represented a city where local data, although seemingly available and good, proved challenging in many aspects to input into HAZUS™. The purpose was to understand both types of circumstances. However, the choice of cities did not allow a good understanding of trends and generalizations, particularly as they relate to building type information. Therefore, further case studies are needed for more cities with good data (like Seattle) to better understand the trends of variations of the default data from local data and generalize the results so that they can be used to truly assess the range of difference of default data from local data. By understanding the variation in more cities, it may be better possible to devise parameters that can be altered to improve default data rather than to input local data.

Third, for the purpose of this dissertation, the building and content exposure were updated simply to reflect updated local square footage information and using HAZUS™ default average values. Although building exposure could be acquired from local sources, the content exposure is difficult to get from local sources. The sensitivity of the model to better building exposure can be further analyzed, particularly if good local data can be available for a few cities. Furthermore, the sensitivity can be analyzed better by changing variables in the building inventory one by one – i.e. first the square footage alone, then the square footage and building type, then square footage, building type and exposure information, etc. to understand the impact of each on the results from the data.

Fourth, this research shows that default data in HAZUS™ do not provide a good picture of the ground reality, at least for large cities and non-residential occupancies. Therefore for all types of damage assessment models, it is important to explore if better square footage information can be inferred from other local data sources such as land use/land cover data, zoning maps and ordinances, etc.

Fifth, more research is needed to establish a suitable data model for building data and assessment data. These data are crucial for disaster damage assessment and for all disaster management purposes. Hence, in order to leverage these data for purposes other than tax and assessment (e.g. planning, disaster management, natural resource management, etc), a data model with good data standards and minimum data needs is needed.

Finally, there is always a need to validate the results from a model (with or without local data) with results from real events. This may be better accomplished with floods and hurricanes rather than earthquakes because of the frequency of occurrences of these hazards. However, whatever the disaster, the inventory remains the same and can inform vastly on the sensitivity and uncertainty involved.

Appendix A: Classification in HAZUS™

Table 1: Occupancy Classes in HAZUS™

Label	Occupancy Class	Example Descriptions
	Residential	
RES1	Single Family Dwelling	House
RES2	Mobile Home	Mobile Home
RES3	Multi Family Dwelling RES3A Duplex RES3B 3-4 Units RES3C 5-9 Units RES3D 10-19 Units RES3E 20-49 Units RES3F 50+ Units	Apartment/Condominium
RES4	Temporary Lodging	Hotel/Motel
RES5	Institutional Dormitory	Group Housing (military, college), Jails
RES6	Nursing Home	
	Commercial	
COM1	Retail Trade	Store
COM2	Wholesale Trade	Warehouse
COM3	Personal and Repair Services	Service Station/Shop
COM4	Professional/Technical Services	Offices
COM5	Banks	
COM6	Hospital	
COM7	Medical Office/Clinic	
COM8	Entertainment & Recreation	Restaurants/Bars
COM9	Theaters	Theaters
COM10	Parking	Garages
	Industrial	
IND1	Heavy	Factory
IND2	Light	Factory
IND3	Food/Drugs/Chemicals	Factory
IND4	Metals/Minerals Processing	Factory
IND5	High Technology	Factory
IND6	Construction	Office
	Agriculture	
AGR1	Agriculture	
	Religion/Non/Profit	
REL1	Church/Non-Profit	
	Government	
GOV1	General Services	Office
GOV2	Emergency Response	Police/Fire Station/EOC
	Education	
EDU1	Grade Schools	
EDU2	Colleges/Universities	Does not include group housing

Table 2: Building Types in HAZUS™

No.	Label	Description	Height			
			Range		Typical	
			Name	Stories	Stories	Feet
1 2	W1 W2	**Wood, Light Frame (≤ 5,000 sq. ft.)** **Wood, Commercial and Industrial (> 5,000 sq. ft.)**		1 - 2 All	1 2	14 24
3 4 5	S1L S1M S1H	**Steel Moment Frame**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 13	24 60 156
6 7 8	S2L S2M S2H	**Steel Braced Frame**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 13	24 60 156
9	S3	**Steel Light Frame**		All	1	15
10 11 12	S4L S4M S4H	**Steel Frame with Cast-in-Place Concrete Shear Walls**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 13	24 60 156
13 14 15	S5L S5M S5H	**Steel Frame with Unreinforced Masonry Infill Walls**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 13	24 60 156
16 17 18	C1L C1M C1H	**Concrete Moment Frame**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 12	20 50 120
19 20 21	C2L C2M C2H	**Concrete Shear Walls**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 12	20 50 120
22 23 24	C3L C3M C3H	**Concrete Frame with Unreinforced Masonry Infill Walls**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 12	20 50 120
25	PC1	**Precast Concrete Tilt-Up Walls**		All	1	15
26 27 28	PC2L PC2M PC2H	**Precast Concrete Frames with Concrete Shear Walls**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 12	20 50 120
29 30	RM1L RM2M	**Reinforced Masonry Bearing Walls with Wood or Metal Deck Diaphragms**	Low-Rise Mid-Rise	1-3 4+	2 5	20 50
31 32 33	RM2L RM2M RM2H	**Reinforced Masonry Bearing Walls with Precast Concrete Diaphragms**	Low-Rise Mid-Rise High-Rise	1 - 3 4 - 7 8+	2 5 12	20 50 120
34 35	URML URM M	**Unreinforced Masonry Bearing Walls**	Low-Rise Mid-Rise	1 - 2 3+	1 3	15 35
36	MH	**Mobile Homes**		All	1	10

