UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ANALYSIS OF SPATIAL UNCERTAINTY IN LIDAR-DERIVED BUILDING DATA AND UNCERTAINTY PROPAGATION IN MODELING OF URBAN ATMOSPHERIC DISPERSION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MANG LUNG CHEUK
Norman, Oklahoma
2009

ANALYSIS OF SPATIAL UNCERTAINTY IN LIDAR-DERIVED BUILDING DATA AND UNCERTAINTY PROPAGATION IN MODELING OF URBAN ATMOSPHERIC DISPERSION

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF GEOGRAPHY

BY

Dr. May Yuan, Chair

Dr. Tarek Rashed

Dr. Sally Gros

Dr. Petra Klein

Dr. Jeffrey Basara

Acknowledgements

Give thanks to God! My gratitude is devoted to my advisor, Dr. May Yuan, who supported my study in the University of Oklahoma and gave me motivation throughout the development of dissertation. Without her inspiration and guidance, this dissertation could not be successfully completed. Dr Tarek Rashed, Dr Sally Gros, Dr Petra Klein, and Dr Jeffrey Basara, thanks for being my committee members and each gave me specific comments on how to improve the dissertation. Also, I would like to present my gratitude to all my colleagues in Center for Spatial Analysis, especially Melissa Brown, who generously proofread my drafts many times. My appreciation also goes to my brothers and sisters in the Southern Oklahoma Chinese Baptist Church, who always pray for me when I need it. I would also like to thank my advisor in Hong Kong Baptist University, Dr Kenneth Wong, who granted me a flexible schedule in working. Finally, my wife, Anne Lee, thanks you for your love, patience and continued support in the past five years.

Table of Content

List of Tables

List of Figures

Abstract

Results of environmental models (EMs) are often used to assist decision making. However, EM outcomes vary significantly with different input data, model parameters and model assumptions. Therefore, informed decision making requires an in-depth understanding of how the changes in input data, model parameters and model assumptions influence the model outputs. While EMs are now accustomed to geo-spatial data, the influences of spatial uncertainty are often overlooked. This research examines the influence of spatial uncertainty throughout the three stages of general environment modeling: 1) examine the uncertainty in geo-spatial data as representation of the environment, 2) examine the uncertainty in the linkage between EMs and Geographic Information System (GIS) and, 3) examine and compare the influence of spatial uncertainty with the uncertainty of model parameters. LiDAR data and urban atmospheric dispersion model (UADM) are used as a use case, to demonstrate the methods and benefits of examining the influence of spatial uncertainty toward EMs.

Chapter 1 : Introduction

1. Introduction

Environmental modeling commonly contributes to decision making processes in many application domains. Many important policies and decisions are made based on the results of environmental modeling (King and Kraemer, 1993). City planners use hydrological models to determine flood plain areas (Maidment, 1993). Emergency managers employ atmospheric dispersion models to evacuate citizens during accidental toxic material release (NRC, 2003). Environmental agencies apply land use change models to study deforestation and suggest future development plans (Moran and Brondizio, 1998).

However, the outcomes of environmental models may not be sufficiently reliable for decision makers due to three kinds of uncertainty. First, uncertainty is inevitably associated with input geo-spatial data: How accurate are the geo-spatial data representing the environment? Second, uncertainty can be introduced from data processing: How does the data change during data conversion between GIS and EMs? Third, uncertainty attaches with the modeling parameters: How well our knowledge on phenomena is represented in a model? Without in-depth understanding of the uncertainty from these aspects, the outcomes of the environmental models are questionable.

In order to assure the outputs of environmental model are reliable, scientists validate their environmental models with observation data. On the other hand, to assure the input data are accurate, scientists also validate the input data through accuracy assessments. However, accuracy assessment of input data are becoming more and more difficult since the input data in the environmental models are becoming more complicated with the blossoming of geo-spatial technologies (i.e. Geographic Information Systems, Global Positioning Systems and Remote Sensing). Input data for environmental models are now available in different formats, multi-dimensional (3D or even 4D), and with fine resolutions (Burrough and McDonnell, 1998).

Geo-spatial data has become one of the standard input requirements for environmental modeling. Hydrological modeling requires digital elevation data, urban atmospheric dispersion modeling demands building dimension and location, and land use change modeling needs land use and population data. Even with a great variety of data sources, geo-spatial data still may not meet the specific input requirement of environmental models because of the different paths in development (Fedra, 1993). Environmental modelers focus more on the model performance while geographic information scientists focus more on spatial representation of reality. Use of different data models and formats challenge

data-model integration and therefore, geo-spatial data often demands conversion before entering into the environmental models.

Besides the input data, model parameters are another source of input uncertainty. Using atmospheric dispersion modeling as an example, uncertainty of model parameters includes variations in wind speed and direction, wind profile formula, atmospheric stability, surface roughness and other parameters. Each model may have a distinct set of parameters, and the uncertainty from parameters is commonly handled by conducting sensitivity analysis. However, most analyses focus on comparing uncertainties of different model parameters while uncertainty in input data (i.e. geo-spatial data) is overlooked.

A comprehensive understanding on the influence of uncertainty is beneficial to both decision makers and modelers. However, due to the complexity of environmental models and geo-spatial data, it is very difficult to analyze the influence of all uncertainties involved in environmental modeling. Nevertheless, this research proposes a three-stage approach to examine the influence of uncertainty in environmental modeling. To test the concept and methodology, this study use urban atmospheric dispersion modeling as an example. This method analyzes the uncertainty encountered in data-model integration in three common stages: 1) uncertainty arising from data gathering using LiDAR data

as example, 2) uncertainty arising from data conversion between Geographic Information System (GIS) and urban atmospheric dispersion models (UADMs), and 3) the uncertainty of the model outcomes arising from the uncertainty input.

2. Background

This study first defines the term “uncertainty” because for different disciplines the term “uncertainty” may have distinctive meanings and contain different components. In environmental modeling, uncertainty can be considered in various categories. The United States Environmental Protection Agency (USEPA, 1992) divided the uncertainty of environmental models into “scenario uncertainty”, “parameter uncertainty,” and “model uncertainty” referring to missing information on exposure and dose, model parameters, and gaps in scientific theory, respectively. However, Cullen and Frey (1999) characterized uncertainty by “input uncertainty” and “model uncertainty”, where “model uncertainty” arises from limited understanding of the model structure, and model detail as well as limited model validation, extrapolation, model boundaries, and model scenarios. “Input uncertainty” is due to empirical quantity errors such as measurement errors of pollutant concentrations.

In Geographic Information Science, uncertainty has been defined as a "discrepancy between geographic data in GIS and the geographic reality these data are intended to represent" (Zhu, 2005). Uncertainty in spatial data also refers to "spatial uncertainty" which includes measures of accuracy, statistical precision, bias in initial values, and estimated statistical coefficients in prediction, as well as estimation of error in the final output of modeling results (Mowrer, 2000).

Moreover, in GIS and Remote Sensing literature, spatial uncertainty relates to terms such as *accuracy, error, incompleteness, precision, randomness, bias, and data quality*. These terms are commonly used to describe the nature of spatial uncertainty and data quality. In the following paragraphs, these terms are defined and their relationships to spatial uncertainty provide a foundation to further examine the spatial uncertainty.

Accuracy refers to how close a measurement is to the reality it represents. In contrast, *error* refers to a discrepancy between a measurement and the reality it represents. Often, the term *error* is interchangeable with the term *uncertainty* in the literature. While the terms are very similar in meanings, they represent different approaches in measuring the associated discrepancy. *Error* is used when measuring discrepancy in absolute term while *uncertainty* is used when measuring discrepancy in relative term (Goodchild, Buttenfield, and Wood, 1994).

Error can be *random* or *systematic*. *Random error* implies that the discrepancy between a measurement and the reality it represents is caused by chance, and cannot be attributed to any underline processes. *Systematic error* implies a consistent discrepancy between measurement and the reality, which is also called *bias* (Mowrer, 2000).

The concept of *data quality* contains two meanings and seven dimensions. It can refer to the accuracy of data production or the fitness of use, which depends on application (Mead, 1982). The seven dimensions that contribute to *data quality* include lineage, position accuracy, attribute accuracy, completeness, logical consistency, semantic accuracy and temporal information (Morrison, 1995). *Uncertainty* is closely related to *data quality*. Data with greater *uncertainty* tend to be less useful and less accurate, and hence poorer in *data quality*. *Incompleteness* is an element of *data quality* which means lacking a part of data and is a cause for *error* or *uncertainty*. *Precision* refers to the degree of detail in measurement; higher precision means more detailed in measurement (Goodchild, 1993). However, higher precision does not necessarily mean lower *uncertainty*. For example, we can measure a building height in meters up to 12 decimal places but it can be far away from actual building height.

This study focuses on spatial uncertainty with respect to a "discrepancy between geographic data in GIS and the geographic reality these data are intended to represent"(Zhu, 2005), and the parameters uncertainty defined by USEPA. In particular, the study evaluates the spatial uncertainties in position (i.e. location of buildings) and attributes (i.e. height of buildings), and uncertainty in the meteorological variables of wind speed and wind direction.

A comprehensive understanding on how various kinds of spatial uncertainty contribute to the model results is necessary and beneficial to both decision makers and modelers. Particularly, in the field of urban atmospheric dispersion modeling (UADM), such understanding is desirable as UADM requires a large number of parameters and detailed spatial data. For example, emergency managers need to know the uncertainty of modeling results in case of hazardous material releases because uncertainty in one parameter or input data may significantly change the model results (NRC, 2003). Ultimately, the goal of the study is to improve the understanding of the modeling results with uncertainty and facilitate decision making under uncertainty.

A common approach is to test the influences of different model parameters through Monte Carlo simulation or sensitivity analysis. Many researchers focus on how uncertainties in model parameters affect UADM results. For example,

Sullivan et.al. (2004) examined uncertainty in wind speed, wind direction, atmospheric stability, and emission rate in two dispersion models (i.e. ISCST3 and TOXST model). The results indicate that the model output is sensitive to uncertainty in emission rate. Manomaiphiboon and Russell (2004) examined the uncertainty in friction velocity, mean surface turbulent heat flux magnitude, surface roughness height, and mean surface temperature. They identified uncertainty of friction velocity as the most influential factor among other meteorological variables. Levy et. al. (2002) determined that the results of the CALPUFF dispersion model were moderately insensitive to the parameterization of chemical mechanism, wet/dry deposition, background concentration, and size of the receptor region. Overall, the significance of uncertainty in model input parameters tends to vary according to the model and the applications. Few studies, however, have focused on the uncertainty of model parameters compared to uncertainty from spatial data, and the uncertainty effects on the results of UADM.

With the growth of utilizing geo-spatial data in environmental models, decision makers and modelers face not only the challenge of uncertainty from model parameters, but also the challenge of uncertainty embedded in the geo-spatial data. In the field of atmospheric dispersion study, building data is one

of the fundamental inputs. In the past, building data could only be obtained through land survey or acquiring building footprints from the building owners. Either process was time consuming and labor intensive. Advances in remote sensing technologies afford detailed and accurate building data, including height, with active sensors such as LiDAR (Light Detection and Ranging). LiDAR is a fast and relatively inexpensive method to capture detailed terrain elevation up to one-meter resolution. With these advantages, the LiDAR technique has been applied to a broad range of environmental applications such as forestry, urban planning and coastal morphological study (Hill, Graham, and Henry, 2000). Nonetheless, more detailed data does not guarantee less uncertainty. In fact, studies show that LiDAR accuracy may depart significantly from the estimate in complex terrain structure such as urban environments (Hopkinson et al., 2001; Ahokas, Kaartinen, and Hyyppa, 2003; Schenk, Csatho, and Lee, 1999). While many researchers have identified the possibility of greater LiDAR error than the overall estimated error in relation to vegetation (Huising and Pereira, 1998; Hopkinson et al., 2001), the error distribution in relation to urban environments is not yet fully understood.

Furthermore, the uncertainty in LiDAR data is blurred by the linkage between GIS and EMs, such as UADM. A linkage between GIS and

environmental models is necessary since they do not share common data models. Similar to most of the EMs, the format of geo-spatial data needs to be converted before ingesting to the UDAMs. Although data formats vary between different models, conversions are made mostly between two main data models: vector and raster models. A vector model represents objects as points, lines and polygons whereas, a raster model represents objects as groups of grid cells or pixels. Uncertainty associated with vector-raster conversion have been identified and studied in the field of GIS (Congalton, 1997; Wedhe, 1982; Piwowar, Ledrew, and Dudycha, 1990; Bregt et al., 1991). Most studies focused on examining the uncertainty in two-dimensional surface. However, in UADM, building data are in three-dimensional surface, as well as the uncertainty.

In response, this study primarily focuses on uncertainty arising from LiDAR-derived building data and examines its subsequent use in UADM, using the QUIC (Quick Urban and Industrial Complex) dispersion model as an example of UADM. This study also compares the uncertainty arising from LiDAR data with the uncertainty introduced by meteorological parameters.

3. Research Hypothesis

This study hypothesizes that the location uncertainty in geographical data is inevitable and it will propagate through the linkage between GIS and UADM. As a result, the uncertainty will alter the outcome of UADM.

> "Location uncertainty in geographical data will propagate and alter the outcome of UADM."

4. Statement of Research Problems

With the significant amount of uncertainty resulting from the spatial data, linkage between GIS and UADM, and model parameters in UADM, deterministic results of pollutant concentration without associated uncertainty information are not reliable to decision makers. Therefore, there is a need to understand how each source of uncertainty contributes to the results of UADM in urban environments. While the research takes the use-case of UADM, the approach developed is applicable to evaluate uncertainty in general GIS-Environmental modeling applications.

The goal of this research is to examine how spatial uncertainty in LiDAR-derived building data affects the result of the UADM in response to the urban environment. To achieve this goal, this study identified the following objectives:

1. Examine and quantify the spatial uncertainty of LiDAR-derived building data in relation to urban environments, using CBD in Oklahoma City as an example;

2. Examine and quantify the spatial uncertainty resulting from the linking GIS and UADM, using ArcGIS and QUIC dispersion model as an example;

3. Examine and quantify the effects of the spatial uncertainty in LiDAR-derived building data toward UADM, and compare the effects with meteorological parameters uncertainty, using QUIC dispersion model as an example.

5. Research Design

5.1 Conceptual flowchart

The concept of the research design is shown in Figure 1.1. First, the sources of spatial uncertainty in LiDAR-derived building data are identified. Then, with ground observations, the spatial uncertainty in statistical terms such as mean and standard deviation of discrepancies in building heights and locations are quantified. This study also examines the spatial distribution of uncertainty and summarizes the urban environments where spatial uncertainty is greater than average. After that, using ESRI ArcGIS and the QUIC dispersion model as examples, a linkage between GIS and UADM is created. As the linkage involves a transformation of building data, this study also quantifies the spatial uncertainty caused by the data transformation. Finally, the building data and model

meteorological parameters are perturbed based on the spatial uncertainty quantified above and ground observation in Joint Urban 2003 Atmospheric Dispersion Study (JU2003). Under the same dispersion scenario, this study generates two sets of simulations: one set with uncertainty from model meteorological parameters only and the other set with both spatial and meteorological sources of uncertainty. The final results are obtained by comparing and examining the uncertainty of the QUIC's results with the ground observations from JU2003 in Oklahoma City. The study area, rationale and summative synopses for each objective are described and explained as follows. Detailed research procedures and analytical outcomes are documented in subsequent chapters.

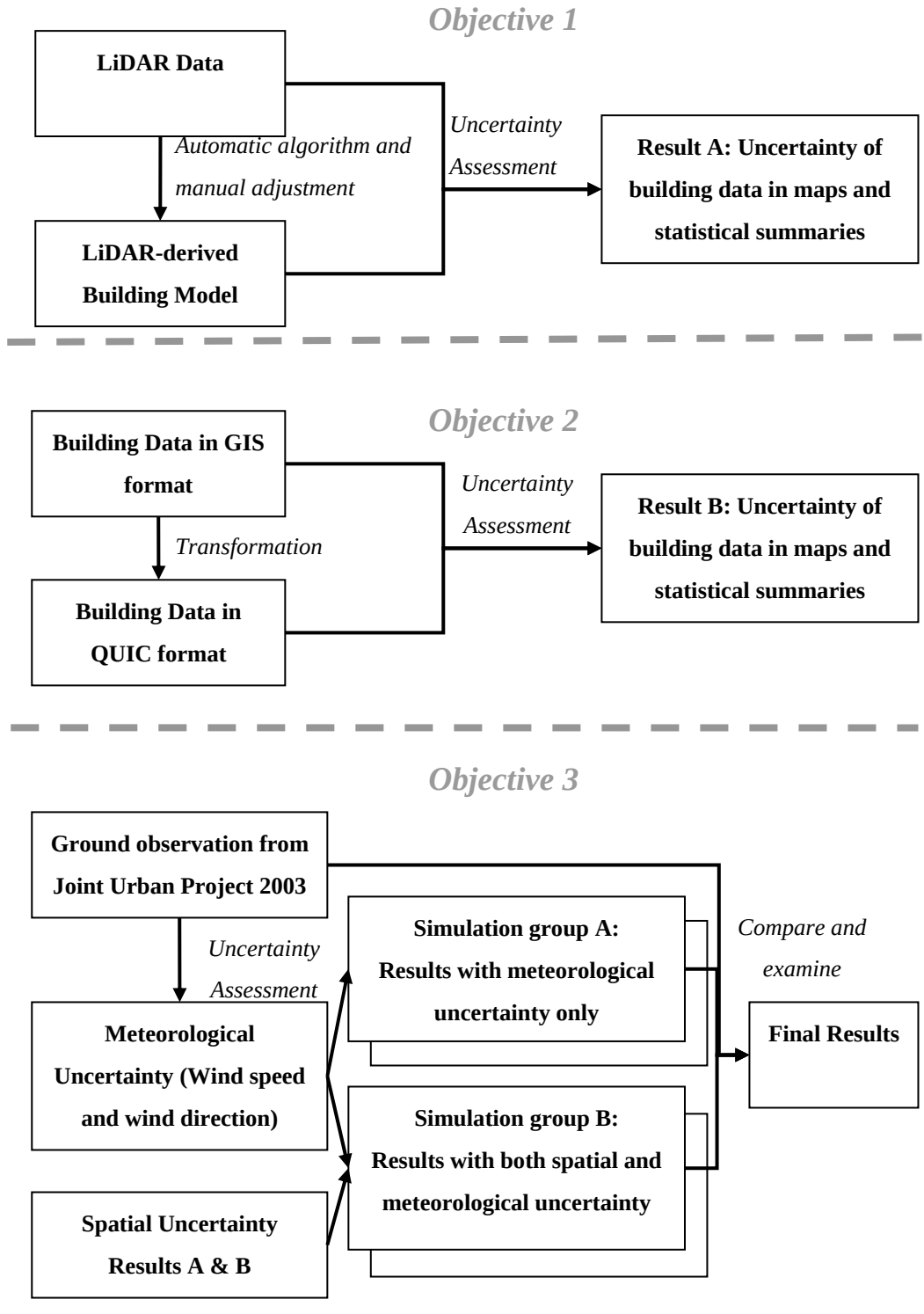


Figure 1.1. The conceptual flowchart of the research design.

5.2 Study area

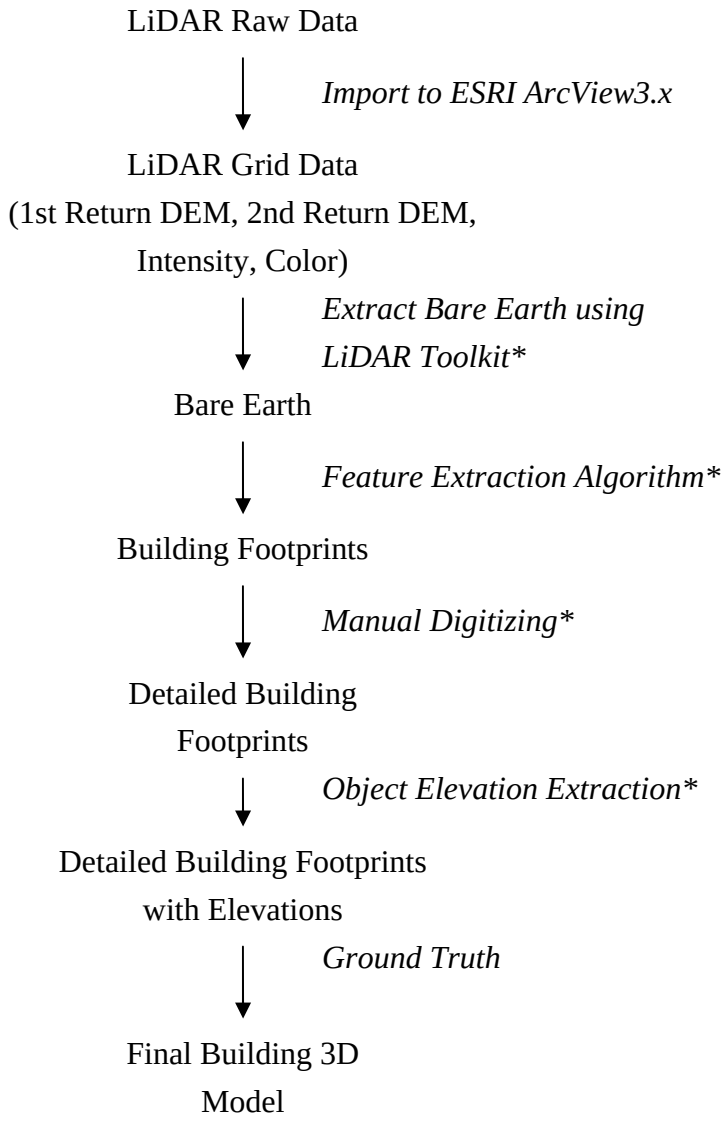
The study area for the first and second objective is approximately 0.5 square kilometers of flat terrain which is located within the Central Business District (CBD) of OKC (Figure 1.2). The city center is well defined and similar other CBDs dominated by high-rise buildings surrounded by open area and low-rise commercial buildings. Due to the nature of the dispersion, computation power and limited field data, a smaller extent (about 612 x 830 meters) of the study area is used in the last objective of the research.