Table 3: Source of Default Data in HAZUS™

Label	Occupancy Class	Source of Data		
		Census		Dun and Bradstreet
		Unit of Data	Conversion Factor	SIC Code
	Residential			
RES1	Single Family Dwelling	# of Units	variable	
RES2	Mobile Home	# of Units	1000 sq. t./unit	
RES3	Multi Family Dwelling	# of Units	1000 sq. t./unit	
RES4	Temporary Lodging			70
RES5	Institutional Dormitory	# in Group Quarters	700 sq. ft./person	
RES6	Nursing Home			8051, 8052, 8059
	Commercial			
COM1	Retail Trade			52, 53, 54, 55, 56, 57, 59
COM2	Wholesale Trade			42, 50, 51
COM3	Personal/Repair Services			72,75,76,83,88
COM4	Prof./Technical Services			40, 41, 44, 45, 46, 47, 49, 61, 62, 63, 64, 65, 67, 73, 78 (except 7832), 81, 87, 89
COM5	Banks			60
COM6	Hospital			8062, 8063, 8069
COM7	Medical Office/Clinic			80 (except 8051, 8052, 8059, 8062, 8063, 8069)
COM8	Entertainment & Rec.			48, 58, 79, (except 7911), 84
COM9	Theaters			7832, 7911
COM10	Parking			
	Industrial			
IND1	Heavy			22, 24, 26, 32, 34, 35 (except 3571, 3572), 37
IND2	Light			23, 25, 27, 30, 31, 36 (except 3671, 3672, 3674), 38, 39
IND3	Food/Drugs/Chemicals			20, 21, 28, 29
IND4	Metals/Minerals Processing			10, 12, 13, 14, 33
IND5	High Technology			3571, 3572, 3671, 3672, 3674
IND6	Construction			15, 16, 17
	Agriculture			
AGR1	Agriculture			01, 02, 07, 08, 09
	Religion/Non/Profit			
REL1	Church/ N.P. Offices			86
	Government			
GOV1	General Services			43, 91, 92 (except 9221, 9224), 93, 94, 95, 96, 97
GOV2	Emergency Response			9221, 9224
	Education			
EDU1	Schools			82 (except 8221, 8222)
EDU2	Colleges/Universities			8221, 8222

Appendix B1: Survey Questionnaire

Doctoral Dissertation Title: Disaster Damage Estimation – Data Needs vs. Ground Reality
Sudha Maheshwari, Dept of Urban Planning and Policy Development, Rutgers University, NJ

Digital Geographic Data Survey

I. Participant Information

1) **Survey Contact**

2) **Survey Contact's Title/Position**

3) Name of Organization
Enter the name of the organization along with the specific department that you represent. Eg City of Los Angeles, Department of Information Technology, GIS Division.

4) Organization's Address

5) Phone

6) Fax

7) Email

8) Which of the following categories of GIS users do you consider yourself?

    1. Using GIS for all its functionalities - creating and manipulating spatial databases, spatial analysis, and display
    2. Using GIS for analysis and display but not in creating or maintaining spatial databases
    3. Using GIS for viewing data only
    4. Very rarely doing anything with GIS

9) Have you heard of HAZUS, the GIS-based tool for disaster damage assessment developed by Federal Emergency Management Agency and National Institute for Building Sciences?

    a. Yes b. No.

10) If yes, which of the following categories of Hazus user do you consider yourself?

1. Use it extensively in my job
2. Have assisted other people in using in by providing GIS help
3. Have attended training or learnt it by reading the manual but never used it
4. Have never used HAZUS

II. Organizational Details

11) Which of the following categories best identifies your organization?

    1. Municipal Agency
    2. Sub County Region
    3. Single County
    4. Multiple County Organization or Consortium
    5. State
    6. National
    7. Other - Please specify__________________________

12) If your organization encompasses multiple jurisdictions, counties, or states, please specify these in detail.

13) Is there any group/consortium that coordinates geographic data in your jurisdiction?

    1. Yes 2. No

14) **Is your organization a part of any such geographic data coordinating group?**

    1. Yes 2. No

15) List the top three geographic data coordinating groups your organization works with most:
    1. ________________________________
    2. ________________________________
    3. ________________________________

16) Is there a Geographic Information Systems department in your organization? If so, please include the department under which it is and a contact name and information.

17) Is there an emergency/disaster management department in your organization? If so, please include the department under which it is and a contact name and information.

III. Digital Geographic Information Data

Overview

18) Which of the following geographic datasets is available in any standard GIS format for your city? Also fill in the name of the department in your organization that is responsible for the dataset and a contact person if you know.

In the "Available" field ,underneath please write Y if data is available currently, N if data is not available currently, C if data is currently under creation, and D if you do not know about the availability of the dataset?

Geographic Dataset	**Available**	**Department**	**Contact**	**Phone/Email**
1. Digital Parcel Maps				
2. Building Footprints				
3. Land/Tax Records				
4. Street Centerlines				
5. Transportation				
6. Water Network				
7. Wastewater Network				
8. Zoning, Landuse				
9. Historical resources				
10. Hazardous waste sites				
11. Educational Institutions				
12. Emergency Facilities (Fire, Emergency Operation Centers, Police Stations, etc.)				
13. Aerial/Ortho Photos				
14. Geodetic Control/Engg data				

Detailed information on above datasets

Please answer the following questions providing detailed information on some of the above datasets. If you are not in a position to answer the questions, please check on the box and continue to the next dataset. However, if you know a contact person or another department that can provide this information, please fill in this before moving to the next dataset.