Figure 1.2. The study area located in the Central Business District of Oklahoma City, Oklahoma. Study area for objective 3 is shown in white boundary. Data source: U.S. Geological Survey, capture date: 2002-03-26.

5.3 Objective 1: Identifying and quantifying the sources of spatial uncertainty in LiDAR-derived building data

This method begins by identifying the source of spatial uncertainty in the spatial data. With a list of procedures that describe how building data is derived from the LiDAR data, sources of spatial uncertainty are identified (Figure 1.3). For LiDAR-derived building data, each procedure contributes to spatial uncertainty. Three main sources of spatial uncertainty have been identified: extraction algorithm, manual digitizing, and LiDAR raw data. Results from feature extraction vary depending upon the chosen feature extraction algorithm. Manual extraction results in inconsistent digitized boundaries. Furthermore, accuracy of LiDAR raw data varies with response to different ground surfaces.

Each source of uncertainty influences different aspects of spatial data. Manual digitization creates uncertainty in building boundaries (i.e. location); extraction algorithms create uncertainty both in building boundaries and building height (i.e. attribute) because the final building heights are determined by algorithms that extract building footprints and bare-earth elevation. LiDAR data create uncertainty in building attributes because water and asphalt can absorb LiDAR signals and create incomplete coverage.



**Uncertainty may introduce during these processes.*

Figure 1.3. Procedures of generating building model from LiDAR data.

For this study, the uncertainties of these different sources are quantified by the following methods:

1) To quantify spatial uncertainty of building location arising from the extraction algorithm, this study calculates the differences between the x, y coordinates of automatically generated building footprints and the x, y coordinates of the final building model. For spatial uncertainty of building attribute, this study compares the building heights using different extraction algorithm options.

2) To quantify spatial uncertainty resulting from manual digitizing, this study calculates the differences between the x, y coordinates of a building boundary digitized by 30 people on the same computer screen at the same scale (i.e. 1:1000).

3) To quantify the spatial uncertainty of building elevation caused by LiDAR data, this study examines the building heights in the field where great elevation differences are found between final model and LiDAR data.

Finally, the mean and standard deviation of the differences are calculated. Results are shown as maps and distribution graphs that can be used to simulate uncertainty for the final objective.

5.4 Objective 2: Create a linkage between GIS and UADM dispersion model and examine the spatial uncertainty associated with the linkage

After quantifying the spatial uncertainty of LiDAR-derived building data, conversion procedures are necessary to link GIS data into an UADM. ESRI ArcGIS 9 and the QUIC dispersion model are used to illustrate the creation of a linkage between GIS and UADM. ArcGIS 9 is a popular commercial GIS

software developed by the Environmental Systems and Research Institute (ESRI, Redlands California). QUIC stands for Quick Urban and Industrial Complex dispersion modeling system and is a new light-weight dispersion model developed by Los Alamos National Laboratory (Pardyjak and Brown, 2002; Williams, Brown, and Pardyjak, 2002). QUIC runs in the MATLAB software package and uses a diagnostic wind field model (QUIC-URB) that has been coupled with a Lagrangian dispersion model (QUIC-PLUME). QUIC has been selected for its ability to produce rapid predictions of atmospheric dispersion in urban areas.

This study employs a tight-coupling strategy for the linkage between ArcGIS and QUIC due to time and skill constraints. Other possible strategies are loose-coupling and full integration. Compared to loose-coupling, tight coupling integrates data and user interfaces to reduce demand of user input during the linkage. On the other hand, tight coupling enables both GIS and environmental models to run independently and allow maximal flexibility for linkage in comparison to a full integration. The tight-coupling strategy transfers data between GIS and UADM through an alternate user-interface. A user-interface was designed using Visual Basic (VB) scripts within ArcMap because Visual Basic provides access to spatial data as well as spatial analysis tools. However, ArcMap and QUIC do not share the same building data model. Therefore, the

building data are modified before importing to QUIC, and hence additional spatial uncertainty is introduced. For instance, QUIC only accepts rectangles or circles as building shapes while ArcMap can store irregular building shapes. As a result, the building data are modified in two ways before importing to QUIC. First, the building data are converted to a gird according to user-defined resolution. Second, the building grids are converted back to polygons and divided into rectangles. The logic of the algorithm is shown in Figure 1.4. For transferring data from QUIC to ArcGIS, this function is already available inside QUIC dispersion model. Since the output data model are the same between QUIC and ArcGIS (i.e. both use raster model), no uncertainty assessment is needed.

During the data transformation from ArcGIS to QUIC, spatial uncertainty is again introduced. Building location and height vary according to the user-defined resolution. To quantify spatial uncertainty arising from change of resolution, this study first produce building data at one to twelve meter resolution in QUIC's format, then calculate the differences between the x, y coordinates of building boundaries at the different resolutions, using one meter resolution as the default. Same as objective 1, the mean and standard deviation of the differences are calculated. Results are shown as maps and distribution graphs, which can be used to simulate uncertainty for the final objective.

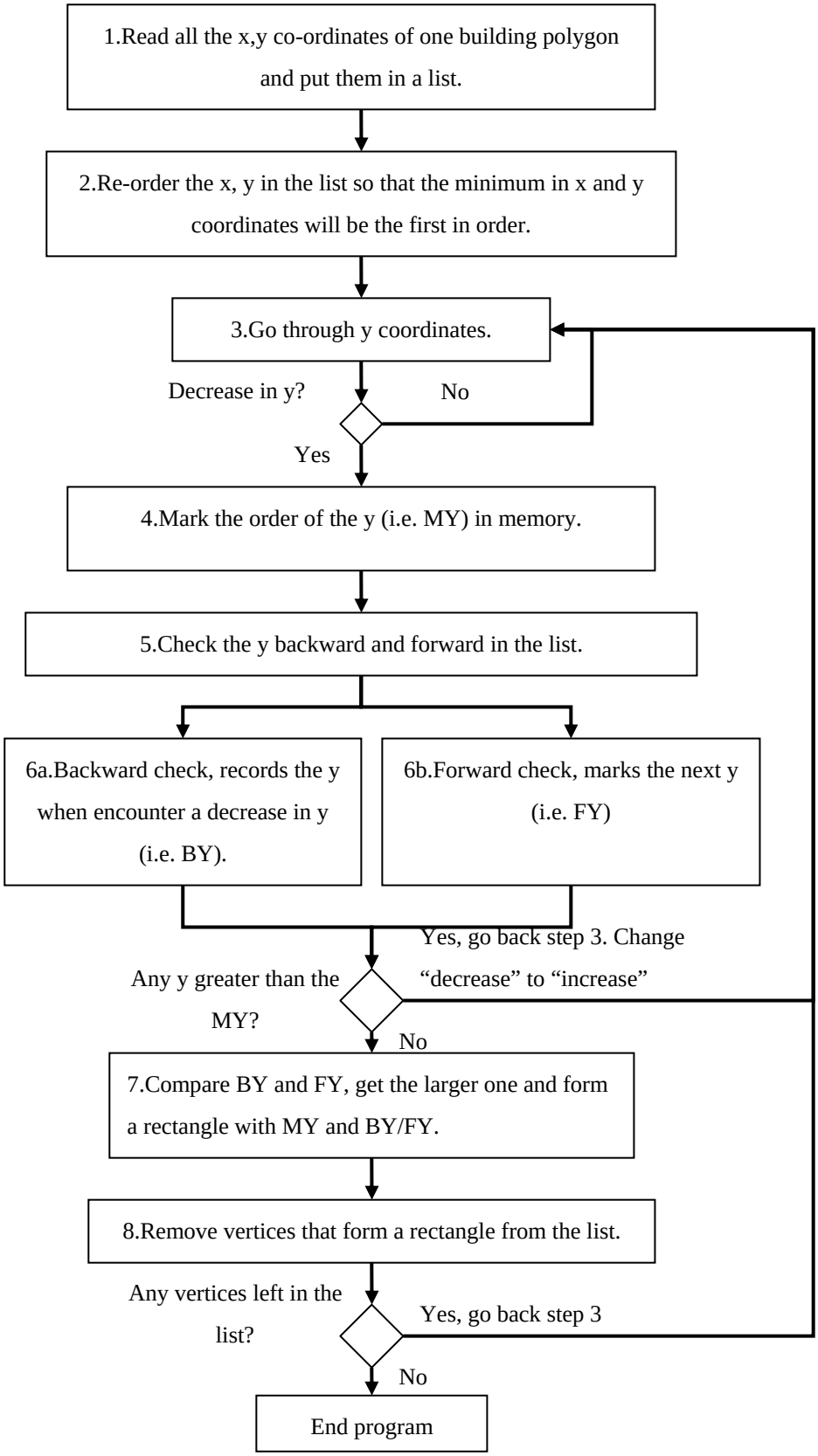


Figure 1.4. The logic of algorithm that divide buildings into rectangular blocks.

5.5 Objective 3: Examining the effect of spatial uncertainty towards the outcomes of QUIC dispersion modeling

With the uncertainty information from above, the final step is to examine how the spatial uncertainty of building data affects the results of the QUIC dispersion model. Two approaches are commonly used to examine the uncertainty: 1) the Monte-Carlo method and, 2) Taylor series analysis. The Monte-Carlo method processes the model with numerous perturbed copies of original input data and then examines the model outcome. It requires a large amount of computing time. Complimentarily, Taylor series analysis estimates the uncertainty of model outcomes by evaluating the derivatives of the output function. While the Taylor series analysis requires less computation by the substitution of mathematical formulas, it is limited to deterministic environmental models. Since the QUIC dispersion model is a stochastic model, this study employs a Monte-Carlo method to examine the effect of spatial uncertainty on the QUIC dispersion model.

The Monte-Carlo method starts by perturbing input data of the model. The input data are divided into three categories: 1) spatial, 2) meteorological and, 3) scenario. The spatial category includes building location and building dimension. The meteorological category includes wind direction, wind speed, reference wind speed, wind profile and reference wind height. The scenario category includes

emission source material, emission amount, release type, source type and emission location. However, in this study building location in spatial category and wind direction and wind speed in meteorological category are perturbed because these input data are the most common input for UADM. Uncertainty of building location is determined from objectives one and two, while uncertainty of wind speed and wind direction are based on ground observations from Joint Urban 2003 project (JU2003).

After perturbation of input data, sixty dispersion simulations according to one of the experimental setups at JU2003 (Table 1) are generated with the consideration of computation time and output data size. JU2003 is one of the largest atmospheric dispersion experiments in the United States. It aims to advance the knowledge of the contaminant movement inside an urban environment and thus improve atmospheric dispersion modeling. The JU2003 includes intensive measurements of meteorological variables and chemical tracers in the downtown area of Oklahoma City, Oklahoma. More than 200 portable wind sensors and tracer gas samplers were placed within the city for this experiment. With the ground observations from JU2003, the simulation results can be validated.

Table 1.1. The fixed input parameters for QUIC dispersion model.

Group	Parameters	Values
Meteorology Information	Wind profile	Power law
	Profile parameters	
	Uref(m/s)	1.5
	Zref(m)	10
	exp	0.3
Release Information	Gas material	Gasideal
	Total Mass release (g)	1000
	Release type	Instantaneous
	Source type	Point
	Source location (UTM)	X: 634775.02567
		Y: 3925884.38097
		Z:1.9
Simulation Information	No of Particles	2000
	Simulation time step (s)	5
	Simulation duration (s)	1200
	Particles output frequency (s)	30
	Concentration averaging time (s)	60
	Start time for concentration averaging (s)	0

Sixty simulations are further divided into two groups in order to examine the influence of spatial uncertainty. Each group contains thirty simulations. The first group of simulations is generated with perturbed meteorological data while the second group of simulations is generated with perturbed spatial and meteorological data. To identify the influence of spatial uncertainty, the perturbed meteorological data are the same for both groups of simulations. Differences in

two groups of simulations are mapped to show the spatial distribution. Moreover, the simulation results are validated with the ground observations from JU2003 because results from the simulations alone do not distinguish whether the spatial uncertainty of building data and the model parameters contributes positively or negatively to the dispersion model. With these methods, how spatial uncertainty may affect the accuracy of QUIC dispersion models in an urban area can be understood.

6. Organization of the Dissertation

The dissertation addresses the broad issue of spatial uncertainty in GIS and urban dispersion modeling in three stand alone articles that are embodied with introduction and conclusion chapters. The three objectives described earlier form the three respective chapters and each chapter will be prepared in forms ready for submissions to academic journals. This introductory chapter and the conclusion chapter in the end of the dissertation provide a common research prelude, summaries for coherence and integration of findings from the three objectives, as well as suggestions for future research.

References

Ahokas, E., H. Kaartinen, and J. Hyyppa. 2003. A quality assessment of airborne laser scanner data. Paper read at ISPRS Working Group 3 Workshop, 3-D reconstruction from airborne laserscanner and InSAR data, 8-10 October 2003, at Dersden, Germany.

Bregt, A. K., J. Denneboom, H. J. Gesink, and Y. v. Randen. 1991. Determination of rasterizing error: a case study with the soil map of the Netherlands. *International Journal of Geographic Information Systems* 5:361-368.

Burrough, P. A., and R. A. McDonnell. 1998. *Principles of Geographical Information Systems*. New York: Oxford University Press.

Congalton, R. G. 1997. Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion. *Photogrammetric Engineering & Remote Sensing* 63 (4):425-434.

Cullen, A. C., and H. C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment*. New York: Plenum Press.

Fedra, K. 1993. GIS and environmental modeling. In *Environmental Modeling with GIS*, eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert. New York: Oxford.

Goodchild, M. F. 1993. Data models and data quality: Problems and prospects. In *Environmental Modeling with GIS*, eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, 94-103. New York: Oxford.

Goodchild, M. F., B. Buttenfield, and J. Wood. 1994. Introduction to visualizing data quality. In *Visualization in Geographical Information Systems*, eds. H. M. Hernshaw and D. J. Unwin, 141-149. New York: John Wiley and Sons.

Hill, J. M., L. A. Graham, and R. J. Henry. 2000. Wide-area topographic mapping and applications using airborne LIght Detection and Ranging (LIDAR) Technology. *Photogrammetric Engineering & Remote Sensing*:908-914, 927, 960.

Hopkinson, C., L. E. Chasmer, G. Zsigovics, I. F. Creed, M. Sitar, P. Treitz, and V. M. Rober. 2001. Errors in LiDAR ground elevation and wetland vegetation height estimates. In *WGV/6 Workshop: 3rd Dynamic & Multi-dimension GIS*. Bangkok, Thailand.

Huising, E. J., and L. M. G. Pereira. 1998. Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry & Remote Sensing* 53:245-261.

King, J. L., and K. L. Kraemer. 1993. Models, Facts and the Policy Process: The Political Ecology of Estimated Truth. In *Environmental Modeling with GIS*, eds. M. F. Goodchild, B. O. Parks and L. T. steyaert, 353-360. New York: Oxford University Press.

Levy, J. I., J. D. Spengler, D. Hlinka, D. Sullivan, and D. Moon. 2002. Using CALPUFF to evaluate the impacts of power plant emissions in Illinois: model sensitivity and implications. *Atmospheric Environment* 36:1063-1075.

Maidment, D. R. 1993. GIS and hydrological modeling. In *Environmental Modeling with GIS*, eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, 147-167. New York: Oxford University Press.

Manomaiphiboon, K., and A. G. Russell. 2004. Effects of uncertainties in parameters of a Lagrangian particle model on mean ground-level concentrations under stable conditions. *Atmospheric Environment* 38:5529-5543.

Mead, D. 1982. Assessing data quality in geographic information systems. In *Remote Sensing for Resource Management*, eds. C. Johannsen and J. Sanders, 51-59. Iowa: Soil Conservation Society of America.

Moran, E. F., and E. Brondizio. 1998. Land-use change after deforestation in Amazonia. In *People and Pixeld: Linking Remote Sensing and Social Science*, eds. D. Liverman, E. F. Moran, R. R. Rindfuss and P. C. Stern, 94-120. Washington D.C.: National Academy Press.

Morrison, J. L. 1995. Spatial data qaulity. In *Elements of Spatial Data Quality,* eds. S. C. Guptill and J. L. Morrison. New York: Elsevier Science.

Mowrer, H. T. 2000. Introduction: the past, present, and future of spatial uncertainty analysis. In *Quantifying Spatial Uncertainty in Natural Resources: Theory and Application for GIS and Remote Sensing*, eds. H. T. Mowrer and R. G. Congalton, xv-xxiv. Chelsea, Michigan: Ann Arbor Press.

NRC, ed. 2003. *Tracking and predicting the atmospheric dispersion of hazardous material releases : implications for homeland security* Washington D.C.: National Academic Press.

Pardyjak, E., and M. J. Brown. 2002. Fast-response modeling of a two building urban street canyon. Paper read at 4th AMS Symposium on the Urban Environment, May 20-24, at Norfolk, Virginia, USA.

Piwowar, J. M., E. F. Ledrew, and D. J. Dudycha. 1990. Integration of spatial data in vector and raster formats in geographical information system. *International Journal of Geographic Information Systems* 4:429-444.

Schenk, T., B. Csatho, and D. C. Lee. 1999. Quality control issues of airborne laser ranging data and accuracy study in an urban area. Paper read at ISPRS Workshop: Mapping Surface Structure and Topography by Airborne and Spaceborne Lasers, at La Jolla, USA.

Sullivan, D. A., M. T. Holdsworth, and D. J. Hlinka. 2004. Monte Carlo-based dispersion modeling of off-gassing releases from the fumigant metam-sodium for determining distances to exposure endpoints. *Atmospheric Environment* 38:2471-2481.

USEPA. 1992. Guidelines for Exposure Assessment.

Wedhe, M. 1982. Grid cell size in relation to errors in maps and inventories produced by computerized map processing. *Photogrammetric Engineering & Remote Sensing* 48 (8):1289-1298.

Williams, M., M. J. Brown, and E. R. Pardyjak. 2002. Development of a dispersion model for flow around buildings. Paper read at 4th AMS Symposium on the Urban Environment, at Norfolk, Virginia, USA.

Zhu, A.-X. 2005. Research Issue on uncertainty in geographic data and GIS-based analysis. In *A Research Agenda for Geographic Information Science*, eds. R. B. McMaster and E. L. Usery, 167-223. Boca Raton, Fla: CRC Press.

Chapter 2 : Assessing Spatial Uncertainty of LiDAR-derived Building Model: A case study in Downtown Oklahoma City

Abstract

Light Detection and Ranging (LiDAR) technology enables cost-effective rapid production of digital models that capture topography and vertical structures of surface features at a fine spatial resolution. This capability has promoted LiDAR applications for mapping terrain, buildings, forest stands, and coastal features that cannot be adequately captured by other remote sensing means over a large area. However, in complex terrain, LiDAR data and LiDAR-derived products may contain significant uncertainty. This research provides a simple method to assess the spatial uncertainty of LiDAR-derived building model, using downtown Oklahoma City as an example. Results indicate that significant uncertainty could be found in urban environment where: 1) building structures are complex, 2) buildings are constructed with reflective materials, and 3) vegetation grows near-by. In addition, cities under fast development also challenge the accuracy assessment of 3D building models. To conclude, this study suggest: 1) careful pre-flight planning before data collection, 2) improve the feature extraction algorithm if possible, 3) use of other remote sensing data, and 4) accuracy assessment on suggested urban environments to reduce the spatial uncertainty of LiDAR data and LiDAR-derived products.

1. Introduction

Light Detection and Ranging (LiDAR) technology, also known as Laser Altimetry or Airborne Laser scanner, enables cost-effective rapid production of digital models that capture topography and vertical structures of surface features at a fine spatial resolution (Flood, 2001). The capability has promoted LiDAR applications for mapping terrain, buildings, forest stands, and coastal features that cannot be adequately captured by other remote sensing means over a large area. With such an advantage, LiDAR data have been applied as basic input data for a wide range of environmental models, such as models for estimating forest biomass, measuring coastal erosion, and calculating atmospheric dispersion in urban area.