19) **Parcel Maps**

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

   Name ____
   Department ____
   Phone No ____
   Email ____

3. What is your organization's role with respect to digital parcel map data for your jurisdiction? **Select as many as applicable.**

   a. Using
   b. Creating
   c. Updating and maintaining
   d. Distributing
   e. None

4. Is there FGDC complaint metadata available for this dataset?

5. What is the status of the creation of parcel maps for your jurisdiction?

   a. Completed
   b. Work in progress
   c. Planned
   d. Not planned

6. Which year was your parcel map created? If Work in Progress, please put in the anticipated month and year of completion.

7. When was this dataset last updated? Please specify the month and year.

8. What is the current geographical extent of this dataset for your jurisdiction? **Please select only one**. For the definition of your jurisdiction please refer to question 12 above.

    a. Entire Jurisdiction
    b. > **two-thirds** of the jurisdiction
    c. **one-third** to **two-thirds** of the jurisdiction
    d. < **one-third** of the jurisdiction

9. What attribute data is attached or could be attached with the parcel features? **Please select as many as necessary**.

    a. Parcel Id
    b. Assessor's Parcel Number (APN)
    c. All tax assessor's data linked with APN/Parcel Id
    d. Partial tax assessor's data linked with APN/Parcel Id
    e. Photos of individual properties
    f. Custom created building characteristic information
    g. Custom created land-use information
    h. Custom created zoning information
    i. Other – Please specify ___________________________________

10. What format is used in your organization for this dataset? **Please select as many as necessary**.

    a. ArcView shapefiles
    b. ArcInfo coverages
    c. MapInfo
    d. GeoMedia
    e. AutoCAD
    f. MicroStation
    g. Smallworld GIS
    h. Other – Please specify ___________________________________

11. Which of the following best represents the positional accuracy of this dataset? **Please select only one.**

    a. 95% of the points within 3 ft
    b. 95% of the points between 3ft and 6 ft
    c. 95 % of the points between 6ft and 15ft
    d. 95% of the points outside 15ft
    e. Not sure

12. If GIS parcel data is not available currently, what mode are the parcel maps in? **Please select as many as necessary.**

    a. Paper Maps
    b. Scanned raster maps
    c. Geo-referenced scanned raster maps
    d. Legal description
    e. Other
    f. Don't know

13. Has the parcel data layer ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

    a. Yes
    b. No
    c. Don't know

    If yes, please specify ____________________________________

14. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

15. Do you normally charge for this data? If so, how much?

20) **Building Footprints**

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

   Name ____________________________________
   Department ____________________________________
   Phone No ____________________________________
   Email ____________________________________

3. What is your organization's role with respect to digital building footprint data for your jurisdiction? **Please select as many as necessary**.

   a. Using
   b. Creating
   c. Updating and maintaining
   d. Distributing
   e. None

4. Is there FGDC complaint metadata available for this dataset?

5. What is the status of the creation of building footprints for your jurisdiction?

   a. Completed
   b. Work in progress

c. Planned
d. Not planned

6. Which year was your building footprint data created? If Work in Progress, please put in the anticipated month and year of completion.

7. When was this dataset last updated? Please specify the month and year.

8. What is the current geographical extent of this dataset for your jurisdiction? Please check only one. For the definition of your jurisdiction please refer to question 12 above.

    a. Entire Jurisdiction
    b. > **two-thirds** of the jurisdiction
    c. **one-third** to **two-thirds** of the jurisdiction
    d. < **one-third** of the jurisdiction

9. What attribute data is attached with the building footprints? If the data is not attached but is available elsewhere and can be linked through some ID, consider it as attached data. **Please select as many as necessary**.

    a. Parcel Id
    b. Assessor's parcel number (APN)
    c. All tax assessor's data linked with APN/Parcel Id
    d. Partial tax assessor's data linked with APN/Parcel Id
    e. Photos of individual properties
    f. Custom created building characteristic information
    g. Custom created land-use information
    h. Custom created zoning information
    i. Other – Please specify ___________________________________

10. What format is used in your organization for this dataset? **Please select as many as necessary.**

    a. ArcView shapefiles
    b. ArcInfo coverages
    c. MapInfo
    d. GeoMedia
    e. AutoCAD
    i. MicroStation
    j. Smallworld GIS
    k. Other – Please specify ___________________________________

11. Which of the following best represents the positional accuracy of this dataset? **Please select only one.**

    a. 95% of the points within 3 ft

b. 95% of the points between 3ft and 6 ft
c. 95 % of the points between 6ft and 15ft
d. 95% of the points outside 15ft
e. Not sure

12. Has the building footprint data layer ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

    a. Yes
    b. No
    c. Don't know

    If yes, please specify ___________________________________

13. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

14. Do you normally charge for this data? If so, how much?

21) Land and Tax Records

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

    Name ____
    Department ____
    Phone No ____
    Email ____

3. What is your organization's role with respect to land and tax records data for your jurisdiction? **Select as many as applicable.**

    a. Using
    b. Creating
    c. Updating and maintaining
    d. Distributing
    e. None

4. Is there FGDC complaint metadata available for this dataset?

5. When was this dataset last updated? Please specify the month and year.

6. Which year was the last tax assessment done?

7. Has the tax assessor's data been linked to the GIS parcel maps? **Please select only one.**

    a. Completed
    b. Work in progress
    c. Planned
    d. Not planned

8. Does the assessor's data for your city include complete records for properties that are not taxed such as religious, educational, or government properties, etc.?

9. What attributes about buildings does your tax assessor's data carry in digital format? **Please select as many as necessary**.

    a. Building use
    b. Building height
    c. Building square footage
    d. Year of construction/age of building
    e. Building material – e.g. brick, concrete, wood, etc.
    f. Type of building structure – e.g. load bearing wall, steel, wood, or RCC frame
    g. Assessed value of property
    h. Number of occupants in building

10. In your opinion, how complete is the assessor's data for your city? **Please select only one.**

    a. Covers > 90% of the parcels in the jurisdiction
    b. 75% - 89% of the parcels
    c. 50% - 75% of the parcels
    d. Less than 50% of the parcels
    e. Don't know

11. What format is used in your organization for this dataset? **Please select as many as necessary**

    a. Relational Database (such as Access, Oracle, SAP, etc.)
    b. Mainframe (such as COBOL, PASCAL, or other format)
    c. Flat File (Excel or other)
    f. Application
    g. Other – Please specify __________________________________
    f. Don't know

12. Has the tax assessment data ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

    a. Yes
    b. No
    c. Don't know

    If yes, please specify ___________________________________

13. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

14. Do you normally charge for this data? If so, how much?

22) **Street Centerlines**

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

    Name ____
    Department ____
    Phone No ____
    Email ____

3. What is your organization's role with respect to street centerline data for your jurisdiction? **Please select as many as necessary.**

    a. Using
    b. Creating
    c. Updating and maintaining
    d. Distributing
    e. None

4. Is there FGDC complaint metadata available for this dataset?

5. What is the status of the creation of street centerlines for your jurisdiction? **Please select only one.**

    a. Completed

b. Work in progress
c. Planned
d. Not planned

6. Which year was your street centerline data created? If Work in Progress, please put in the anticipated month and year of completion.

7. When was this dataset last updated? Please specify the month and year.

8. What is the current geographical extent of this dataset for your jurisdiction? **Please select only one.** For the definition of your jurisdiction please refer to question 12 above.

    a. Entire Jurisdiction
    b. > **two-thirds** of the jurisdiction
    c. **one-third** to **two-thirds** of the jurisdiction
    d. < **one-third** of the jurisdiction

9. Does your dataset on streets include street addresses?

    a. Yes b. No

10. What format is used in your organization for this dataset?

    a. ArcView shapefiles
    b. ArcInfo coverages
    c. MapInfo
    d. GeoMedia
    e. AutoCAD
    h. MicroStation
    i. Smallworld GIS
    j. Other – Please specify ____________________

11. Which of the following best represents the positional accuracy of this dataset? Please select only one.

    a. 95% of the points within 3 ft
    b. 95% of the points between 3ft and 6 ft
    c. 95 % of the points between 6ft and 15ft
    d. 95% of the points outside 15ft
    e. Not sure

12. Has the street centerline data layer ever been used in your organization for disaster management purposes (i.e response, preparedness, recovery of mitigation)?

a. Yes
b. No
c. Don't know

If yes, please specify ____________________________________

13. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

14. Do you normally charge for this data? If so, how much?

23) **Transportation**

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

Name ____________________________________
Department ____________________________________
Phone No ____________________________________
Email ____________________________________

3. What is the status of the following transportation layers for your jurisdiction? **Please select only one for each layer.**

Transportation Layer	Completed	Work in progress	Planned	Not Planned
a. Roads				
b. Railroads				
c. Light Rail				
d. Waterways				
e. Airports				
f. Ports				
g. Bridges				
h. Tunnels				

4. Is there FGDC complaint metadata available for these dataset?

5. What is the source of the available transportation layers for your jurisdiction?

Transportation Layer	Updated TIGER/other agency data (specify)	Digitized by your agency	Other methods (specify)	Don't know
a. Road				
b. Railroads				
c. Light Rail				
d. Waterways				
e. Airports				
f. Ports				
g. Bridges				
h. Tunnels				

6. What is the current geographical extent of transportation datasets for your jurisdiction? **Please select only one**. For the definition of your jurisdiction please refer to question 12 above.

    a. Entire Jurisdiction
    b. > **two-thirds** of the jurisdiction
    c. **one-third** to **two-thirds** of the jurisdiction
    d. < **one-third** of the jurisdiction

7. Have any of the transportation data layer ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

    a. Yes
    b. No
    c. Don't know

    If yes, please specify ___________________________________

8. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

9. Do you normally charge for this data? If so, how much?

24) Public Utilities (Water and WasteWater)

1. I do not know much about these datasets.

2. The following person/department may be able to provide you with more information:

a. <u>Water</u>

Name ______________________________
Department ______________________________
Phone No ______________________________
Email ______________________________

b. <u>Wastewater</u>

Name ______________________________
Department ______________________________
Phone No ______________________________
Email ______________________________

3. What is the status of the creation of digital public utilities data for your jurisdiction?

Utility Layer	Completed	Work in progress	Planned	Not Planned
a. Water				
b. Wastewater				
c. Communication				

4. Is there FGDC complaint metadata available for this dataset?

5. What is the current geographical extent of this dataset for your jurisdiction? Please check only one. For the definition of your jurisdiction please refer to question 12 above.

   a. Entire Jurisdiction
   b. > **two-thirds** of the jurisdiction
   c. **one-third** to **two-thirds** of the jurisdiction
   d. < **one-third** of the jurisdiction

6. Have any of the utility data layer ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

   a. Yes,
   b. No
   c. Don’t know

   If yes, please specify ______________________________

7. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

8. Do you normally charge for this data? If so, how much?

25) **<u>Digital Orthophotography</u>**

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

   Name ____
   Department ____
   Phone No ____
   Email ____

3. What is your organization's role with respect to street centerline data for your jurisdiction? **Select as many as applicable.**

   a. Using
   b. Creating
   c. Updating and maintaining
   d. Distributing
   e. None

4. Is there FGDC complaint metadata available for this dataset?

5. What is the status of the creation of orthophotos for your jurisdiction?

   a. Completed
   b. Work in progress
   c. Planned
   d. Not planned

6. Which year was your latest orthophoto acquired? If new photography is being acquired recently, please also put in the anticipated month and year of completion.