Although LiDAR data provide fine resolution digital terrain models, the data do not guarantee quality results from environmental models. Especially in areas with complex terrain, LiDAR data may contain significant spatial errors. An overall accuracy assessment of a LiDAR dataset says nothing about the spatial distribution of LiDAR errors. In addition, the post data processing also introduces additional uncertainty to the environmental modeling. This study asserts that a spatial accuracy assessment of LiDAR data and post-processing (i.e. LiDAR-derived data) are critical to LiDAR applications, especially in urban area where terrain structure is complex and post-processes are required to generate building models.

This study uses the term, “uncertainty” here instead of “error” because the actual heights of buildings and other geographic features (such as tree canopies) are generally lacking. The term “error” is often interchangeable with the term “uncertainty” in the literature. While the terms are very similar in meaning, they represent different approaches in measuring the associated discrepancy. “Error” is used when measuring discrepancy in absolute terms while “uncertainty” is used when measuring discrepancy in relative term (Goodchild, Buttenfield, and Wood, 1994). Spatial uncertainty, nevertheless, can be assessed by comparison and relative measurements in the field.

With a comprehensive understanding of the spatial uncertainty of LiDAR data and LiDAR-derived data related to urban environments, target locations for groundtruthing and data correction can be identified. Even when a field survey is constrained by the limits of time and resources, an understanding of LiDAR accuracy and post-processing effects on data accuracy serves as a foundation for the interpretation of modeling outcomes, such as urban viewshed analysis and atmospheric dispersion modeling.

To discern spatial uncertainty of LiDAR data and post-processing effects in relation to urban environments, this study first examine the spatial uncertainty of LiDAR data and then the post-processing effects on spatial uncertainty. Next,

processes that can introduce spatial uncertainty to the building model are identified. The spatial uncertainty is related to building location (i.e. x, y coordinates) and height for each process. Finally, this study reviews the spatial distribution of uncertainty embedded in the LiDAR data and post-processes with digital ortho-photos and field measurements. The Central Business District (CBD) of Oklahoma City is used as the study area.

2. Background

LiDAR works similar to traditional radar technology, but rather than radio waves, LiDAR emits beams of light and captures the returned light. Combined with Global Positioning Systems (GPS) and Inertial Navigation Systems (INS), LiDAR theoretically can provide accurate vertical measurements, up to sub-meter accuracy, of ground objects. Raw LiDAR data consist of massive points with x, y, and z co-ordinates and are seldom used as an end product. LiDAR vendors usually provide various levels of LiDAR data to meet the needs of user's applications. Flood (2002) identified five levels of LiDAR deliverables: 1) Basic or "All-Points", 2) Low Fidelity or "First-Pass", 3) High Fidelity or "Cleaned", 4) Features layers, and 5) Fused. Higher-level LiDAR data attempts to provide more accurate information, and are refined with other sources of remote sensing data; and

consequently incurs higher cost and longer delivery time. Level-4 and Level-5 LiDAR products consist of extracted features and are extensively reviewed. LiDAR data from levels one to three are processed at increasing degrees of data filtering but have no classifications or feature identification.

Depending on the LiDAR product level and degree of detail required by the users, many methods have been developed to derive a 3D building model from LiDAR data. General procedures involve separation of non-ground points from ground points, segmentation of different objects on the ground, generation of bare earth elevation, and building boundary detection and regularization. Most procedures are processed through customized automatic algorithms by LiDAR data providers with the assumption that buildings are rectangular in shape and have flat roof surfaces. However, the assumption is violated in downtown areas where buildings have complex structures and roof tops. Manual refinement becomes necessary to identify problematic locations and adjust heights with other data sources.

In addition to the assumption of simplified building geometry, there are other sources of errors or uncertainties in LiDAR data. Through various procedures, errors or uncertainties can be introduced. With extensive use of laser scanning for topographic measurements, Huising and Pereira (1998) outlined three main sources

of error: 1) the laser system, 2) data collection and processing, and 3) the target surface. The laser system can introduce errors caused by laser pulse delay, GPS and INS misalignment. Errors from data collection and processing are due to mistakes in flight line planning, system calibration and data filtering. The reflectivity of target surfaces can induce errors for asphalt and water surfaces which do not reflect laser back to the system. Moreover, Huising and Perreira (1998) classified the terrain surface into six groups: 1) flat paved, 2) flat barren, 3) flat grass and scrubs, 4) hilly paved, 5) hilly barren, and 6) hilly grass and scrub, and examined the error for each surface category. They found that LiDAR data acquired on grass and scrubs terrain were less accurate than other surfaces and errors could reach up to 0.3 meters (30 percent). Similarly, Hopkinson et.al. (2001) studied LiDAR's error on ground elevation and wetland vegetation height. Their results showed largest errors were associated with low shrub, tall vegetation classes and aquatic vegetation. In addition to vertical accuracy, Alharthy et.al. (2004) assessed the planimetric accuracy of LiDAR data on flat terrain and found that the planimetric accuracy varies with the swath width. Larger errors could be found at the end of the swath width. Ahokas et.al. (2003) examined the LiDAR elevation error associated with flight lines, flight altitudes and observation angles. They concluded that flight lines might generate both random and systematic error, higher flight altitudes generated

greater random error, and observation angles induced systematic error up to ten centimeters. Most of these citied studies examined LiDAR's uncertainty regarding to vegetation cover in rural area.

In urban areas, 3D building modeling involves additional sources of uncertainty that are not discussed in the studies above. They come from data post-processing such as feature extraction algorithm and manual digitizing, which outline the buildings footprints and structures. Feature extraction involves numbers of procedure such as data segmentation, filtering, boundaries detection and smoothening, but the algorithms behind the procedures can vary from vendor to vendor. It can be a critical source of uncertainty in 3D building models as a small displacement of the laser footprint can result in a large vertical discrepancies on building edges (Schenk, Csatho, and Lee, 1999). Yet, feature extraction algorithms are hard to evaluate because ground information in urban area is difficult to obtain. For example, tall buildings blockage in urban downtown area limits the use of GPS. Private ownership or security issues restrict access to building roof features, and direct measurements of buildings dimensions. Manual identification of building footprints also introduces uncertainty to the 3D building model that is hard to evaluate without ground truth information. Therefore, 3D building models inherit

not only the uncertainty embedded in LiDAR data, but also uncertainty introduced by the post-processing of the LiDAR data.

Nevertheless, uncertainty assessments of raw LiDAR data have been well addressed in flat terrain. For example, a common approach is to interpolate the raw LiDAR data points into a raster layer and then compare it with existing DEMs from traditional photogrammetry. It is also common to compare the interpolated LiDAR data with ground control points (GCPs). By comparing the differences between raw LiDAR data and existing ground information, mean error with a confidence level can be calculated using statistical techniques. However, different interpolation methods may generate different terrain elevation and result in uncertainty up to one meter (Smith, Holland, and Longley, 2003, 2004).

The other way to assess the accuracy of raw LiDAR data is to measure the relative accuracy, by comparing overlapping strips and calculating the relative offset in height for the same area (Latypov, 2002; Mass, 2002). Obviously, good "relative accuracy" using the method only means better consistency within the laser system.

This research focuses on assessing the spatial distribution of LiDAR uncertainty and LiDAR-derived building model uncertainty, using a LiDAR dataset covering the Central Business District in downtown Oklahoma City as example. To understand the cause of uncertainty distribution, this research further examine the

urban environments (i.e. building characteristics, arrangements and vegetation growth) in the study area. The following section describes the research design, followed by results and discussion.

3. Research design

This study first describes the LiDAR data and other data sources, then the study area, and the procedures to generate a 3D building model from the LiDAR data. Following that, adjustment strategies to develop a refined 3D building model are elaborated and the LiDAR data is compared with the refined model as a basis to assess LiDAR uncertainty. The results of feature extraction and manual digitizing with the refined model are compared to examine the spatial uncertainty of LiDAR-derived building data. Finally, this study compares the uncertainty embedded in LiDAR data and uncertainty in LiDAR-derived building data in the study area (Figure 2.1).

3.1 Study area and data sources

The study area (approximately 0.8 square miles) is located in the CBD of Oklahoma City (OKC) with a flat terrain and well-defined central city area (Figure 2.2). Similar to other CBDs, this area is dominated by high-rise buildings in the

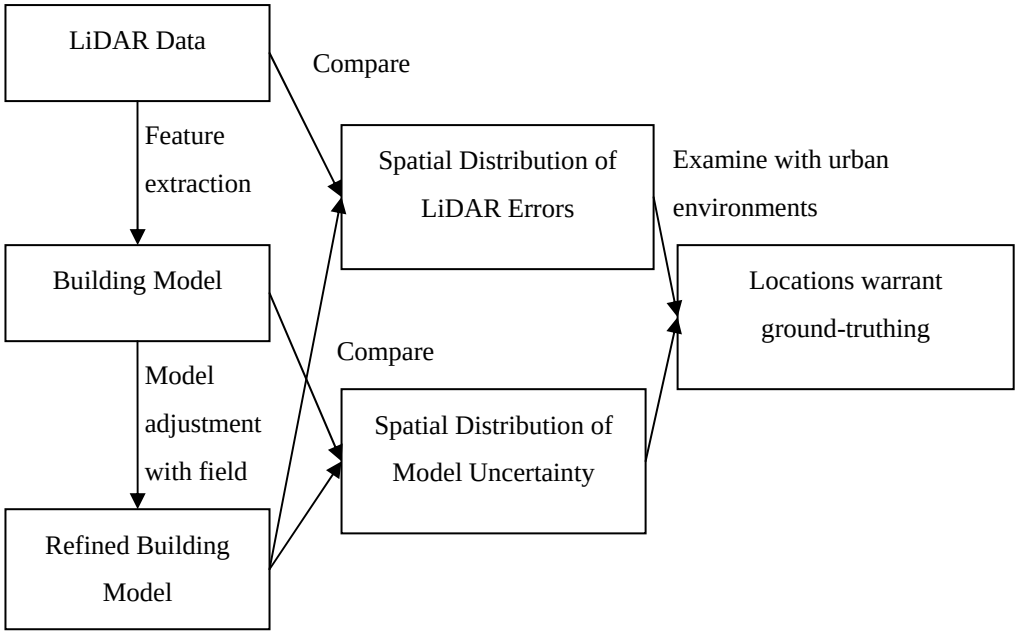


Figure 2.1. The conceptual flow chart of the research design.

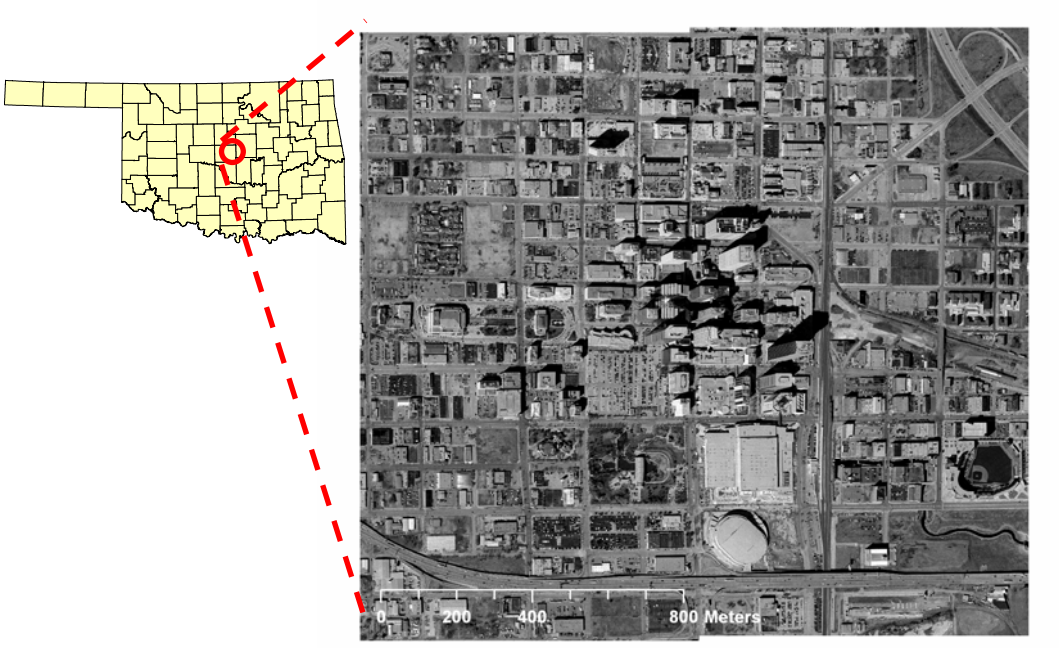


Figure 2.2. Study area: Central Business District of Oklahoma City. Aerial photo from: U.S.G.S., date: March 2002.

center with surrounding open areas and low-rise commercial buildings. With the flat terrain in OKC, LiDAR uncertainty that caused by the terrain roughness can be minimized.

The LiDAR data was collected in late October 2001 by the Optech Airborne Laser Terrain Mapper (ALTM) 2033 sensor at an operation altitude around 2000 to 2500 meters and swath width of 540 meters. The Joint Precision Strike Demonstration (JPSD) Program Office of the U.S. Army executed the flight plan, geo-referenced, and geo-rectified the LiDAR data. LiDAR data products from the JPSD Program Office include the first return DEM, second return DEM, intensity layer and a color LiDAR image in one meter resolution (Figure 2.3). First return DEM records the first return of laser pulse, while the second return DEM captures second return of laser pulse. Therefore, the first return DEM usually reflects the very top part of the object and the second return DEM reflects the lower part of the object if the laser beam can pass through the object. The color LiDAR image shows the lowest elevation in blue and the highest elevation in red. The data come in GeoTIFF format with a header file specifying the co-ordinates and projection information. The general accuracy of the LIDAR data is about 0.3 meters for vertical measurements and 0.5 meters for the horizontal measurements, with a 90 percent confidence level.

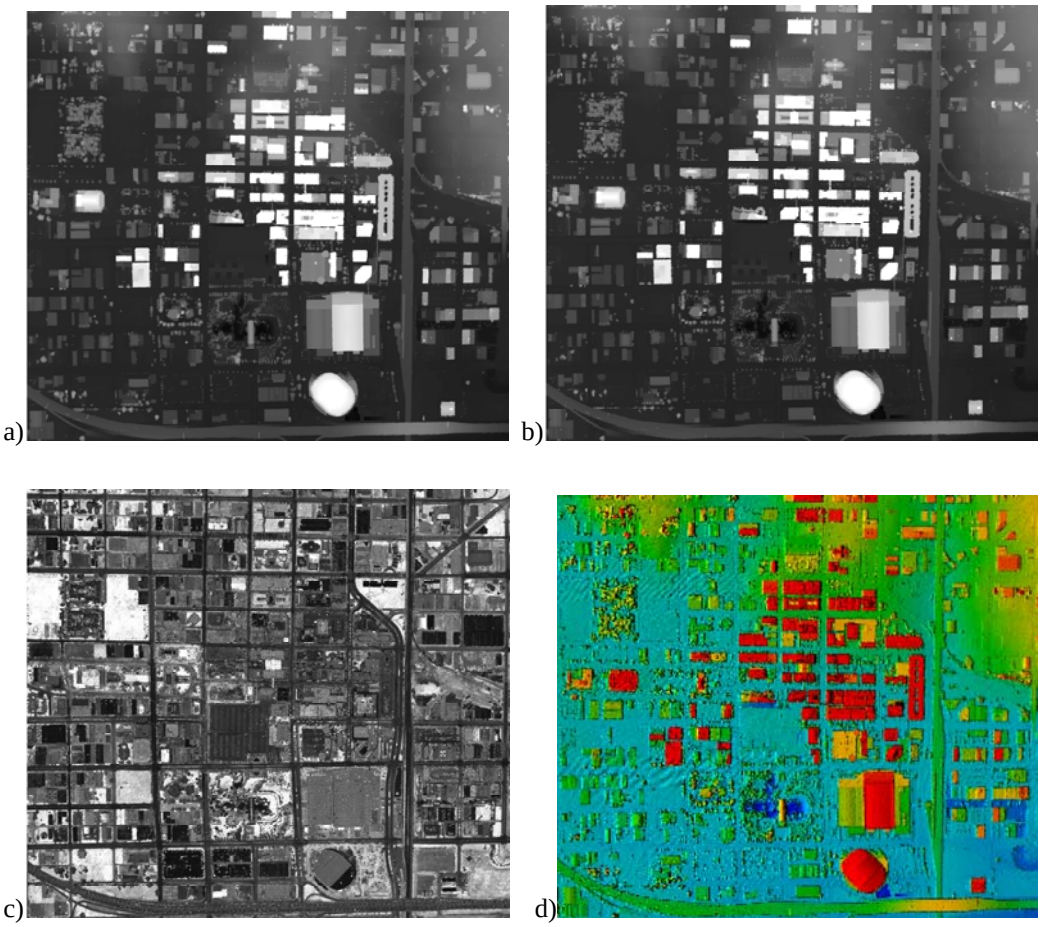


Figure 2.3. The LiDAR data: a) first return DEM, b) second return DEM, c) Intensity layer, d) color-code image.

In addition to the LiDAR data, United States Geological Survey (USGS) high resolution ortho-photographs are used to aid the construction of a refined 3D building model in Oklahoma City (Figure 2.4). The ortho-photograph is an aerial photograph in which distortions caused by terrain and sensor orientation have been removed mathematically. The ortho-photograph used here has 0.3 meter resolution with horizontal accuracy around 1 meter. The image was captured in March 2002.

Figure 2.4. USGS orthophotograhy of the Oklahoma City. Date: March 2002.

3.2 Procedures of generating a 3D building model with LiDAR data

To construct a 3D building model with LiDAR data, the RTV LiDAR Toolkit originally developed by the JSPD program is used. The LiDAR Toolkit is an extension of ArcView 3.x (Environmental System and Research Institute), and it contains functions that include visualization, extract bare earth, buildings, vegetation, roads and network features from LiDAR data. The toolkit has been used in

numerous projects funded by the National Geospatial-Intelligence Agency (NGA), U.S. Geological Survey (USGS) and U.S. Army Topographic Engineering Center (SAIC, 2005).

A bare earth surface for the study area is first computed by the algorithm based on the first and second return DEM. Then, the bare earth is subtracted from the first return DEM to derive object heights. Building footprints are generated by the feature extraction algorithm in LiDAR toolkit. To obtain the detail of building structures within a footprint, this study classifies the objects according to height and converted them into polygons. Next, building layouts and structural details are delineated with the aid of aerial photographs, color LiDAR imagery and field measurements. The refinement procedures include manual adjustment of building footprints based on ortho-photographs and field measurements. The height of each feature is assigned by the majority object height (Figure 2.5). Also, additional field observations and comparative estimates are used to discern ambiguous areas, such as buildings blocked by shadows in the aerial photos or noise in the color LiDAR image. Finally, a refined 3D building model is generated (Figure 2.6).

From the refinement procedures, two main steps that may introduce spatial uncertainty to the 3D building model are identified: 1) feature extraction, and 2) manual digitizing. These two steps are identified because they are primary

procedures in creating a 3D building model regardless of the methods selected for 3D building generation. The following paragraphs describe the methods used to assess and quantify the spatial uncertainty of LiDAR data and LiDAR-derived building dimensions.

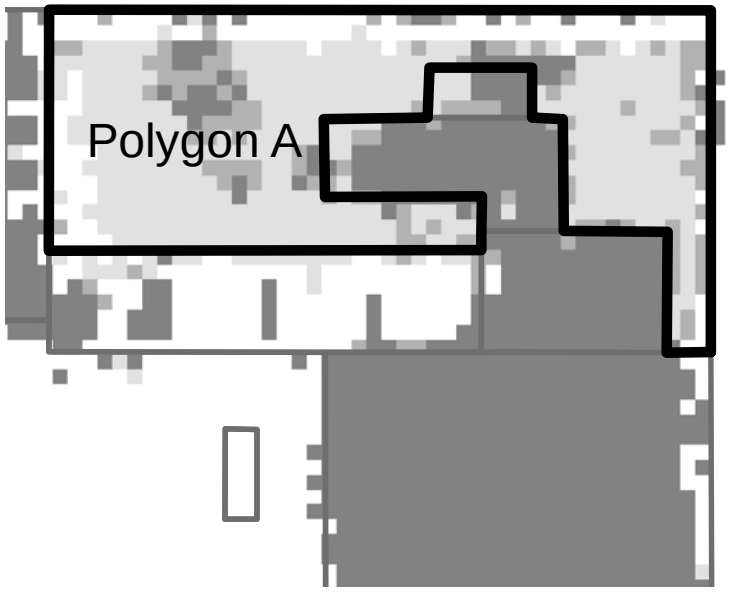


Figure 2.5. Examples of determining building height by the majority rules. White area represents elevation lower than 43.5m; light grey represents elevation between 43.5m to 44.5m; medium grey represents 44.5m to 46.5m; darkest grey represents elevation higher than 46.5m. Therefore, polygon A is assigned 44m.

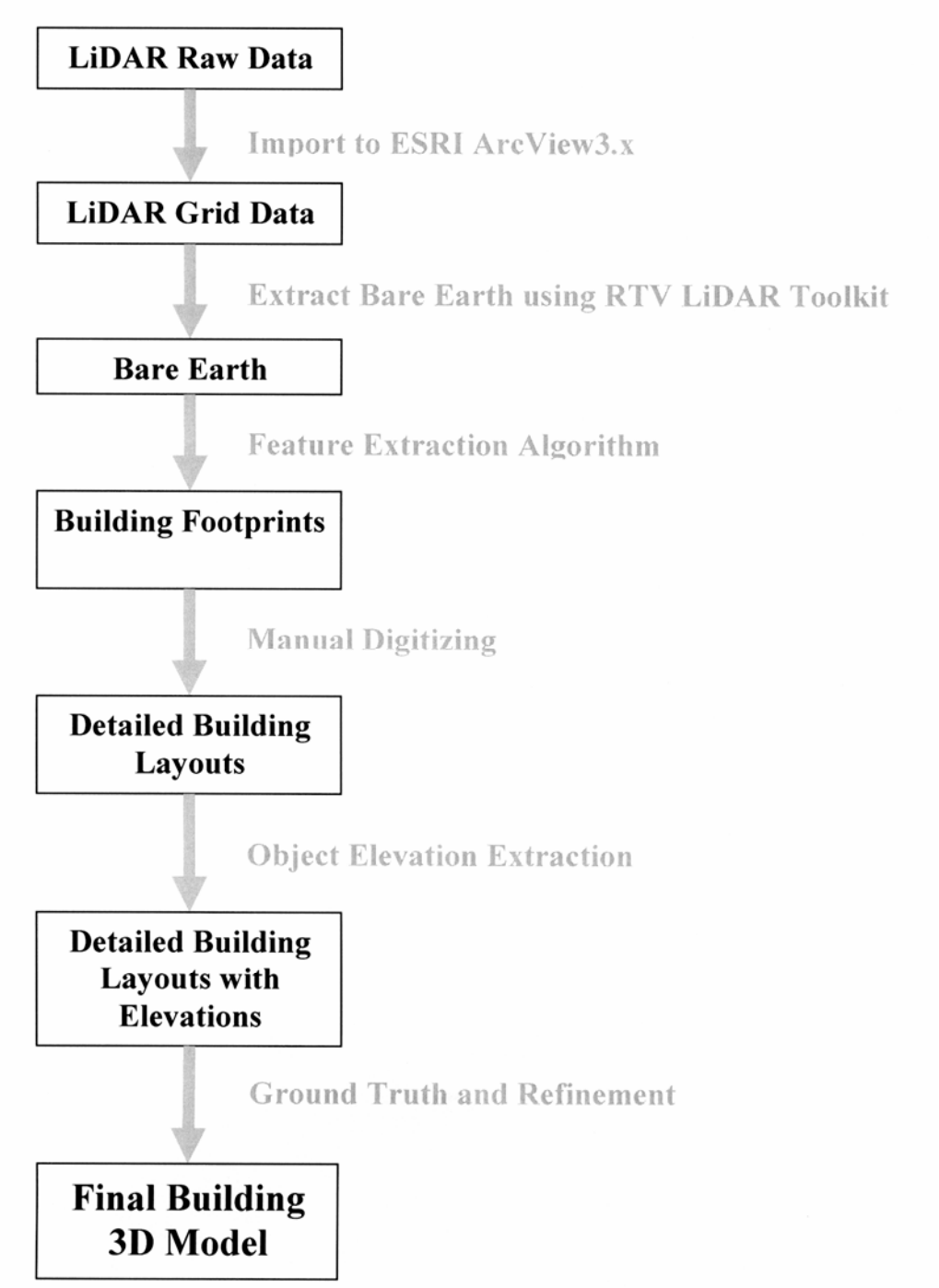


Figure 2.6. The flowchart of generating 3D building model from LiDAR data.

3.3 Spatial assessments of LiDAR data and LiDAR-derived building data

To assess the spatial uncertainty of LiDAR data, the LiDAR data is compared with a refined building model. With adjustments based on ortho-photograph and ground observations, this study assumes the refined 3D building model represents a more accurate building model. Two issues arise when comparing the LiDAR data with the refined 3D building model. The refined 3D building model stores buildings as polygons, which cannot accurately capture elevations along sloping building surfaces (e.g. parking ramps or tilted roofs), as only one elevation can be assigned to a given polygon. Also, the refined model needs to be converted to raster in order to compare with the LiDAR data. Both issues can introduce disparity between the polygon building footprints and LiDAR data. Therefore, the following procedures are utilized to reduce the differences caused by data conversion and data representation.

First, the refined 3D building model is converted to a grid based on the elevation attribute of building polygons. Then, the building height grid is subtracted from the LiDAR data and classified the differences into five classes: 1) 3 meters and above, 2) 1 to 3 meters, 3) 1 to -1 meter, 4) -1 to -3 meters, and 5) -3 meters and less. Positive values indicate that LiDAR height measurements are higher than the heights adjusted in the refined building model, while negative values

indicate the opposite. Most differences with +/- one to three meters (i.e. classes 2 to 4) result from sloped surfaces or equipment on roofs, these differences are acceptable for the 3D modeling. Moreover, differences around one meter of building edges are removed to eliminate the differences caused by vector-raster transformation. Areas with differences more than three meters (i.e. class one and five) are identified as uncertainty for both the LiDAR data and the refined 3D building model. Finally, field observations and assessments are made in order to identify the causes of uncertainty.

For LiDAR-derived data, feature extraction algorithms and manual digitizing both introduce spatial uncertainty to building boundaries. To examine the spatial uncertainty of building boundaries introduced by feature extraction algorithms, this study computes the distances based on x and y co-ordinates between the vertices of automatically generated building footprints and vertices of refined building footprints. Locations with differences greater than two standard deviations from the mean are further examined with field survey and aerial photographs inspection.

To examine the spatial uncertainty introduced from manual digitizing, fifteen graduates and fifteen undergraduates students who study in geography department at the University of Oklahoma are asked to digitize a sample building footprint based on the color LiDAR image (Figure 2.7) at the same scale (i.e. 1:1000). Instructions

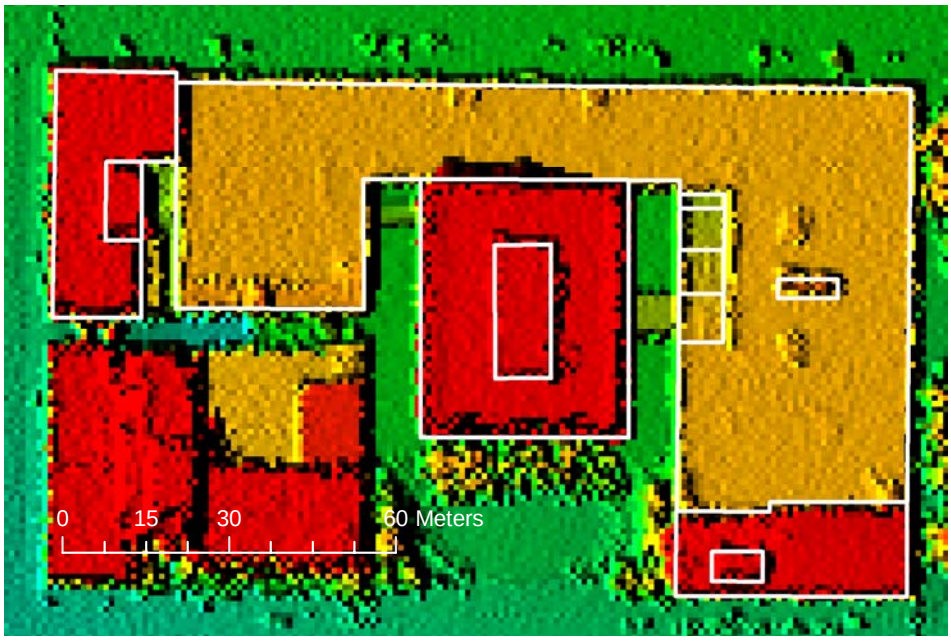


Figure 2.7. The building sample for digitizing test. The building is outlined by white lines. LiDAR color image is used as background image for digitizing test.

on how to digitize building on-screen using ESRI ArcMap and building orth-photograph were given prior to digitizing. A sample building from the CBD was chosen to represent the typical building structure in the study area. It is assumed that spatial uncertainty found in the sample building is representative of uncertainty embedded in the other building footprints captured in the refined building model. The elevation difference between the building structures is significant with the highest elevation at 119 meters and the lowest elevation at four meters. The sample building is also representative of a range of structural

complexity as it contains a parking lot, tilted roof, gaps and vegetation near one of the corners. Then the variance of x, y coordinates digitized from the 30 subjects is computed. Unlike feature extraction algorithms, only a sample of the building is examined because of time and labor constraints. Therefore, it is assumed the spatial uncertainties found in the sample will be the representative of the uncertainties found in the other buildings.

4. Results and Discussions

4.1 Spatial distribution of LiDAR uncertainty

Figure 2.8 shows the spatial distribution of LiDAR uncertainty with elevation differences greater than three meters between the raw LiDAR data and the refined building model. About five percent of the total building area was identified as uncertainty in the LiDAR data. The spatial uncertainty was broadly distributed across the study area. However, larger patches of uncertainty appeared in the outer area while noise-like smaller patches primarily occurred in the inner part of the study area.

Figure 2.8. Area in black show potential LiDAR errors with more than 3 meters differences from the refined building model.

Every patch of LiDAR uncertainty was further inspected in the field. Five urban conditions that were associated with large differences between the LiDAR data and the 3D building model were summarized. The first three conditions suggested LiDAR error, while the last two conditions related to 3D building model construction.

The first condition was found around building gaps. For example, the gap in the First National Center (Figure 2.9) should be at the ground level, but the LiDAR data shown a height around 30 meters in the gap. On the east side of the First National Center building, the gaps in Sante Fe Parking Garage also had similar differences, but this occurred only near the edge of the gap. The LiDAR data suggested a higher elevation at narrow gaps (i.e. either length or width was less than ten meters), where the laser pulse might be reflected back to the sensor before reaching to the ground.

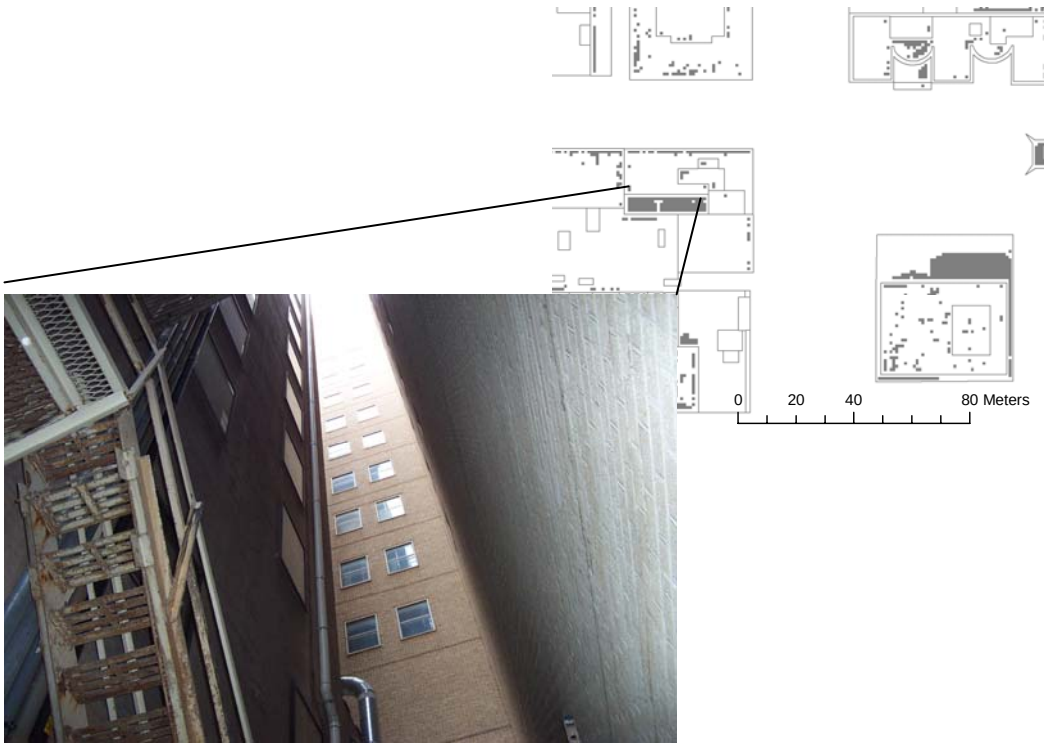


Figure 2.9. LiDAR errors caused by gaps in the First National Center.

The second condition was found around the highest building, the Bank One Tower, in the study area. The Tower is 150 meters tall with a platform at about eleven meters high. However, the LiDAR data suggested a portion of the platform up to 54 meters (Figure 2.10a). In addition, a filtered effect was found in that area (Figure 2.10b). The area was mis-captured by LiDAR data. However, the reason for the error could not be verified with the data provider.

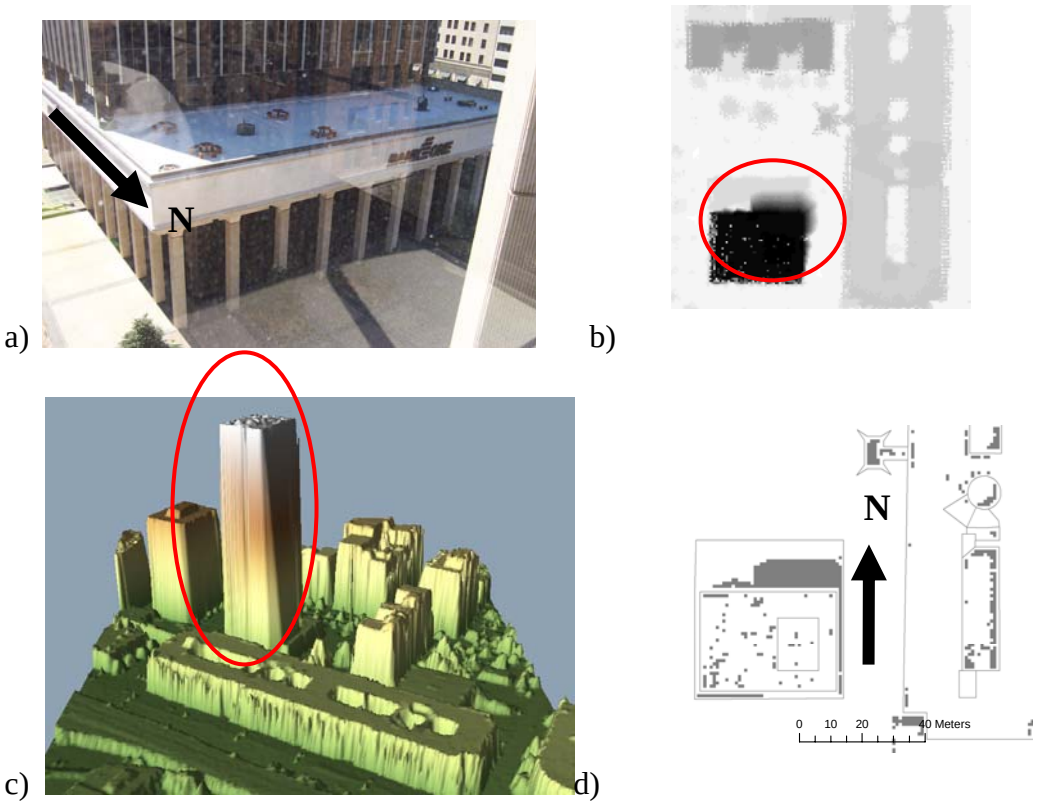


Figure 2.10. LiDAR errors in Bank One Tower. a) Photograph showing the actual view of Bank One Tower, b) a filtered effect of elevation was found in LiDAR data, with color scheme from white to black, represents heights from the lowest to the highest, c) the 3D view of the filtered effect of Bank One Tower, d) area with differences greater than 3 meters between LiDAR data and refined building model.

The third condition was identified around glass surfaces, such as Leadership Square with glass exteriors. Glassy material generated noisy signals in the LiDAR data, but as shown in Figure 2.11, the noise surrounded the glassy building edge instead of the interior of the building footprint.

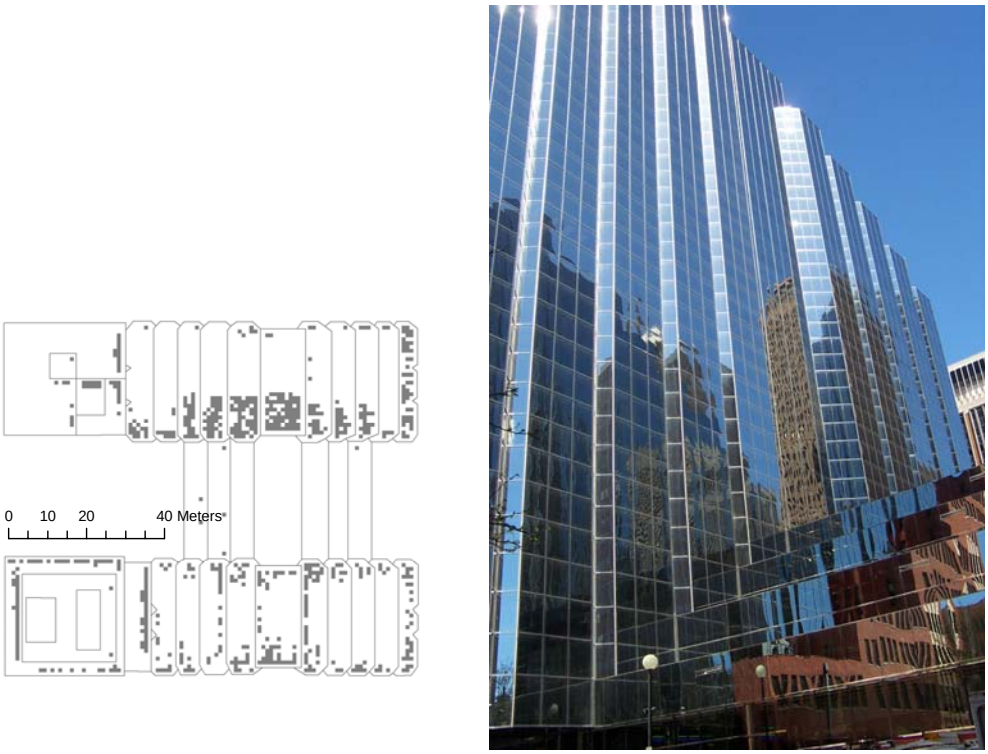


Figure 2.11. Glassy material of the Leadership Square generates noise for the LiDAR data.

The above three conditions introduce much uncertainty in LiDAR data and deserve special attention for ground-truthing. As significant elevation differences and various kinds of building structures are common in urban downtown areas, there are high probabilities for laser reflected back to the sensor before reaching the ground. Anticipation of these three conditions is crucial to flight route planning in

LiDAR data collection. Extra flight routes perpendicular to original flight routes can reduce the chance of missing data or potential error inside building gaps and hence increase the accuracy of LiDAR data. However, if the building is covered by glassy material, data noise may not be avoided and manual adjustments are needed.

Beside the three conditions related to building structures, vegetation on top of buildings and construction sites were two other sources identified in the study area. Vegetation like trees was detected with the feature extraction algorithm. However, when a tree grew over or on top of a building, the building could not be extracted. Moreover, the algorithm proved incapable of detecting building boundaries when vegetation was planted near or around the buildings. An example was found at a drive-through bank located north of the study area (Figure 2.12). Construction sites are another potential source of uncertainty in the 3D building model because building construction might already be completed by the time user received the LiDAR data (Figure 2.13). As a result, the building characteristics were outdated upon the release of the LiDAR data.

Moreover, there are some salt and pepper effect and small patches of LiDAR uncertainty located around the center of the study area. These discrepancies were mainly caused by the vector-to-raster conversion which occurred over the one meter buffer of the building polygons.