7. What is the current geographical extent of this dataset for your jurisdiction? Please check only one. For the definition of your jurisdiction please refer to question 12 above.

   a. Entire Jurisdiction
   b. > **two-thirds** of the jurisdiction
   c. **one-third** to **two-thirds** of the jurisdiction

d. < **one-third** of the jurisdiction

8. Which best represents the spatial resolution of your orthophotos?

   a. Finer than 1 m resolution
   b. 1 meter resolution
   c. 1 – 3 m resolution
   d. 3 m resolution
   e. Don't know

9. Which best represents the accuracy of your orthophotos?

   a. 95% of the points within 3 ft
   b. 95% of the points between 3ft and 6 ft
   c. 95 % of the points between 6ft and 15ft
   d. 95% of the points outside 15ft
   e. Not sure

10. Digital orthophotography ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

    a. Yes
    b. No
    c. Don't know

    If yes, please specify ______________________________

11. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

12. Do you normally charge for this data? If so, how much?

26) Elevation Data

1. I do not know much about this dataset.

2. The following person/department may be able to provide you with more information:

   Name ______________________________
   Department ______________________________
   Phone No ______________________________
   Email ______________________________

3. What is your organization's role with respect to street centerline data for your jurisdiction? **Select as many as applicable.**

    a. Using
    b. Creating
    c. Updating and maintaining
    d. Distributing
    e. None

4. Is there FGDC complaint metadata available for this dataset?

5. What is the status of the creation of elevation data for your jurisdiction?

    a. Completed
    b. Work in progress
    c. Planned
    d. Not planned

6. Which year was your latest elevation data acquired? If Work in Progress, please indicate the anticipated month and year of completion.

7. What is the current geographical extent of this dataset for your jurisdiction? Please check only one. For the definition of your jurisdiction please refer to question 12 above.

    a. Entire Jurisdiction
    b. > **two-thirds** of the jurisdiction
    c. **one-third** to **two-thirds** of the jurisdiction
    d. < **one-third** of the jurisdiction

8. Which best represents the horizontal accuracy of your elevation model?

    a. 95% of the points within 3 ft
    b. 95% of the points between 3ft and 6 ft
    c. 95 % of the points between 6ft and 15ft
    d. 95% of the points outside 15ft
    e. Not sure

9. Which best represents the vertical accuracy of your elevation model?

    a. < 4 ft
    b. 4 – 10 ft
    c. > 10 ft
    d. Don't know

10. Has digital elevation data ever been used in your organization for disaster management purposes (i.e. response, preparedness, recovery or mitigation)?

    a. Yes
    b. No
    c. Don’t know

    If yes, please specify ___________________________________

11. In exchange for results of analyses of an earthquake event using HAZUS 99, would you be willing to share this data?

12. Do you normally charge for this data? If so, how much?

Appendix B2: GIS Diffusion in Selected Cities

Table 1: Summary of GIS Diffusion in Surveyed Cities

City Name	*Type of GIS*	*GIS Mgr/ Coord.*	*Dept housing Enterprise GIS*	*Regional Consortium*	*Intranet Mapping*	*Public Access via Internet*	*Emergency Mgmt*
Atlanta, GA	Enterprise	No	Public Works	Yes	Yes	No	No
Birmingham, AL	Enterprise	Yes	Planning/ Permitting	No	No	No	City/ County EMA
Colorado Springs, CO	Departmental	No		No	No	No	Yes
Honolulu, HI	Enterprise	Yes	Planning/ Permitting	Yes	Yes	Yes	No
Jacksonville, FL	Enterprise	Yes	Info Tech	No	Yes	Yes	Yes
Las Vegas, NV	Enterprise	Yes	Dept of Tech	No	Yes	Yes	No
Long Beach, CA	Enterprise	No	Info Tech	Yes	Yes	No	?
Miami, FL	Departmental	No		Yes	No	No	No
Milwaukee, WI	Enterprise	Yes	Info Tech	Yes	Yes	Yes	Yes
Minneapolis, MN	Enterprise	Yes	Info Tech	Yes	Yes	No	Yes
Newark, NJ	Departmental	No		No	No	No	Yes
Oklahoma City, OK	Enterprise	Yes	Info Tech	No	Yes	No	Yes
Omaha, NE	Departmental	No		Yes	No	No	No
Pittsburgh, PA	Departmental	No		No	No	No	Yes
Portland, OR	Enterprise	Yes	Info Tech	Yes	Yes	Yes	Yes
San Antonio, TX	Enterprise	Yes	Info Tech	No	Yes	No	Yes
Santa Ana, CA	Departmental	Yes		No	Yes	No	Yes
Seattle, WA	Enterprise	Yes	Public Utilities	Yes	Yes	Yes	Yes
Wichita, KS	Enterprise	Yes	Info Tech	No	Yes	No	No