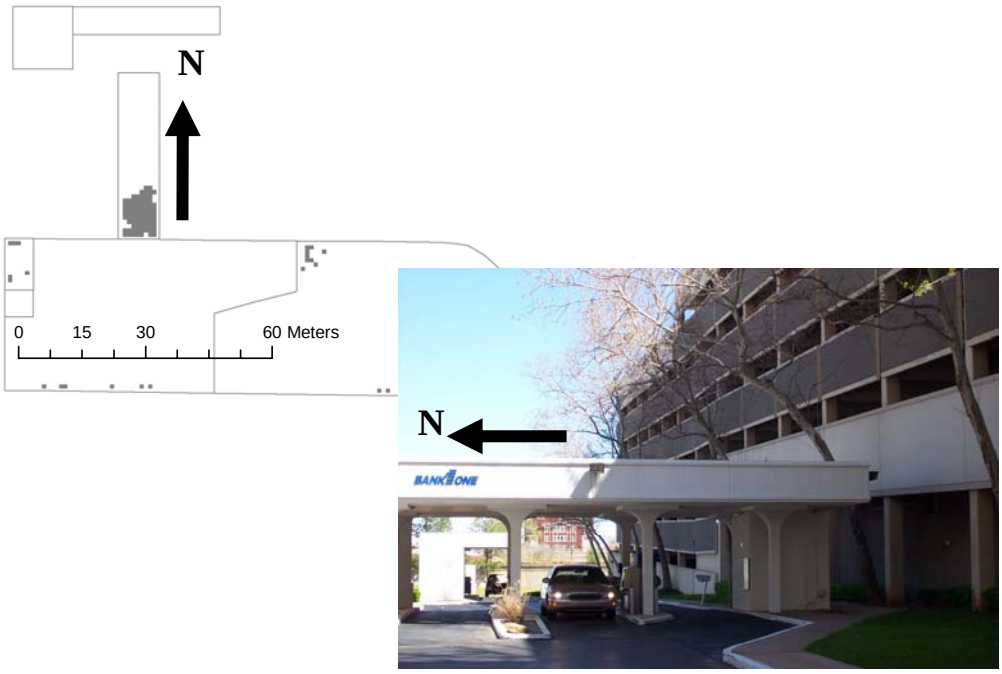


Figure 2.12. Vegetation on top of the drive-through bank causing differences between LiDAR data and building model.

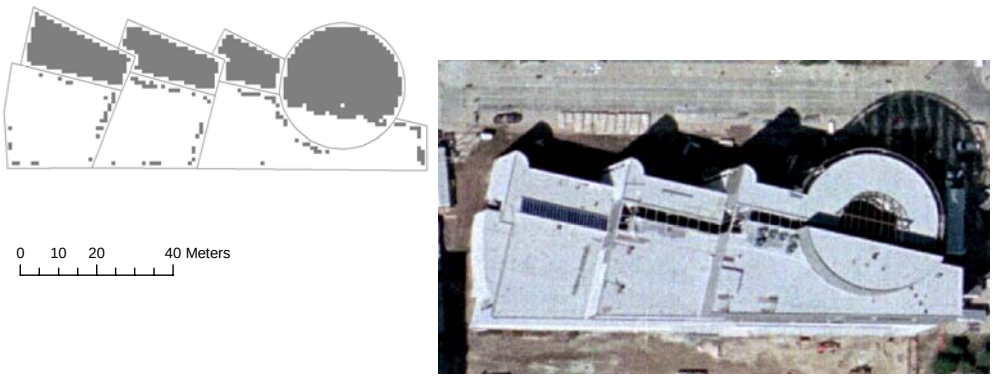


Figure 2.13. Example of construction site in Oklahoma City, it causes uncertainty in LiDAR data and feature extraction algorithm.

4.2 Uncertainty from feature extraction algorithm

Figure 2.14 shows an overlay of automatically generated building footprints and refined building footprints. Light grey footprints represented buildings delineated by both the automatic algorithm and the refined model, which was about 91 percent of total building footprint area. Dark grey footprints were defined by the refined model but not by the automatic algorithm, while black footprints were extracted by the automatic algorithm but not by the refined model. About 8.7 percent of the total refined building footprint area was not detected by the automatic algorithm and they were mainly located around the north-west and south-east parts of the study area. Of the total automatically generated footprint area, 5.3 percent were not classified as building in the refined model, and these footprints appeared on the southern edge of the study area.

Figure 2.15 shows the Euclidean distances between building vertices in the refined model and the automatically generated building footprints. Small black dots are used to represent distances of less than one standard deviation; medium black circles represent distances between one and two standard deviations; and large black circles represent distances greater than two standard deviations. Comparatively large distances were found on the east, south and west part of the

study area. Overall, the average distance between the automatic footprints and the refined footprints was 4.3 meters with a standard deviation of 12.8 meters.

Figure 2.14. Overlay automatic building footprint with refined building footprint. Grey shows area of intersection; black shows area that are classified as building by algorithm but not by the refined model; medium-dark grey shows area that are classified as building by refined model but not the algorithm.

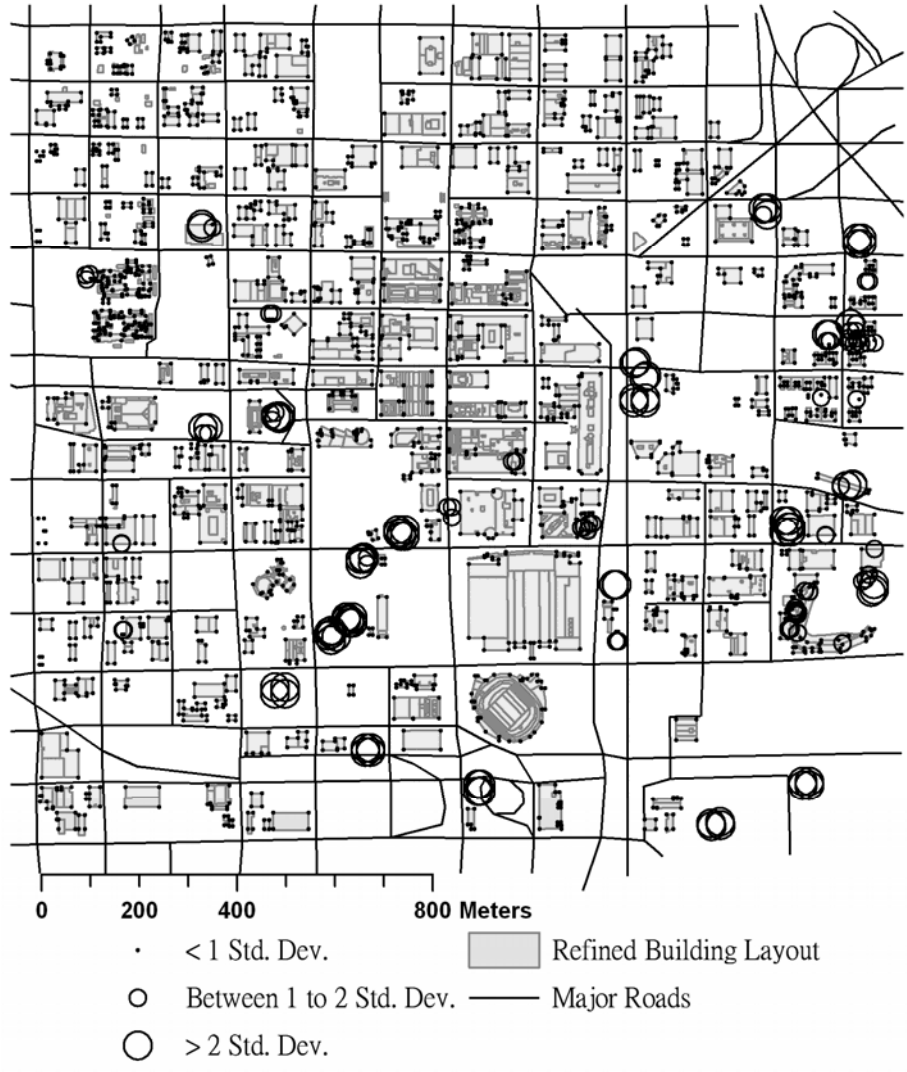


Figure 2.15. The distribution of differences between automatic building footprint and refined building footprint; measured by the distances between vertices of two building footprints. Black dots represent distances less than 1 standard deviation (<17.1m); medium circles represent distance within 1 to 2 standard deviations (17.1 – 29.9m); and large circles represents distance greater than 2 standard deviation (>30m).

Cases with differences greater than two standard deviations were due to mis-classification of trees or small rectangular objects as buildings by the feature extraction algorithm. Examples of such mis-classification include an individual tree inside the park, cargo containers, bridges and cover parking canopy (Figure

2.16). These errors are commission errors resulted by the feature extraction algorithm. If the commission error is omitted, the average and standard deviation values will be greatly reduced to 2.29 meters and 3.5 meters respectively. Unlike the spatial distribution of LiDAR uncertainty, the above results suggest that the major spatial uncertainty of feature extraction is distributed around the outer range of the study area where parks, cargo containers are most likely present. Moreover, construction sites and complex building structures are also causes of spatial uncertainty; one example of complex building structure is the baseball stadium on the east side of the study area (Figure 2.17).

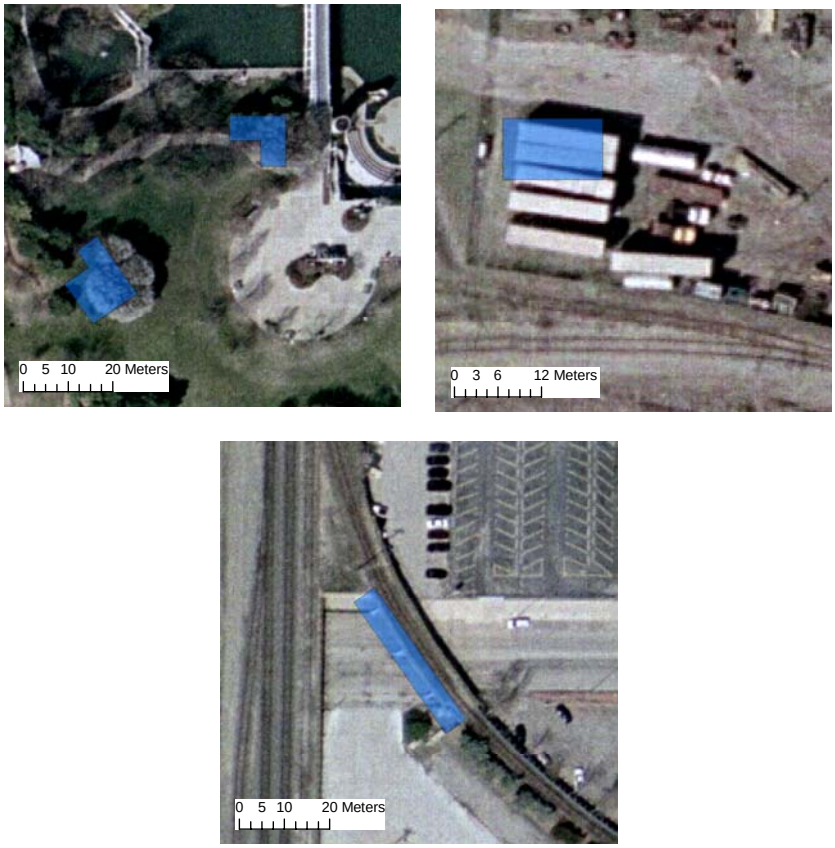


Figure 2.16. Examples of objects that are mis-classified by the feature extraction algorithm. Top: trees (left) and cargo (right); bottom: bridge.

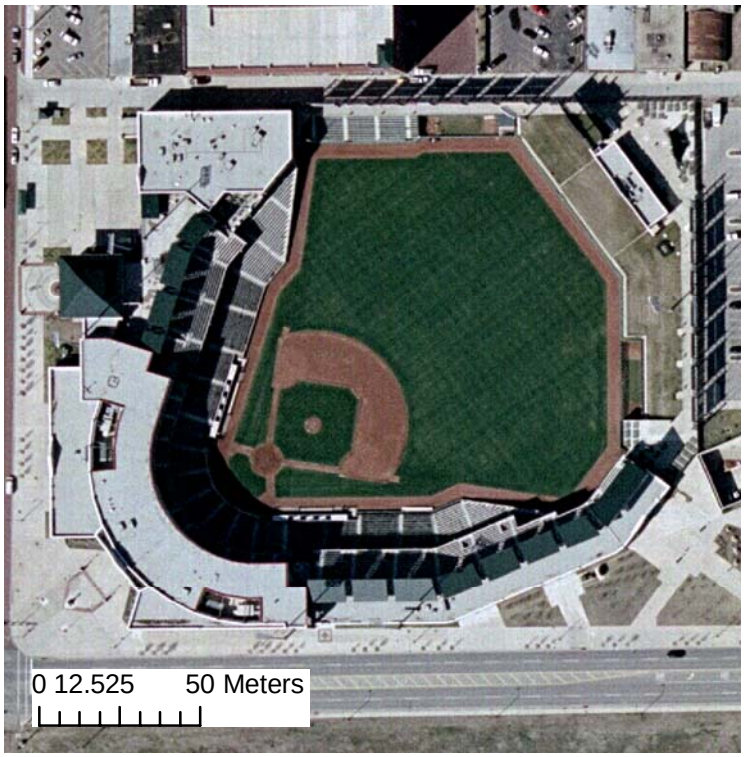


Figure 2.17. Example of complex building structure which is difficult to be detected by feature extraction algorithm. A softball stadium in downtown Oklahoma City.

4.3 Spatial uncertainty from manual digitizing

Figure 2.18 shows the results of differences in manual digitizing by thirty students. Each of the seventeen building corners was labeled with an ID number. Among all the corners, corners five, and thirteen through sixteen showed a more dispersed pattern than the others. The variance of x and y co-ordinates for each corner were computed (Table 1). On average, the digitizing variances were 1.01 square meters for x coordinate and 0.63 square meters for y coordinate. The greatest digitizing variance for x coordinate, 3.56 square meters, occurred at corner

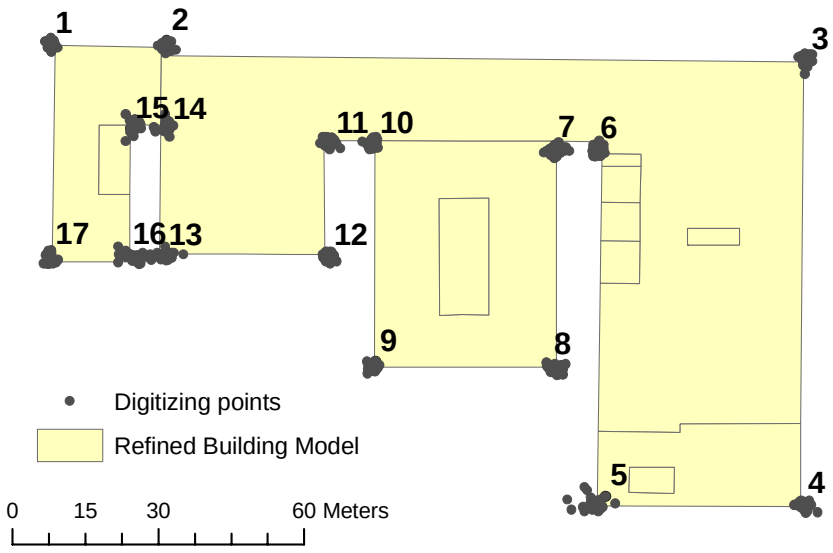


Figure 2.18. Uncertainty of manual digitizing. Samples from thirty students.

Table 2.1. The variances of x, y coordinates computed from samples of manual digitzing.

ID	Variance in x	Variance in y
1	0.23	0.35
2	0.86	0.63
3	0.50	0.71
4	0.79	0.44
5	3.05	1.08
6	0.49	0.64
7	1.29	0.62
8	1.31	0.50
9	0.45	0.41
10	0.50	0.46
11	0.69	0.49
12	0.47	0.30
13	1.25	0.34
14	0.72	1.00
15	0.65	1.37
16	3.56	0.66
17	0.34	0.65
Average	1.01	0.63

sixteen, while the greatest digitizing variance for y coordinate, 1.37 square meters, occurred at corner fifteen. Corner five had the second greatest digitizing variance in both x and y coordinates (i.e. 3.05 and 1.08 square meters respectively).

High spatial uncertainty could be found at corner five because there was vegetation near the corner of the building where the boundary becomes blurry. Corners thirteen to sixteen also contain significant spatial uncertainty as they were close to each other. When compared to the distances between corners six, seven, ten and eleven, the distances between corners thirteen to sixteen were smaller. Distances between corners six and seven, ten and eleven were around ten meters while the distances between corners fourteen and fifteen, thirteen and sixteen were only around six meters. Therefore, vegetation cover near the building and gaps smaller than six meters tended to generate a greater degree of spatial uncertainty during manual digitizing.

4.4 The influences of urban environments on the spatial uncertainty of LiDAR-derived building model

To summarize, three main factors related to urban environments have notable influences on LiDAR-derived building data: 1) vegetative interference, 2) complexity of building structures and, 3) building materials. Vegetation

interference introduces spatial uncertainty during LiDAR data acquisition, feature extraction and manual digitizing. Although trees can be distinguished by feature extraction algorithms, trees also mask out building footprints where trees grow adjacent to buildings or grow over the roofs.

Complex building structure, such as building with gaps of less than six meters, can prohibit the penetration of laser pulses, consequently resulting excessive LiDAR elevation values. Small gaps blur building boundaries and also challenge identification of building boundaries during manual digitization. Moreover, most feature extraction algorithms assume a rectangular building shape, and hence are inadequate in classifying complex buildings structure. In addition, rectangular non-building objects (e.g. cargo containers and bridges) may also lead to mis-classification.

The third major factor, building material, results from reflective characteristics of surface materials. Glassy material reduces the accuracy of LiDAR data, feature extraction algorithms and manual digitizing as it introduces noise to the LiDAR data and hence generates blurred boundaries.

Besides the three major factors, many cities are undergoing urban development or urban renewal projects. These fast-changing city landscapes pose

additional challenges in developing 3D building models as buildings in a construction site would have been completed after LiDAR data collection.

Urban environments are complex and dynamic. To reduce the spatial uncertainty of a 3D building model, several suggestions are offered. First, pre-flight planning in data acquisition helps to reduce LiDAR uncertainties. Lower flight attitude, smaller swath angles and perpendicular flight routes also help to reduce LiDAR uncertainty. Second, even though it is hard or impossible to modify feature extraction algorithms as they are usually provided by the system vendors, the spatial uncertainty can be reduced by referencing other remote sensing data such as InSAR and ortho-photographs. Data fusion techniques have been suggested as a critical research area for understanding urban environments (Gamba and Houshmand, 2002). Finally, resources can be focused on accessing the quality of model in locations where: 1) building structures are complex, 2) buildings are constructed with reflective materials and, 3) vegetation growth near the buildings.

5. Conclusion

LiDAR data receive wide recognition in terrain modeling because of fine resolution and high elevation accuracy. However, in an urban environment where buildings vary greatly in height, shape and building material, uncertainty of LiDAR

data can reach ten meters or more. Moreover, vegetation near or grow over buildings reduce the accuracy of feature extraction and manual digitizing.

Pre-flight planning in data acquisition and data fusion techniques always help to reduce spatial uncertainty in 3D building models. However, assessment is necessary to justify whether the additional flight route or data collection is worthy. Improvements to the feature extraction algorithm also help to reduce the spatial uncertainty in a 3D building model only if user has the skills and access to modify the algorithm. When time and labor are limited, LiDAR data users can focus their assessments on buildings with glassy surface materials, gaps smaller than ten meters, and with vegetation near-by.

Since the study site, CBD of OKC, is situated in a very flat area, the methods and findings here can only attribute spatial uncertainty to building characteristics and surrounding environments. Further research could examine the spatial uncertainty of LiDAR and LiDAR-derived data in cities with complex terrain such as San Francisco and Hong Kong.

Acknowledgement:

This research was part of the US DOE/DTRA JOINT URBAN 2003 Field Experiment in Oklahoma City. Follow-up research was supported by Oklahoma NASA Space Grant. May Yuan's effort is also supported by the National Science

Foundation (SBE: 0416208). The contents of the research are solely the responsibility of the authors and do not necessarily represent the official view of DOE, Oklahoma NASA Space Grant, or NSF. Dr. Petra Klein and Dr. Jeff Basara provided helpful comments on the research development. The authors would also like to thank Melissa Brown at Center for Spatial Analysis for reviewing the article and the anonymous reviewers for their comments.

References

Ahokas, E., H. Kaartinen, and J. Hyyppa. 2003. A quality assessment of airborne laser scanner data. Paper read at ISPRS Working Group 3 Workshop, 3-D reconstruction from airborne laserscanner and InSAR data, 8-10 October 2003, at Dersden, Germany.

Alharthy, A., J. Bethel, and E. M. Mikhail. 2004. Analysis and accuracy assessment of airborne laserscanning system. Paper read at XXth ISPRS Congress, Commission II, 12-23 July 2004, at Istanbul, Turkey.

Flood, M. 2001. Laser altimetry: From Science to commercial LiDAR mapping. *Photogrammetric Engineering & Remote Sensing*:1209-1217.

———. 2002. Product definitions and guidelines in specifying LiDAR deliverables. *Photogrammetric Engineering & Remote Sensing* 68 (12):1230-1234.

Gamba, P., and B. Houshmand. 2002. Joint analysis of SAR, LIDAR and aerial imagery for simultaneous extraction of land cover, DTM and 3D shape of buildings. *International Journal of Remote Sensing* 23 (20):4439-4450.

Goodchild, M. F., B. Buttenfield, and J. Wood. 1994. Introduction to visualizing data quality. In *Visualization in Geographical Information Systems*, eds. H. M. Hernshaw and D. J. Unwin, 141-149. New York: John Wiley and Sons.