Table 2: Availability of Core Dataset by City

	Parcel with Tax Attribute	**Bldg Footprint**	**Street Centerline**	**Ortho**	**Topography**
Atlanta, GA	Yes	Yes	Yes	Yes	Yes
Birmingham, AL	Yes	Yes	Yes	Partial	Partial
Colorado Springs, CO	Yes	Yes	Yes	Yes	Yes
Honolulu, HI	Yes	Partial	Yes	Partial	Yes
Jacksonville, FL	Yes	Planned	Yes	Yes	Yes
Las Vegas, NV	Yes	Yes	Yes	Yes	Yes
Long Beach, CA	Yes	Yes	Yes	Yes	Partial
Miami, FL	Yes	Partial	Yes	Yes	Planned
Milwaukee, WI	Yes	Yes	Yes	Yes	Yes
Minneapolis, MN	Yes	Yes	In Progress	Yes	Yes
Newark, NJ	Yes	Yes	In Progress	Yes	Partial
Oklahoma City, OK	Yes	Yes	Yes	Yes	Yes
Omaha, NE	Yes	No	Yes	Yes	Partial
Pittsburgh, PA	Yes	Yes	Yes	Yes	Yes
Portland, OR	Yes	Yes	Yes	Yes	Yes
San Antonio, TX	Yes	Partial	Yes	Yes	Yes
Santa Ana, CA	Yes	No	Yes	Yes	Yes
Seattle, WA	Yes	Yes	Yes	Yes	Yes
Wichita, KS	Yes	No	Yes	Yes	Partial

Appendix C: City of Seattle Case Study

Table 1: Mapping of Local Structure Type to HAZUS™ Structure Types

Local Structure	Source of Assessor	Count	HAZUS Structure Type
Wood (<50%)	Residential Table	124,080	Wood
Masonry (>= 50%)	Residential Table	15,150	Masonry
Structural Steel	Commercial Table	210	Steel
Reinforced Concrete	Commercial Table	765	Concrete
Masonry	Commercial Table	5,673	Masonry
Wood Frame	Commercial Table	10,491	Wood
Prefab Steel	Commercial Table	717	Precast
Unknown	Commercial Table	2,269	Unknown
Total		159,355	

Appendix D: City of Long Beach Case Study

Table 1: Construction/Structure Type Information in Local Data

Type of Construction	# of Records
Brick	65
Concrete	295
Concrete Block	35
Frame	13,234
Heavy	1
Light	23
Log	4
Manufactured	6
Masonry	118
Metal	25
Special	4
Steel/Heavy	2
No Data	74,205

Table 2: Values in Frame_Code Field

FRAME_CODE	# of Records
Concrete	99
Masonry	310
Steel	90
Wood	3,274
No Data	84,244

Appendix E: Comparison of Seattle and Long Beach

Table 1: Average square footage of various occupancies for Seattle and Long Beach

		Seattle		Long Beach	
Specific Occupancy	**Description**	**Average Default Size**	**Average Local Size**	**Average Default Size**	**Average Local Size**
RES1	Single Family Dwelling	1,600	1,632	1,600	1,634
RES2	Manuf. Housing	1,161	4,567	1,065	11,318
RES3A	Duplex	3,253	1,978	3,065	1,864
RES3B	Triplex / Quads	3,537	3,271	3,098	2,837
RES4	Temporary Lodging	176,889	46,218	184,657	16,822
RES5	Institutional Dormitory	27,177	21,271	26,401	1,229
RES6	Nursing Home	27,104	36,413	25,053	14,197
COM1	Retail	338,951	10,133	722,556	7,573
COM2	Wholesale Trade	35,076	20,998	36,981	22,220
COM3	Personal and Repair Services	11,436	5,213	11,703	4,450
COM4	Professional/Technical Services	105,525	42,365	113,411	13,918
COM5	Banks	4,188	7,761	4,323	10,620
COM6	Hospital	61,679	131,157	55,016	29,533
COM7	Medical Office/Clinic	7,446	18,241	7,940	1,505
COM8	Entertainment & Recreation	5,029	11,925	5,117	4,998
COM9	Theaters	16,913	39,616	31,600	23,733
IND1	Heavy	34,414	17,962	33,684	16,676
IND2	Light	42,815	18,173	34,568	8,390
IND3	Food/Drugs/Chemicals	62,846	11,048	93,600	11,625
IND4	Metals/Minerals Processing	40,180	13,668	50,750	13,987
IND5	High Technology	#DIV/0	21,362	#DIV/0	61
IND6	Construction	82,870	10,380	155,025	8,782
AGR1	Agriculture	120,460	7,442	#DIV/0	3,454
REL1	Religion	24,650	11,042	26,079	11,793
GOV1	General Services	11,484	27,651	11,216	13,293
GOV2	Emergency Response	35,200	#DIV/0	9,267	#DIV/0
EDU1	Grade Schools	539,450	64,978	365,450	39,370
EDU2	Colleges/Universities	58,227	#DIV/0	54,625	182,194

List of References and Citations

Alexander, David E. 1995. A Survey of the Field of Natural Hazards and Disaster Studies. In *Geographical Information Systems in Assessing Natural Hazards*. ed. Carrara A. and Guzzeti F. Kluwer, 1-20. Netherlands: Academic Publishers.

Applied Technology Council (ATC). 1985. *ATC-13*: *Earthquake Damage Evaluation Data for California*. Applied Technology Council.

Bernknopf, Richard L. Laura B. Dinitz, Sharyl J.M. Rabinovici, and Alexander M. Evans. 2001. A Portfolio Approach to Evaluating Natural Hazard Mitigation Policies: An Application to Lateral-Spread Ground Failure in Coastal California. *International Geology Review* 43: 424-44

Burby, Raymond J., Robert E Deyle, David R.Godschalk and Robert B. Olshansky. 2000. Creating Hazard Resilient Communities through Land-Use Planning. *Natural Hazards Review* 1 (2): 99-112.

California State University. 2000. *Annual Financial Report*. California State University, Long Beach.

Carrara, Alberto, and Fausto Guzzetti. 1995. *Geographical Information Systems in Assessing Natural Hazards*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Chan, T.O. and I.P. Williamson. 1999. The Different Identities of GIS and GIS Diffusion, *International Journal of Geographic Information Science* 13 (3): 267-281.

City of Chicago. 2002. *Rules and Regulations Pursuant to Chapter 13-78 of the Municipal Code of Chicago "High-Rise Buildings – Emergency Procedures" and Associated Code Provisions.* On-line. Available from Internet, http://cityofchicago.org/Fire/ProposedRules/Rules9.pdf, accessed 14 April 2003.

Consortium for International Earth Science Information Network (CIESIN). 1995. *Thematic Guide to Integrated Assessment Modeling of Climate Change*. On-line. Available from Internet, http://sedac.ciesin.org/mva/iamcc.tg/TGHP.html, accessed 22 September 2007.

Coppock, Terry J. 1995. GIS and Natural Hazards: An Overview From a GIS Perspective. In *Geographical Information Systems in Assessing Natural Hazards*, eds. A. Carrara and F. Guzzetti. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Degg, Martin. 1992. Natural Disasters: Recent Trends and Future Prospects. *Geography* 77(336): 198-209

Deyle, Robert E., Steven P. French, Robert B. Olshansky and Robert G. Paterson. 1998. Hazard Assessment: The Factual Basis for Planning and Mitigation. In *Cooperating with Nature: Confronting Natural Hazards with Land-Use Planning for Sustainable Communities*, ed. Raymond J. Burby, 119-166. Washington DC: Joseph Henry Press.

Dymon, Ute. 1990. *The Role of Emergency Mapping in Disaster Response.* Quick Response Report # 42. University of Colorado at Boulder: Natural Hazards Center.

Dymon, Ute. 1993. *Map Use During and After Hurricane Andrew*. Quick Response Report # 60. University of Colorado at Boulder: Natural Hazards Center.

Dymon, Ute. 1999. *Effectiveness of Geographic Information Systems (GIS) Applications in Flood Management During and After Hurricane Fran.* Quick Response Report # 114. University of Colorado at Boulder: Natural Hazards Center.

Eichenbaum Jack, Fran Pearl, and John F. Ryan. 1993. Assessment Information: Raw Resource in the Municipal Jungle. *Property Tax Journal.* 1(3): 89–124.

Eichenbaum, Jack. 2002. CAMA, GIS, and the Recovery of New York City. *Assessment Journal*, 9(3): 17-22.

Elber, Gail. 2002. A National Map or a Federal Map? *GeoSpatial Solutions*. October 1. On-line. Available from Internet, http://www.geospatial-online.com/geospatialsolutions/article/articleDetail.jsp?id=31971, accessed 22 September 2007.

FEMA. 2005. On-line. Available from Internet, http://www.fema.gov, accessed on 15 October 2005.

FEMA. 2003. *Earthquake Loss Estimation Methodology HAZUS™*. User's Manual.

Gartner. 2002. *Government GIS Survey: Fall 2002*. On-line. Available from Internet, http://www4.gartner.com/Init, accessed on 10 April.

Greene, R.W. 2002. *Confronting Catastrophe: A GIS Handbook*. Redlands California: ESRI Press.

Heiken, Grant, Gregory A. Valentine, Meredith Brown, Denise George, and Steen Rasmussen. 2000. *Modeling Cities: The Los Alamos Urban Security Initiative.* Paper presented at the American Geophysical Union (AGU) Spring Meeting Sessions, Washington D.C. May 30-June 3.

Maheshwari, S., Jones, E. and Rasmussen, S. 1999. *Urban Security Inititative: Modeling the 'Urban System of Systems'*. Paper presented at the Second Conference on Use of GIS and Remote Sensing for Disaster Management held in George Washington University, Washington D.C. Jan 19-21.

Masser, Ian. 1992. The Diffusion of GIS in British Local Government. In *Diffusion and Use of Geographic Information Technologies.* eds. Masser I. & Onsrud H. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Masser, Ian. and Harlan Onsrud. 1992. *Diffusion and Use of Geographic Information Technologies*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Masser, Ian., Heather Campbell, and Massimo Craglia. 1996. *GIS Diffusion: The Adoption and Use of Geographic Information Systems in Local Government in Europe.* London: Taylor & Francis

Mitchell, James K. 1995. Coping with Natural Hazards and Disasters in Megacities: Perspectives on the Twenty-first Century. *Geojournal*. 37 (3): 303-311

Morentz, James, W. 1986. Emergency Management is Information Management. *In Terminal Disasters: Computer Applications in Emergency Management*. ed. Sallie A. Marston, University of Colorado, Boulder: Institute of Behavioral Science.

Nordenson, G., Deodatis, G., Tantala, M., and Kumpf, A. 1999. *Earthquake Loss Estimation Study for the New York City Area*. 1ST Year Technical Report prepared for MCEER. On-line. Available from Internet, http://www.nycem.org/techdocs/lossEstYr1/default.asp, accessed 22 September, 2007.

Parson, Edward E., and Karen Fisher-Vanden. 1997. Integrated Assessment Models of Global Climate Change. *In Annual Review of Energy and Environment.* eds. Robert H. Socolow, Dennis Anderson, and John Harte, 589-628. Palo Alto, CA: Annual Reviews Inc.