Hopkinson, C., L. E. Chasmer, G. Zsigovics, I. F. Creed, M. Sitar, P. Treitz, and V. M. Rober. 2001. Errors in LiDAR ground elevation and wetland vegetation height estimates. Paper read at International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, at Bangkok, Thailand.

Huising, E. J., and L. M. G. Pereira. 1998. Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry & Remote Sensing* 53:245-261.

Latypov, D. 2002. Estimating relative lidar accuracy information from overlapping flight lines. *ISPRS Journal of Photogrammetry & Remote Sensing* 56:236-245.

Mass, H.-G. 2002. Methods for measuring height and planimetry discrepancies in airborne laserscanner data. *Photogrammetric Engineering & Remote Sensing* 68 (9):933-940.

SAIC. 2005. *RTV LiDAR Toolkit* 2005 [cited 4 April 2005]. Available from http://www.tec.army.mil/td/tvd/survey/RTV_LIDAR_Toolkit.html.

Schenk, T., B. Csatho, and D. C. Lee. 1999. Quality control issues of airborne laser ranging data and accuracy study in an urban area. Paper read at ISPRS Workshop: Mapping Surface Structure and Topography by Airborne and Spaceborne Lasers, 9-11 Nov 1999, at La Jolla, USA.

Smith, S. L., D. A. Holland, and P. A. Longley. 2003. Investigating the spatial structure of error in digital surface models derived from laser scanning data. Paper read at ISPRS Working Group 3 Workshop, 8-10 October 2003, at Dersden, Germany.

———. 2004. The importance of understanding error in lidar digital elevation models. Paper read at XXth ISPRS Congress, 12-23 July 2004, at Istanbul, Turkey.

Chapter 3 : Spatial Uncertainty of 3D Data from Coupling Geographic Information Systems and Urban Atmospheric Dispersion Model: An Example using ArcGIS and QUIC

Abstract

Understanding spatial uncertainty is fundamental to Geographical Information System (GIS) applications and environmental modeling. In addition to uncertainty inherited in the input data, additional spatial uncertainty is introduced through necessary data conversion and manipulation to couple a GIS and an environmental model. Spatial uncertainty may vary depending on applications. This research investigates the common sources and patterns of spatial uncertainty involved in coupling GIS and EMS, particular with a focus on 3D building data. While several researchers have attempted to address spatial uncertainty for 2D data, there is no systematic research to examine spatial uncertainty associated with conversions of 3D GIS data for environmental modeling. ArcGIS (Environmental Systems and Research Institute Inc., Redlands, California) and the QUIC atmospheric dispersion model (Los Alamos National Laboratory) are chosen as an example to examine the extent and ramification of spatial uncertainty from GIS-EMS coupling. Typical of most environmental models, QUIC imposes specific requirements on spatial data, and therefore, conversion of GIS data to meet the requirements is inevitable. Like many data

conversion procedures, spatial data resampling is needed to convert ArcGIS data for QUIC. Consequently, spatial resolution (or cell size) plays a key role in the introduction of spatial uncertainty in GIS-EMS coupling. Furthermore, footprints and the height dimension of building data required by the QUIC model contribute additional complexity to spatial uncertainty. To account for three-dimensional data and spatial resolution, this study analyzes spatial uncertainty in building footprint, building location, and building volume across twelve spatial resampling resolutions during data conversion to meet QUIC data input requirements. Results show a linear relationship between the mean shifting distance of building location and spatial resampling resolution. As the spatial resolution increases from one meter to twelve meter, measurements of building footprint and volume vary from one to three percent, while eighty percent of footprint area and building volume remains unchanged. Elongated buildings or linear urban structures, such as skywalks, may be missing after the conversion. The study first presents a novel algorithm to convert 3D GIS building data for atmospheric dispersion modeling and then systematically examines spatial uncertainty introduced during data conversion. While the study is only an example to explore the manifestation of spatial uncertainty of 3D data in coupling GIS and environmental modeling, the commonality of resampling in data conversion and the complexity of resampling

3D building data warrant the research findings relevant to many GIS-EMS coupling applications.

1. Introduction

Many environmental modeling systems (EMS) take geo-spatial data as a fundamental part of input data. Although geospatial data are easily accessible, these EMSs are likely to have unique data requirements that are not directly compatiable with existing geospatial data. Since EMSs have limited capability in handling geospatial data, the burden is mostly carried out through GIS transformation of geospatial data. To overcome the frequent needs for intensive labor and time on data transformation, linkages of customized programs can facilitate automation of data exchange between the EMSs and Geographic Information Systems (GIS).

Two general approaches of linkages can be identified: 1) coupling and, 2) integration (Bernard and Kruger, 2000; Goodchild, 1996; Fedra, 1996). Coupling means linking two systems through data conversion while integration means linking two systems based on the same data model and functionality. Each approach carries its own advantages and limitations, which have been well discussed in the literature (Martin et al., 2005; Fedra, 1996). Although the integration approach is ideal to fully facilitate GIS and EMS interactions and minimize data loss between the two systems, this approach is not always practical because of high level of development cost and skills. Instead, the coupling

approach is more commonly used than integration approach because coupling only requires development in the data conversion algorithm, without the needs to disassemble and reassemble one system into the other. However, geospatial data conversion inevitably results in spatial uncertainty, and therefore, proper interpretation of modeling output depends upon a good understanding of how spatial uncertainty may be introduced through data conversion process.

This paper follows the definition of spatial uncertainty as "discrepancy between geographic data in GIS and the geographic reality these data are intended to represent" in Geographic Information Science (Zhu, 2005). Distinguished from the definition of "error", the term "uncertainty" is used when true values of the discrepancy are not available (Goodchild, Buttenfield, and Wood, 1994). Longley et.al. (Longley et al., 2001) explain that spatial uncertainty appears as soon as we conceptualize the reality, and it continues to propagate through the data life cycle including data capture, storage, spatial analyses and modeling. In addition, the attributes and geospatial representation of data may be altered through different stages of the data life cycle.

Understanding spatial uncertainty is fundamental to GIS applications and environmental modeling. This study asserts that a comprehensive assessment of spatial uncertainty during GIS-EMS linkage is necessary. In urban atmospheric

dispersion modeling (UADM), for example, uncertainty in geographic data can be a critical issue. A small change in building location may lead to a significant difference in how pollutants disperse within an urban area. Moreover, understanding the spatial uncertainty in geospatial data provides crucial information to further assess the behaviors of UADMs. Subsequent GIS analyses, such as overlaying dispersion estimates with demographic data, are often used for emergency planning. Knowledge about spatial uncertainty from GIS-UADM coupling aids interpretation of the dispersion outcome and vulnerable population to hazardous chemical exposures. Therefore, this paper aims to first couple a GIS and an UADM, using ESRI ArcGIS and the Quick Urban and Industrial Complex (QUIC) atmospheric dispersion modeling system as an example and then assess the spatial uncertainty that arises during the coupling.

2. Background

2.1 Overviews of ArcGIS and QUIC

This study uses ArcGIS and QUIC dispersion model as an example to demonstrate the spatial uncertainty involved during linkage of GIS and environmental modeling systems. First, the data models and formats used by ArcGIS and QUIC are described, and a linkage approach is identified. After that,

a literature review provides insights in assessing the spatial uncertainty arising from the linkage.

ArcGIS is a popular commercial GIS software developed by the Environmental Systems and Research Institute (ESRI, Redlands California) to store, manipulate and display spatial data from different sources and at different resolutions. Primary components of the software consist of ArcMap, ArcCatalog and ArcToolbox. The main functions of ArcMap are to display and query spatial data. ArcCatolog manages spatial databases; whereas ArcToolbox contains data processing and spatial analytical tools.

The QUIC dispersion modeling system is a new dispersion model developed by Los Alamos National Laboratory (Pardyjak and Brown, 2002; Williams, Brown, and Pardyjak, 2002) in MATLAB software with a diagnostic wind field model (QUIC-URB) and a Lagrangian dispersion model (QUIC-PLUME) to provide fast predictions of atmospheric dispersion in urban areas. QUIC windfield and dispersion models take detailed building information to estimate chemical dispersion patterns.

Building data are commonly available in GIS formats. While buildings can be represented in raster or vector data models, the discrete nature of building footprints makes vector representation more appropriate to store building data than

raster representation. Even when LiDAR technology is used to capture building footprints and heights, LiDAR cells represented buildings are delineated to form building polygons and determine building heights. Similarly, ArcMap stores buildings as vector polygons, which are composed of lines and points with x, y co-ordinates. Each polygon is associated with an attribute table which stores other information about the building such as height, base height, name and roof features. In QUIC, buildings are stored as either rectangular or circular shapes. Information of each shape includes building ID, group numbers, types (i.e. either rectangle or circle), width, length, height, xfo (the middle x co-ordinate of the length), yfo (the middle y co-ordinate of the width) and zfo (base height) which are stored as a line of text in a file (i.e. QU_buildings.inp).

The fact that both ArcGIS and QUIC represent building data in vector form suggests that data format will not be the primary source of spatial uncertainty during the coupling of the two systems. Rather, spatial uncertainty is likely to due to the change of shape and location of buildings during the data conversion. Buildings are conceptualized as polygons with discrete boundaries. The main differences are that building data in QUIC is limited by resolution defined by users, and shapes (i.e. either rectangle or circle), while in ArcGIS, building data are not limited by resolutions and shapes.

A linkage also enables exporting dispersion results from QUIC to ArcGIS. Similar to most of the UADMs, both ArcGIS and QUIC utilize raster model to represent the concentration of pollutants in atmosphere because the spatially continuous nature of pollutant dispersion. Although QUIC includes a function that exports dispersion results as ArcGIS raster format, it only exports one layer at a time. An automatic export function that exports all outputs at once is desirable.

2.2 Approach for linking ArcGIS and QUIC

Which linkage approach should be adapted to bridge ArcGIS and QUIC? As mentioned in section 1, integration and coupling are the two main approaches. Although integration is the ideal linkage for GIS and EMS, it may not be practical to rewrite GIS or EMS to integrate one into the other. For instance, most of the GIS and EMS are developed as individual software; integration of them will be similar to re-design a brand new software package. The cost and the risk are too high for most of the users. Therefore, coupling approaches are comparably effective to bridge the two systems while maintain the advantages of each system.

Coupling approach can be further divided into tight-coupling and loose-coupling (Nyerges, 1993). Tight-coupling transfers data between GIS and EMS using the same user-interface, while loose-coupling uses different

user-interfaces. A user-interface is a computer program that interacts between user and computer systems. It can be developed by any computer languages such as Fortune, C++ or Visual Basic. Although coupling approach is very common in application due to low development cost, spatial uncertainty can be introduced and embedded during the data conversion.

2.3 Spatial uncertainty from the linkage

Although different UADM may have different formats for inputting spatial data, they face the similar problem during the linkage. For example, QUIC dispersion modeling system represents buildings as rectangles by default. This default assumption is common to many other UADMs (Vardoulakis et al., 2003). However, buildings can be in many different shapes, and an ArcGIS database represents building as such. When importing building data from ArcGIS to QUIC, procedures are necessary to transfer irregular polygons in GIS to orthogonal shapes of rectangles used in QUIC. This study utilizes rasterization as one of the procedures to convert irregularly shaped buildings in ArcGIS into regular rectangles in QUIC. Consequently, rasterization contributes to the main source of spatial uncertainty during the linkage.

The relationship between rasterization and data accuracy had been well-studied in the GIS literature. When vector polygons are converted to raster cells, there are a number of factors influencing data accuracy. The cell size (i.e. resolution) is always the main issue regarding rasterization. Piwowar (1990) found that accuracy of data decreased as cell size increased during data conversion between vector and raster models. In addition, he recommended using one-fourth of the minimum polygon area as the cell size in order to maintain the integrity of the data. If the cell size was larger than twice of polygon area, the polygon would be absent after the conversion.

Besides the cell size, the position of the grid also induces uncertainty during rasterization. Wedhe (1982) examined the relationship between cell size and map error. He found out that the grid position could determine the presence of a polygon if the cell size was the same as the area of the polygon. In general, he concluded that the grid position was not an important consideration for the overall map accuracy but it was a significant factor for accuracy of individual polygons.

Congalton (1997) explored the errors introduced during rasterization and vectorization in relation to various shapes of simplistic polygons. By computing the area change before and after conversion, he found that as cell size increased,

larger differences in shape and area of original polygons were expected. Moreover, some polygons might disappear after conversion if the area was the same or less than the cell size. This issue is critical when there are linear features with width equal to the cell size. In conclusion, he suggested further study with real world examples.

Bregt et al. (1991) estimated the errors from rasterization using an approach termed the double-conversion method. They converted polygons into raster twice using a desired resolution and a much finer resolution. By comparing the two raster maps, they found a linear relationship between rasterized errors and the boundary index computed by dividing the total boundary length of polygons in centimeters by the total area of the map in square centimeters. However, the linear relationship varied with cell size and rasterization methods, and the relationship only applied to certain resolutions.

While these studies summarize errors or uncertainty introduced by rasterization of polygons at different resolution, these studies are limited to issues on a two-dimensional surface. However, spatial uncertainty arising from coupling between GIS and urban atmospheric dispersion models extends to the third dimension (i.e. building height). This study contributes to the understanding of rasterization and spatial uncertainty in both horizontal and vertical dimensions.

The following sections examine the spatial uncertainty of rasterization during coupling related to: 1) the displacement of building location, 2) the change of building footprint area, and 3) the change of building volume at various resolutions.

3. Method

3.1 Study area and data

The central business district of downtown Oklahoma City is selected as the study area to explore spatial uncertainty due to the coupling of a GIS and an urban dispersion model. Downtown Oklahoma City is located in the central portion of Oklahoma. With an area of about one square mile, downtown Oklahoma City is a typical urban business district with high-rise buildings in the center surrounded by low-rise business buildings and open areas. Its well-defined central business district and flat terrain make it suitable for urban atmospheric dispersion studies and easy for assessing the accuracy of geographic data.

The building model used in this paper was derived from LiDAR data which were collected by the Joint Precision Strike Demonstration (JPSD) Program Office of the U.S. Army in late October 2001. The LiDAR data had one meter resolution with the overall accuracy of the data up to 0.3 meters in vertical

measurement and 0.5 meters in horizontal measurement (i.e. with 90 percent confidence level). The building model was extracted with the aid of the RTV LiDAR toolkit, an extension in ArcView 3.x developed by the JPSD program specifically for the LiDAR acquisition technology. The building model was quality assured with references to aerial photography and ground observations. Corrections were made manually to achieve the highest level of details possible with these auxiliary data.

The final building model is stored as a shape-file with heights in ESRI ArcMap (Figure 3.1). The study area contains 390 buildings and seven skywalks that together are represented by 1465 polygons. The building heights range from three meters to 150 meters among which 90 percent of the buildings are below 53 meters. The area of building footprints ranges from 20 m^2 to 4250 m^2, and volume of buildings ranges from 1,728 m^3 to 27,404 million m^3. Most polygons are rectangular in shape.

3.2 The concept of coupling algorithm

This study developes a coupling algorithm of ArcGIS and QUIC and implements it using Visual Basic for Application (VBA) inside ArcMap. VBA is

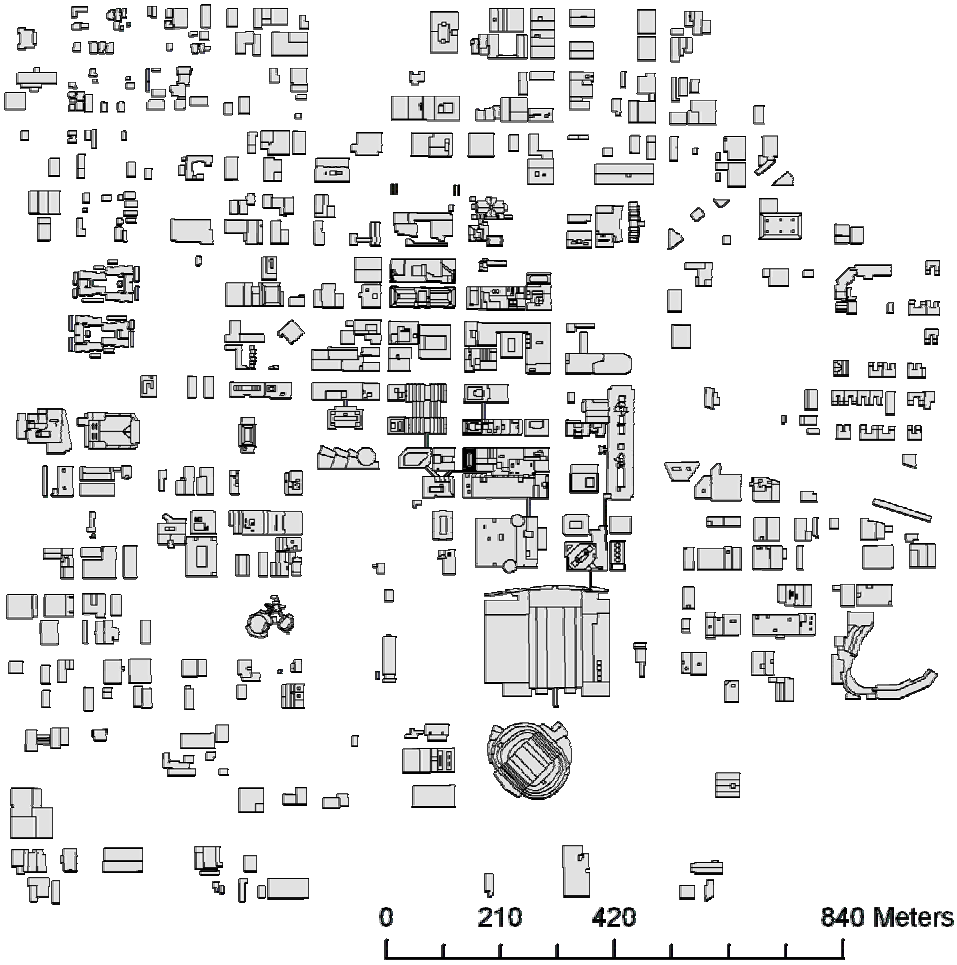
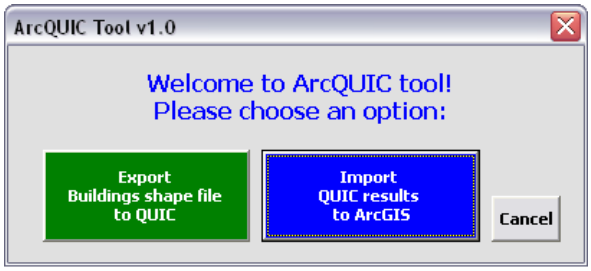


Figure 3.1. The building model developed for this research. The buildings are quality-assured through field surveys and air-photography corrections.

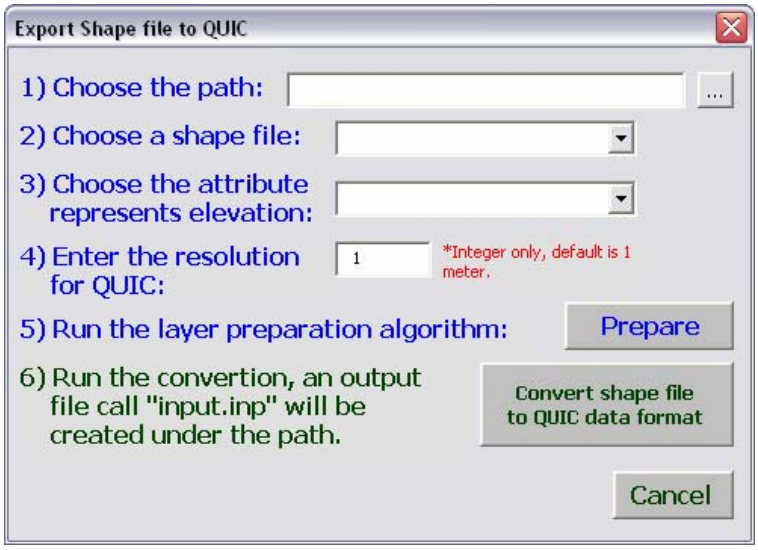
chosen because of its capability to access the spatial data information and spatial analytical functions in ArcMap. The algorithm contains two main functions: one to convert building data from ArcMap shape-files to data format required by QUIC and the other to convert dispersion results from the QUIC format to a grid data format in ArcMap. Since QUIC and ArcMap employ the same underlined data organization (i.e. raster model) for the dispersion results, no data model

transformation is needed from QUIC to ArcMap. Figure 3.2 shows the user interfaces of the conversion algorithm.

a)



b)



c)

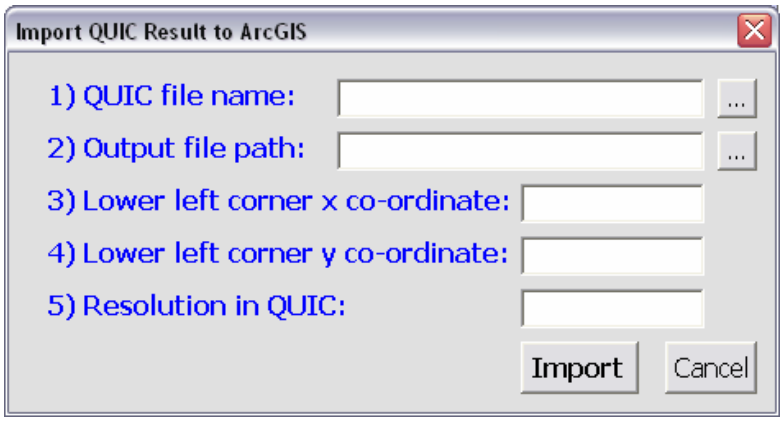


Figure 3.2. The user interfaces of the coupling algorithm: a) Beginning interface, b) interface of exporting shape file to QUIC, and c) interface of importing QUIC results to ArcGIS.

Many buildings exhibit different geometries or spatial configurations over height. For example, some buildings may have extended structures on the ground level or have skywalks at multiples levels. This algorithm reshapes each building horizontally and then vertically to derive a simple rectangular block as required by the QUIC model. The user first determines the vertical and horizontal resolutions for the data conversion. These resolutions are used to create a grid to which a building configuration is modified to the closest grid lines to approximate the building shape with orthogonal angles. Then, the algorithm slices a building horizontally according to the predefined intervals of elevation (i.e. vertical resolution) to derive the spatial configuration of the building at an elevation (i.e. a building slice). At each elevation interval, if the building is not a rectangle, the algorithm will vertically slices the polygon into several rectangles. Finally, the algorithm calculates the dimension (i.e. width, length and the height) of each rectangle and records the dimension data into a QUIC building file. Figure 3.3 presents a simplified example that illustrates the concept of the conversion algorithm to reshape a building slice.

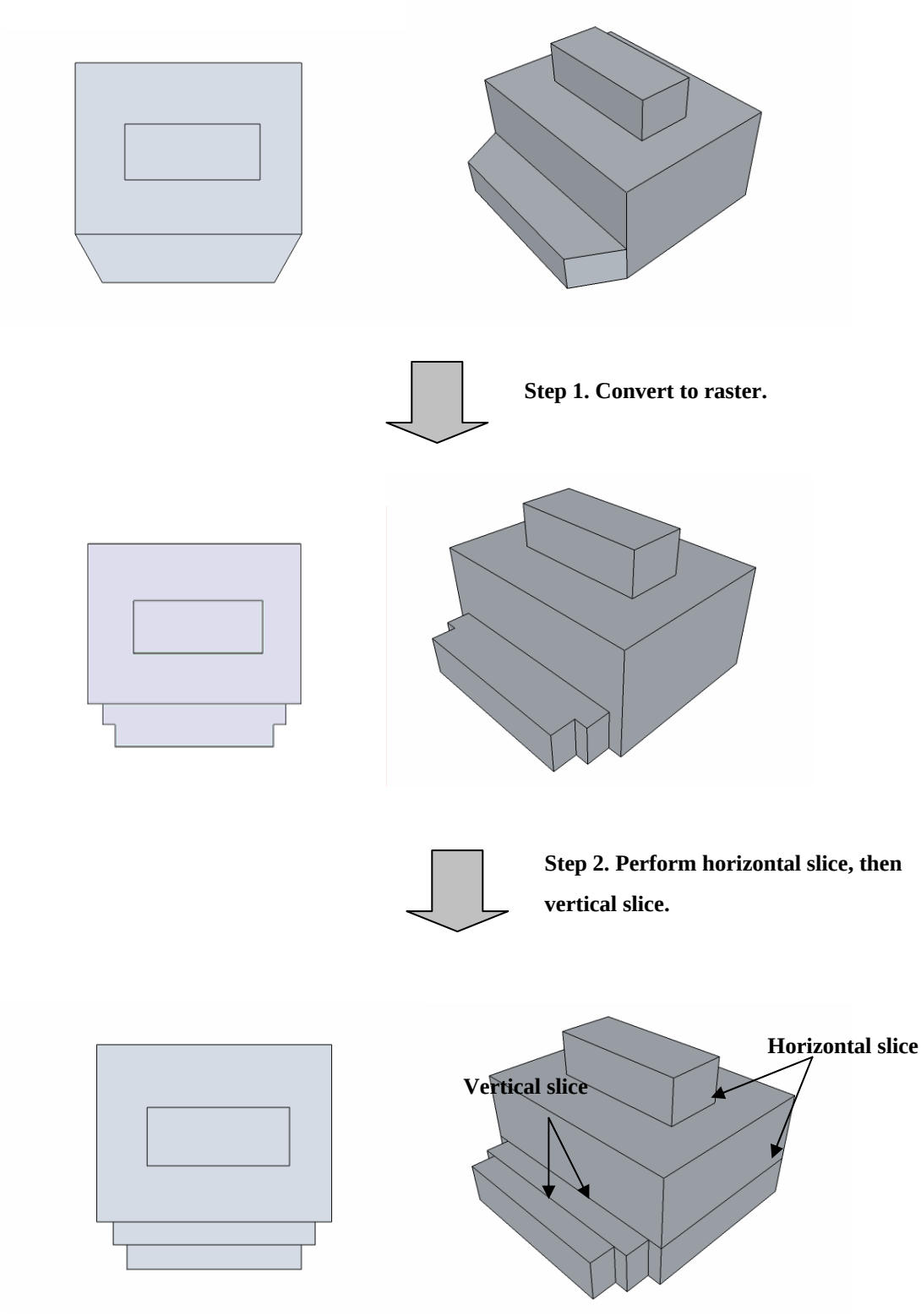


Figure 3.3. A simplified example to demonstrate the concept of data conversion in coupling algorithm. The building polygons are first converted to raster, then sliced into blocks horizontally and vertically.

3.3 Methods of spatial uncertainty assessment

The algorithm was applied to all building in the study area. As buildings were rasterized and approximated into rectangles, the algorithm inevitably changed building location, footprint area and volume, and thus introduced spatial uncertainty to the input data of the QUIC model. First, in order to quantify the building displacement (i.e. how far the location of building shifted after conversion), the Euclidean distance between the building vertices before and after the conversion at various resolutions were measured (Figure 3.4) and then the statistics of the shifting distances, including the mean and standard deviation were calculated. In addition, this study examined the locations where the shifting distances of building vertices were more than twice of the defined resolution. Finally, the angle of displacement was calculated and classified into eight directions.

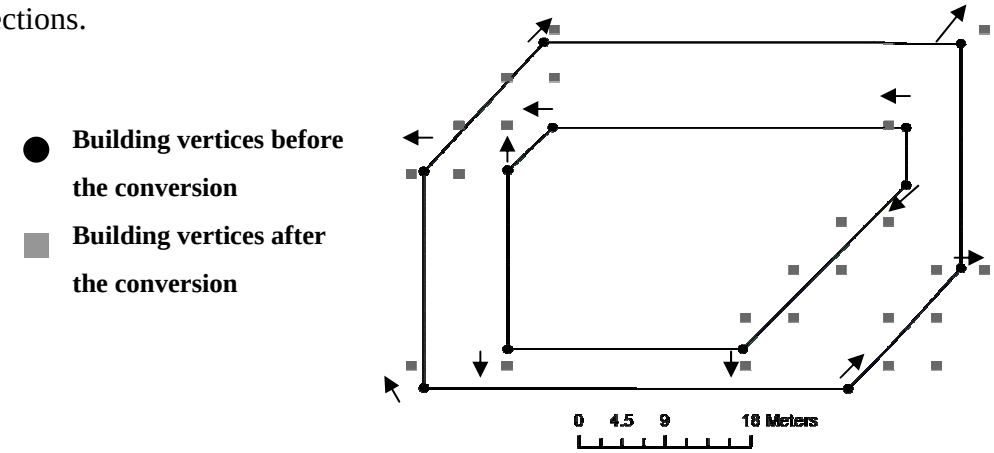


Figure 3.4. The shifting distance that is measured by calculating the distance between the vertices before and after the conversion. Arrows showing the shifting directions.

Second, this study explored the change to building footprints area by computing the difference between building footprint area before and after the conversion. The footprints before conversion with the footprints after conversion were compared at various resolutions by GIS overlays. Then the percentages of footprint area that were unchanged, omitted and committed across various resolutions were calculated. The results were also compared with the shifting distance at various resolutions.

Third, this study explored the change in building volume by computing the percentages of building volume unchanged, omitted and committed after conversion. In order to determine the percentages of building volume unchanged, omitted and committed, the double-conversion method from Bregt et.al. (1991) was utilized. First, the elevation of the building model was converted into a reference raster layer with 0.5-meter resolution (one half of the vertical resolution of the LiDAR data). Then, the building elevation was converted into various raster layers using different horizontal resolutions. By comparing the reference layer to other layers with coarser resolutions, the volume unchanged, omitted and committed across various horizontal resolutions could be computed.

In addition, this study executed the above uncertainty assessments with two rasterization methods that were supported by ArcMap. One was central position

method and the other was dominant unit method. The central position method assigned values to the grids by taking the polygons that fell at the center of the grids. The dominant unit method assigned values to the grids by considering the polygons that shared the dominant unit of the grid.

4. Results and Discussions

4.1 Coupling algorithm

The coupling algorithm worked reasonably well in converting building data from ArcMap shape-file to QUIC format. Table 3.1 shows the conversion algorithm run-time and the number of records after the conversion. The algorithm is ran on a desktop computer with a 3.4 GHz Pentium four CPU, two GB RAM. With the range from thirteen minutes nineteen seconds to one minute forty-two seconds, the processing time decreased as the resolution became coarser. The number of records after the conversion also decreased dramatically from 9258 records at one meter resolution to 477 records at twelve meter resolution (Table 1). Compared to the original number of polygons for buildings (i.e. 1458 polygons), represented in ArcMap shape-file, the QUIC building model required a greater number of records to represent all buildings in the study area than ArcMap when the selected resolution is finer than seven meters.

Table 3.1. The run-time for conversion algorithm and number of records after conversion.

Resolution (meter)	Time for conversion (mm:ss)	Number of records after conversion (Original no of polygons: 1458)
1	13:19	9258
2	08:28	4580
3	05:29	3356
4	03:51	2370
5	03:23	1885
6	03:05	1463
7	03:50	1273
8	02:29	978
9	02:14	815
10	01:57	686
11	01:41	566
12	01:42	477

Even though the algorithm was reasonably well and easily to run via the graphic user interface (GUI) menus, the run time is still unpractical to convert a large number of buildings at a fine resolution. Using the building complex depicted in Figure 3.5 as an example, since it was not in rectangular shape, the algorithm would slice the building into many long narrow rectangles. With such a large number of records, users might not able to run the QUIC dispersion model due to memory shortage when displaying buildings or calculating the wind field at every grid cell. Such challenges were common to many small-scale environmental models that were computationally demanding and could easily be

over-burdened by large data volumes. Utilizing data compaction techniques in slicing the buildings might reduce the number of records. However, at a fine resolution, even though the number of records was reduced, a large number of grid cells for the study area were still necessary. For instance, a hundred m^3 study area required 1,000,000 grid cells at one meter resolution. If two meter resolution was used, the number of grid cells was reduced to 125,000 which were about 12.5 percent of grid cells used at one meter resolution. Also, decrease in spatial resolution, such as from one meter to two meter resolution, could greatly reduce the number of blocks in QUIC to nearly 50 percent (Table 3.1). Therefore, a

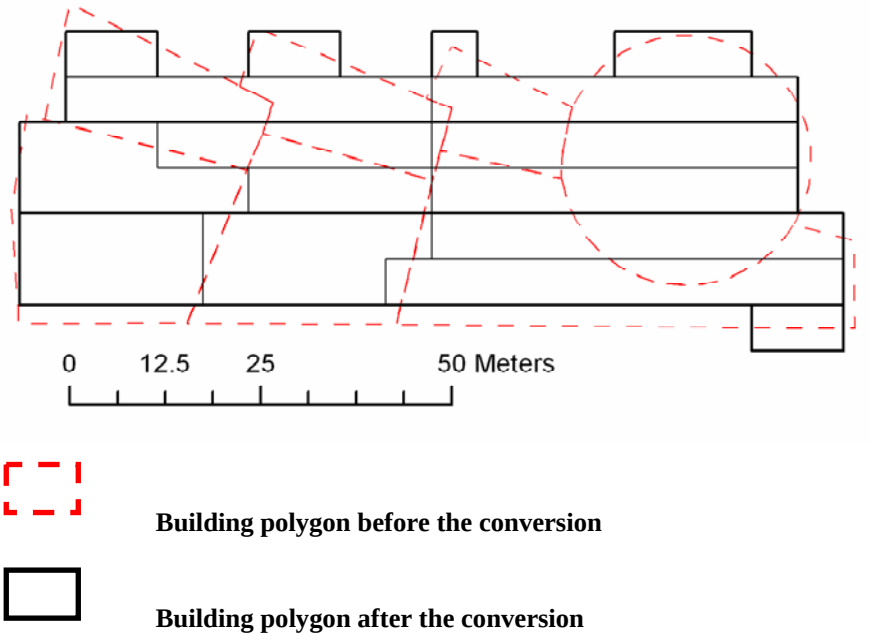


Figure 3.5. An example of building in downtown Oklahoma City using six meter horizontal resolution, which results more number of records after the data conversion.

simpler solution would be either to increase the resolution or to reduce the size of study area.

On the other hand, converting the results from QUIC to ArcMap was easier and more direct than from ArcMap to QUIC because no data conversion was involved. However, ArcMap was not fully capable of displaying a large number of temporal 3D data. Results from urban atmospheric dispersion models always included numerous time steps, each of which consists of chemical dispersion estimates in 3D space. Alternative solutions included displays of chemical dispersion estimates one time step at a time, displays of lows or pattern of changes in chemical concentrations, animation of time-stamped dispersion estimates, and other visualization approaches.

Returning to the research focus on exporting the building model from ArcMap to the QUIC dispersion model, the sensitivity of the building footprint location, area and volume are examined by changes in spatial resolution. The next section presents the results from spatial uncertainty assessment of building footprint location, area and volume.

4.2 Change in building location

The change in building location was represented by the mean distance between building vertices shifted after conversion. The shifting distances were transformed with natural log for normality. Figure 3.6 shows the mean shifting distance of all building vertices after conversion with both methods of rasterization. The mean shifting distance was around 50 percent of the defined resolutions and increased steadily as the spatial resolution became coarser. Figure 3.7 summarizes the shifting directions across various resolutions. The directional shifts of building vertices were similar in proportion, contributing about 12.5 percent for each direction across twelve resolutions. Compare two methods of rasterization, results were almost the same.

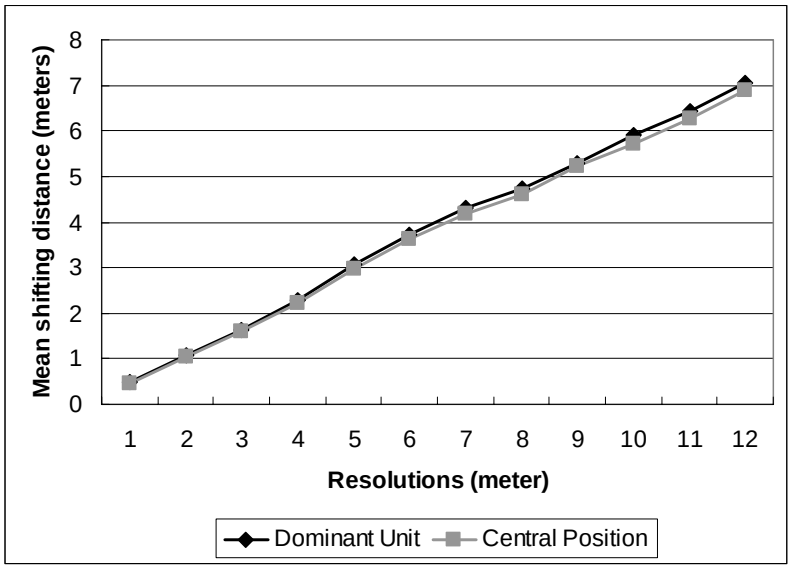


Figure 3.6. The relationship between the change in resolutions of the gird used in conversion and the average shifting distances. Comparing two types of rasterization.

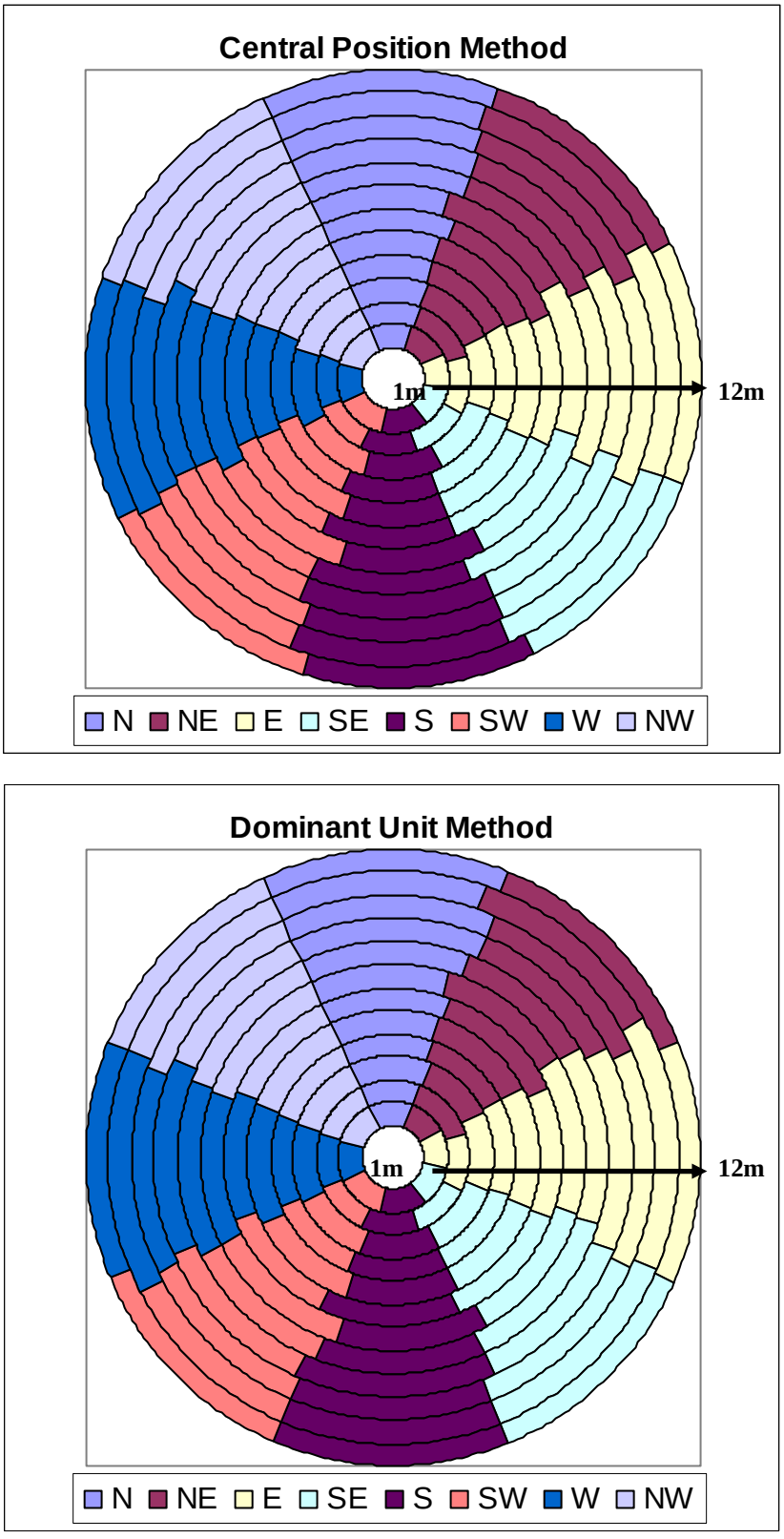


Figure 3.7. The proportion of shifting directions across twelve resolutions, comparing two types of rasterization.

Figure 3.8 depicts the spatial distribution of the building vertices that shifted over two standard deviations (i.e. 2.6 meters) after conversion with one meter horizontal resolution. These building vertices were found across the whole study area. Both methods of rasterization received the almost the same distribution except two vertices that were located at the southern part of the area. Across twelve resolutions, the numbers of buildings vertices with shifting distance exceeded two standard deviations were first increased at two meter resolutions and then gradually declined (Table 3.2). As resolution became coarser, although both methods of rasterization got fewer building vertices with abnormal shifting distance, some abnormal shifting were found at different locations (Figure 3.9). Compare both rasterization methods, extra abnormal shifting distances were found at some buildings with relative small footprint area when dominant unit method was employed. If the building was smaller than the user-defined resolution, it could be absented from the model after rasterization using dominant unit method and hence introduced abnormal shifting distance. However, if central position method was chosen, the building might still exist and the shifting distance would be reduced.