Roberts, Mary Rose. 2007. Urban Legends. In Fire Chief. March 2007. On-line. Available from Internet, http://firechief.com/communications/firefighting_urban_legends/, accessed 22 September 2007

Rumor, Massimo. 1992. The Use of Geographic Information Technology in the City of Padova. In *Diffusion and Use of Geographic Information Technologies.* eds. Masser I. & Onsrud H., 229-244 Dordrecht, The Netherlands: Kluwer Academic Publishers.

Swiatek, Joseph A. and Dean C. Kaul. 1999. *Crisis Prediction Disaster Management*. SAIC Science and Technology Trends II. SAIC. On-line. Available from Internet, http://cats.saic.com/src/technical/catsswi.doc, accessed 22 September 2007.

Seattle Datasheet. 2005. On-line. Available from Internet, http://www.cityofseattle.net/oir/datasheet/economy.htm, accessed 22 September 2007.

Shanley, Lea, Steven J. Ventura, and David A. Hart. 2001. *Wisconsin and National Framework Surveys – Comparison, Validation, and Recommendations.* Project Report submitted to the Federal Geographic Data Committee and University Consortium for Geospatial Information Science. On-line. Available from Internet, http://www.lic.wisc.edu/framework/FrameSur.pdf, accessed 22 September 2007.

Sommers, Rebecca. 1999. Framework Data Survey Preliminary Report, Supplement to *Geo Info Systems*.

Spencer, Laurie. 2003. The Assessor's Database – Why it Must Change for GIS Success? *ImagiNews* – Quarterly Newsletter of Improving Michigan's Access to Geographic Information Networks. March/April.

Tulloch, David L. and Jennifer Fuld. 2001. Exploring County-level Production of Framework Data: Analysis of the National Framework Data Survey. *Journal of the Urban and Regional Information Systems Association*, 13 (2): 11-21.

Tulloch, David.L. 2000. *Exploratory Studies of the NSGIC/FGDC Framework Survey: Looking at the State of the Nation.* Final Report to the University Consortium for Geographic Information Sciences and The Federal Geographic Data Committee.

U.S. Geological Survey. 2001. *The National Map: Topographic Mapping for the 21st Century*. Reston, VA: Office of the Associate Director for Geography, U.S. Geologic Survey. On-line. Available from Internet, http://www.nationalmap.usgs.gov, accessed 22 September 2007.

United Nations Center for Human Settlements (HABITAT). 1996. *An Urbanizing World: Global Report on Human Settlements 1996*. Oxford University Press.

Waugh, William.L. 1995. Geographic Information Systems: The Case of Disaster Management. *Social Science Computer Rev*iew 13(4): 422-431.

Watson, Charles C., Mark E. Johnson. 1999. *Design, Implementation and Operation of a Modular Integrated Tropical Cyclone Hazard Model*. Proceedings of the 23rd Conference on Hurricanes and Tropical Meteorology, Dallas, TX. Jan 10-15.

Wiggins, Lyna. 1992. Diffusion and Use of Geographic Information Systems in Public Sector Agencies in the United States. In *Diffusion and Use of Geographic Information Technologies.* eds. Masser I. & Onsrud H. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Wiggins, Lyna. and Sudha Maheshwari. 1999. *Earthquake Loss Estimation Study for Newark, NJ*. Technical Progress Report. Prepared for the New Jersey State Police Office of Emergency Management.

Curriculum Vita

Sudha Maheshwari

EDUCATION

October 2007	Ph.D. Rutgers, The State University of New Jersey, New Brunswick NJ
September 1994 – August 1996	Master of Regional Planning University of Massachusetts, Amherst MA
July 1988 – June 1993	Bachelor of Architecture Jadavpur University, Kolkata India

PRINCIPLE OCCUPATIONS AND POSITIONS HELD

July 2006 - present	Operations Manager, Sanborn Ann Arbor MI
Summer 2007	Adjunct Faculty, Department of Architecture and Urban Planning University of Michigan, Ann Arbor MI
January 2005 – July 2006	Supervisor, Oakland County Department of Information Technologies, Pontiac MI
December 2001 – December 2004	GIS Project Manager, Oakland County Department of Information Technology, Pontiac MI
August 2001 – December 2001	Adjunct Faculty, Department of Architecture and Urban Planning University of Michigan, Ann Arbor MI
May 2000 – April 2001	GIS Project Manager, Urban Data Solutions, New York NY
September 1998 – August 1999	Graduate Student Intern, Los Alamos National Laboratory, Los Alamos NM
September 1997 – May 1998	Graduate Teaching Assistant, Rutgers University, New Brunswick NJ
September 1996 –	Graduate Research Assistant, Center for Urban Policy Research,

August 1997	Rutgers University, New Brunswick NJ
September 1994 – August 1996	Research Assistant and Teaching Assistant, University of Massachusetts, Amherst MA
January 1994 – July 1994	Architect, InterDesign, Kolkata India

SELECTED PUBLICATIONS

“Diversification of Defense Based Industries in India”. 2003. In A. Markusen and S. DiGiovanna eds. *From Defense to Development: International Perspectives on Realizing the Peace Dividend*. Routledge. New York.

“Integrated Modeling of Earthquake Impacts to the Electric-Power Infrastructure: Analyses of an Elysian Park Scenario in the Los Angeles Metropolitan Area.” (with L. J. Dowell). 1999. Los Alamos Technical Paper. LA-UR-99-4429. Also presented at the *Annual Conference of Urban and Regional Information Systems, Chicago, August 20-25*.