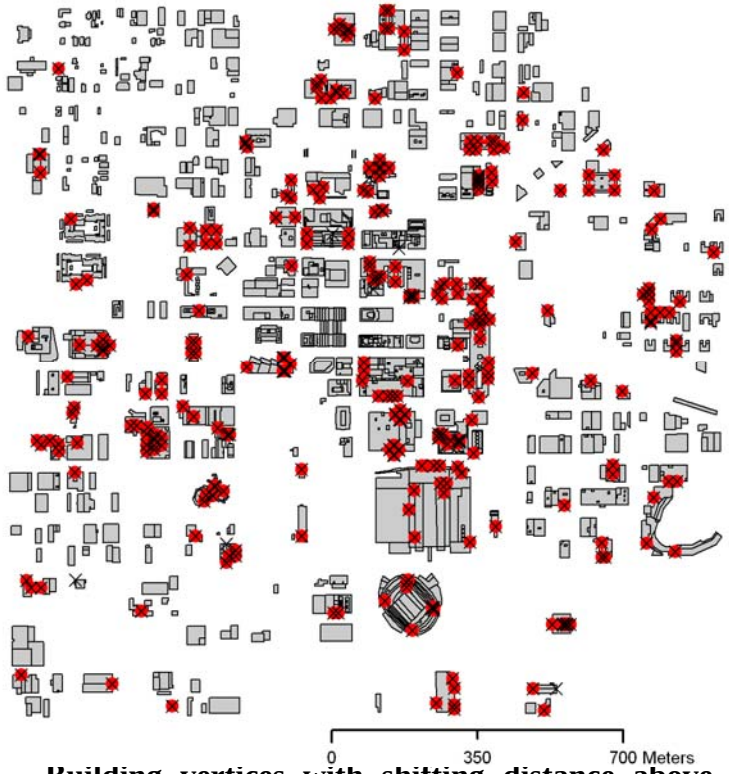


Figure 3.8. Building vertices with shifting distance above two standard deviations after conversion, using one meter user-defined spatial resolution. Red dots represent dominant unit rasterization while black crosses represent central position rasterization.

Table 3.2. The number of building vertices with shifting distances above two standard deviations across twelve resolutions.

Resolution (meters)	Maximum area method	Cell center method
1	339	347
2	462	469
3	404	408
4	335	339
5	207	232
6	171	189
7	129	161
8	163	192
9	135	133
10	142	133
11	133	129
12	120	112

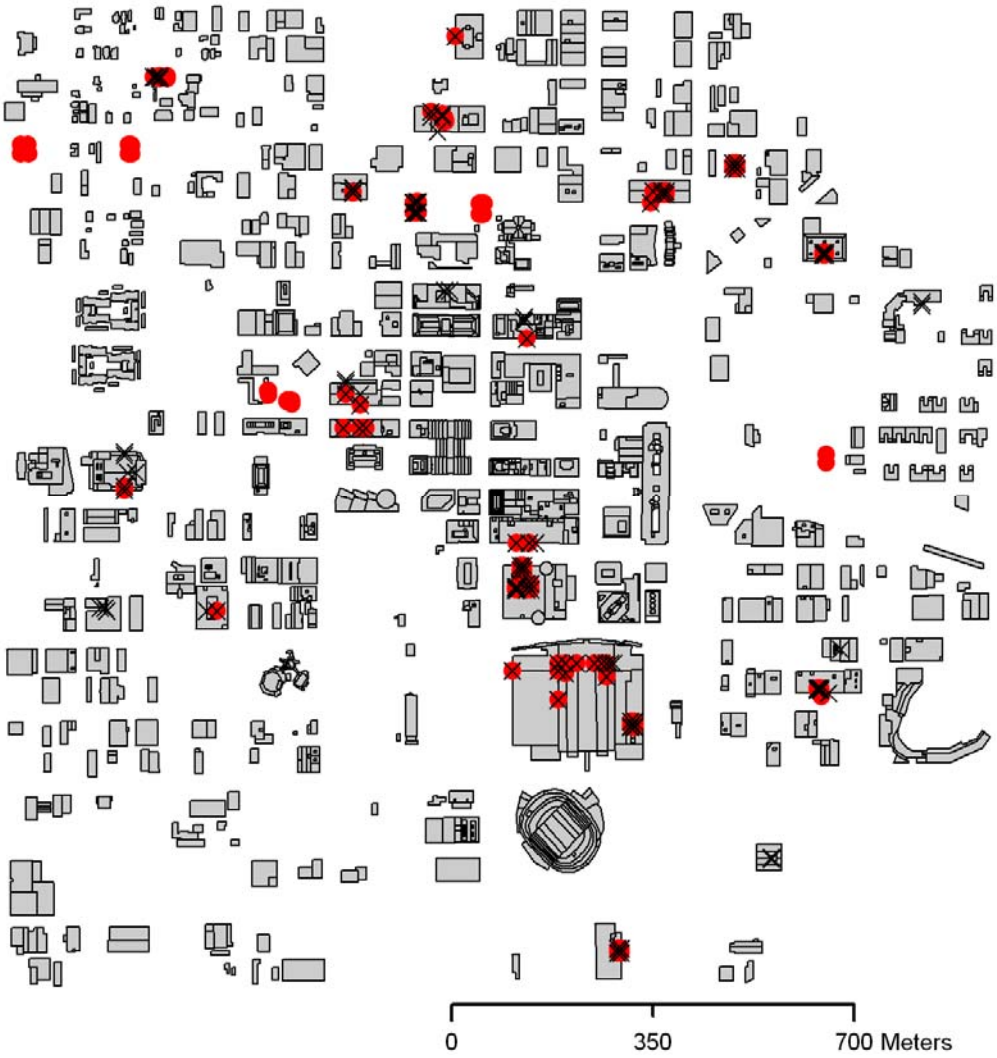


Figure 3.9. The location of building vertices with shifting distance above two standard deviations, at twelve meter resolution. Red dots represent results using maximum area method while black crosses represent results using cell center method.

Buildings with great shifting distance after conversion might greatly influence the results of atmospheric dispersion model. This study summarized building characteristics with great shifting distance by examining the relationship

between the shifting distance and building characteristics. First, compactness and edge index (Wentz, 2000) were calculated to characterize the buildings. Compactness also called circularity ratio which compared the area of a shape to the area of a circle. It was calculated by:

4π (area) / (perimeter)^2

, where compactness of one represented a circle and zero represented an infinitely long and narrow shape. Edge index characterized the roughness or smoothness of a shape by:

2* log (perimeter) / log (area)

, where larger edge index represented a shape with rougher edge.

Second, the sum of the shifting distance for each building polygons after conversion was calculated. Figure 3.10 shows the scatter plots of two indexes against the sum of shifting distance after conversion at one meter resolution using dominant unit method. No obvious relationship could be found between the sum of shifting distance and shape compactness. For edge index, buildings with greater edge index tended to have a smaller sum of shifting distance. However, the range of the shifting distance was large for buildings with small edge index. Similar results were observed from central position method. Although no significant relationship could be found between shifting distance and two shape

indexes, the buildings with abnormal shifting distances were visually examined and three building characteristics that accounted for the results were summarized.

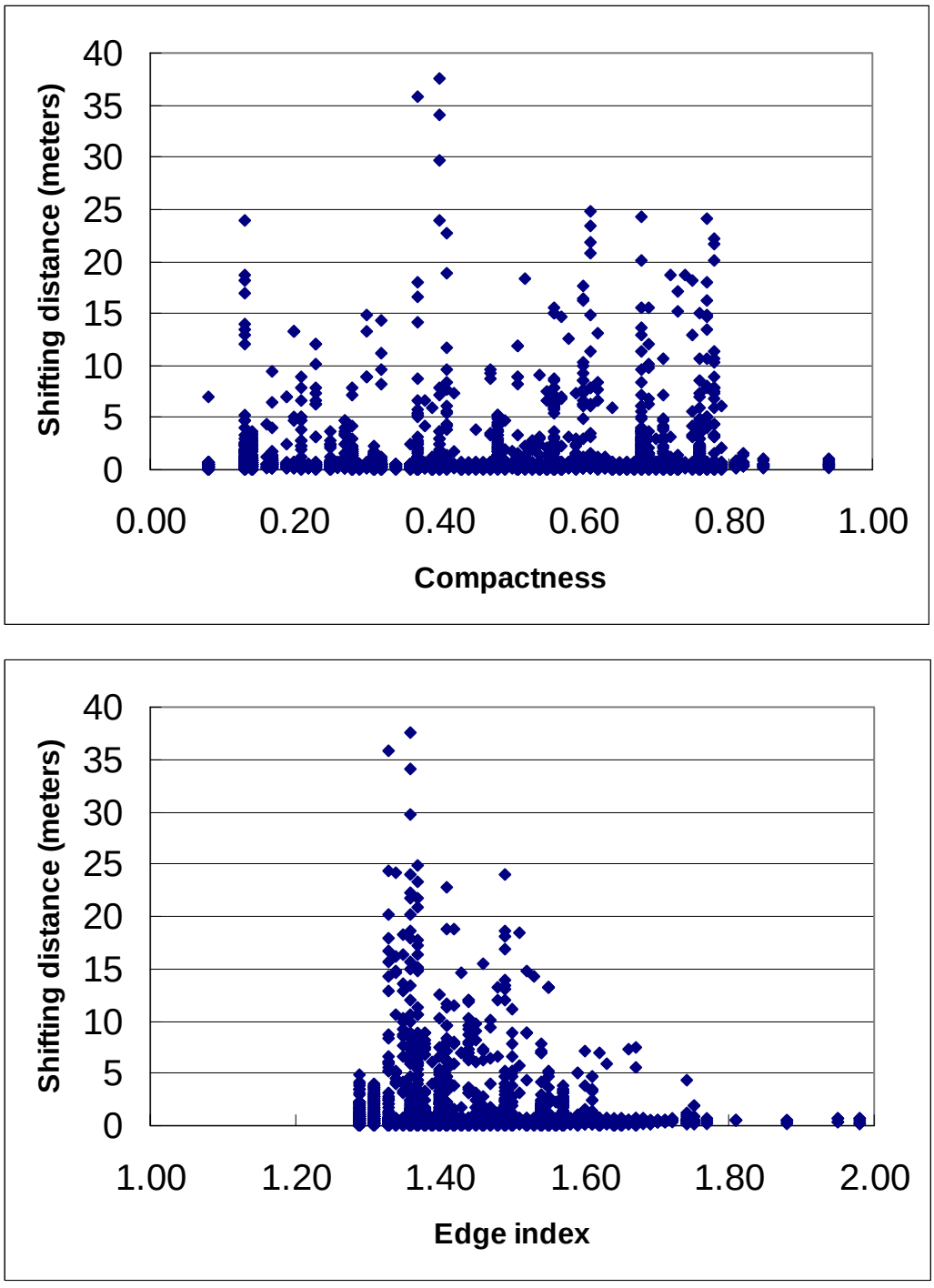
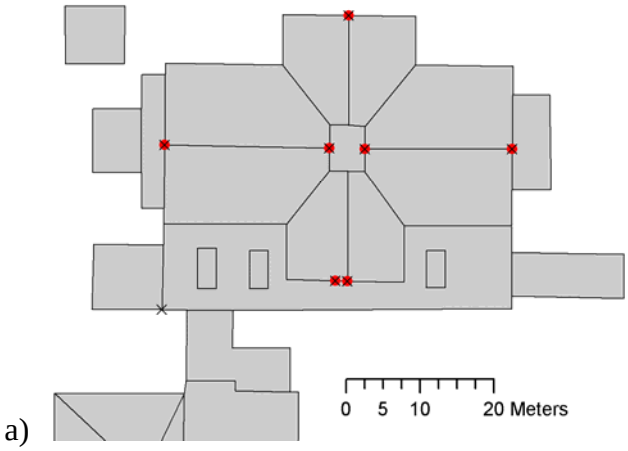
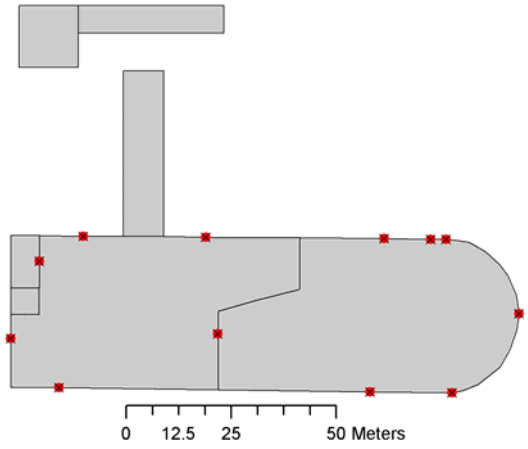


Figure 3.10. The scatter plot of shifting distance and two shape indexes at one meter resolution, using dominant unit method.

The first building characteristic is “sloping roof top”. As Figure 3.11a shows, that buildings with sloping roofs could be represented by triangular shape in ArcMap with heights recorded as minimum, maximum, and the slope of the roof. However, when converted to a rectangular shape in QUIC’s format, those vertices must be eliminated and resulted in significant shifting distances after data conversion.



a)



b)

Figure 3.11. Two examples of building vertices with shifting distances above two standard deviations at one meter user-defined resolution: a) building with sloppy root-top and, b) building with excessive vertices.

The second characteristic resulted in high shifting distances was the presence of excessive building vertices. For example, Figure 3.11b shows a building polygon with excessive building vertices. After conversion, those vertices disappeared and hence the shifting distance greatly increased.

The third characteristic accounted for the differences between dominant unit method and central position method. When the building was smaller than the user-defined resolution, it might still be recognized as a building by central position method but not by dominant unit method. A significant shift might occur if it absented from the model after rasterization. However, if there were buildings near this absented building, the shifting distance of the absented building would be reduced since the algorithm searched for the closest building vertices and calculated the shifting distance.

Although the above three characteristics of buildings could not be quantified with the shape index, two of them could be easily discovered using spatial editing and spatial query functions in GIS software except the sloppy roof top of buildings, which could be hard to detect if no information about the roof was obtained during data collection.

Finally, with a systematic analysis of the building displacements due to data conversion could be used to model the spatial uncertainty for building location

in urban dispersion modeling. By calculating the shifting distance in x and y co-ordinates from all buildings, a frequency distribution of x, y shifting for all resolutions could be computed (Figure 3.12). The frequency distribution provided the basis to simulate uncertainty of building location when we studied the influence of spatial uncertainty on the urban atmospheric dispersion model.

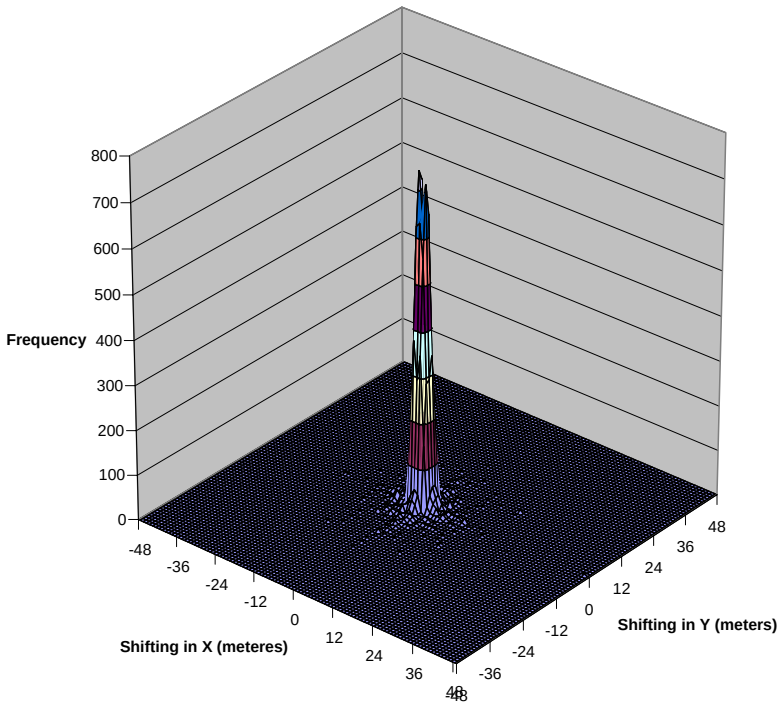


Figure 3.12. The frequency distribution of shifting distance in x and y co-ordinates at one meter resolution.

4.3 Change in building footprint area

Figure 3.13 illustrates the percentage changes in total building footprint area and the unchanged building footprint area across twelve resolutions, two methods of rasterization. As resolution became coarser, the total footprint area after conversion changed within +/- one percent. The building footprint area remained unchanged steadily decreased from 98.6 percent to 84.6 percent and 81.7 percent for central position method and dominant unit method respectively. The changes in the percentage of building footprint area omitted and committed were very similar for central position method, where they both increased from 1.3 percent to around fourteen percent and fifteen percent (Figure 3.14). However, for the dominant unit method, the percentages changes in building footprint area omitted and committed were increased from one percent to eighteen percent and eleven percent respectively. The changes in area omitted were larger while the changes in area committed were smaller at twelve meter resolution.

As resolution became coarser, the interior of building footprints remained unchanged and the edge of buildings became less complex. The building footprints omitted and committed were distributed around the edges of the original building footprints (Figure 3.15). Coarser resolution contributed a larger area

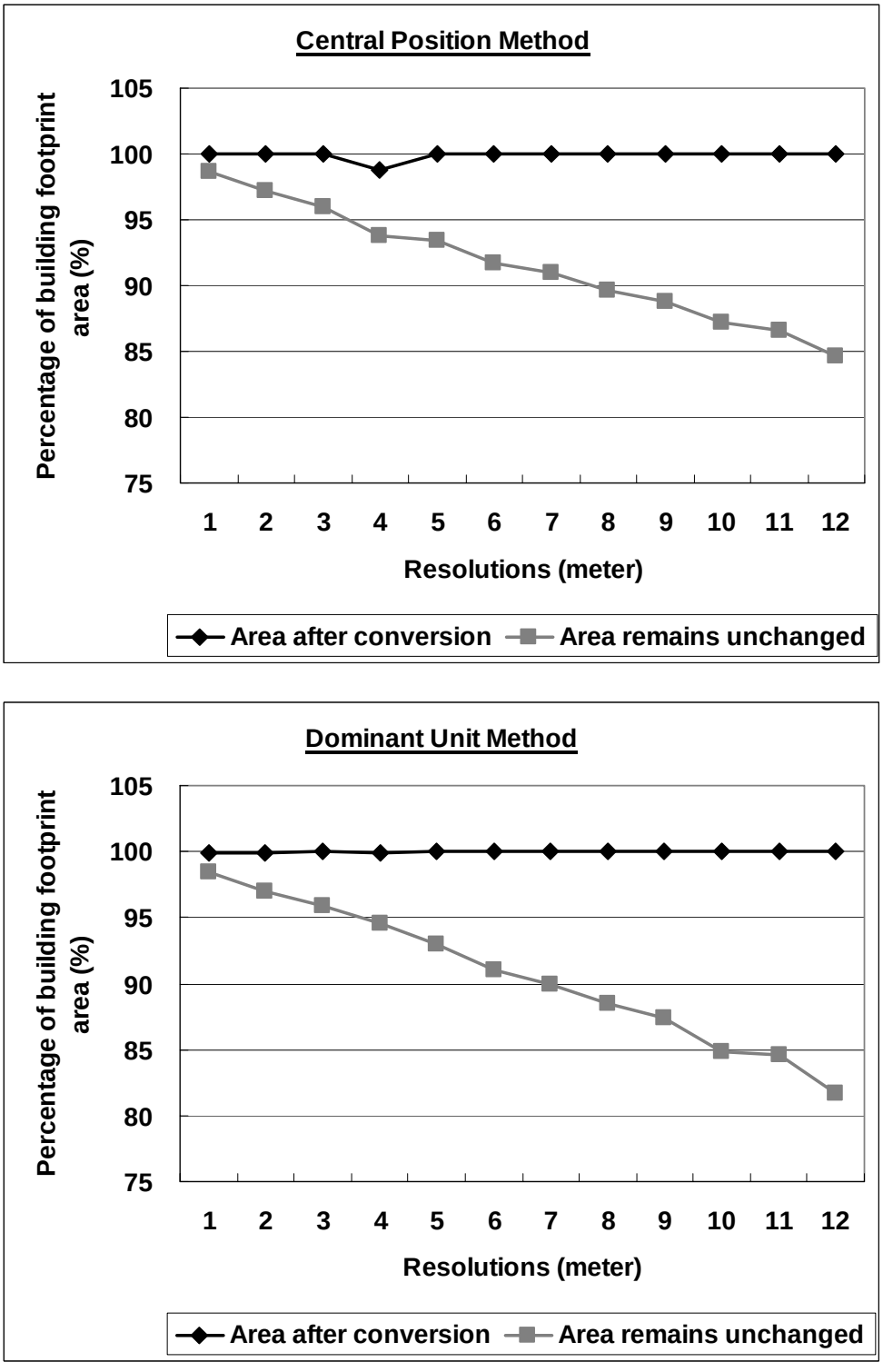


Figure 3.13. Percentages of building footprint area changed (black line) and unchanged (grey line) after conversion across twelve resolutions, comparing two methods of rasterization.

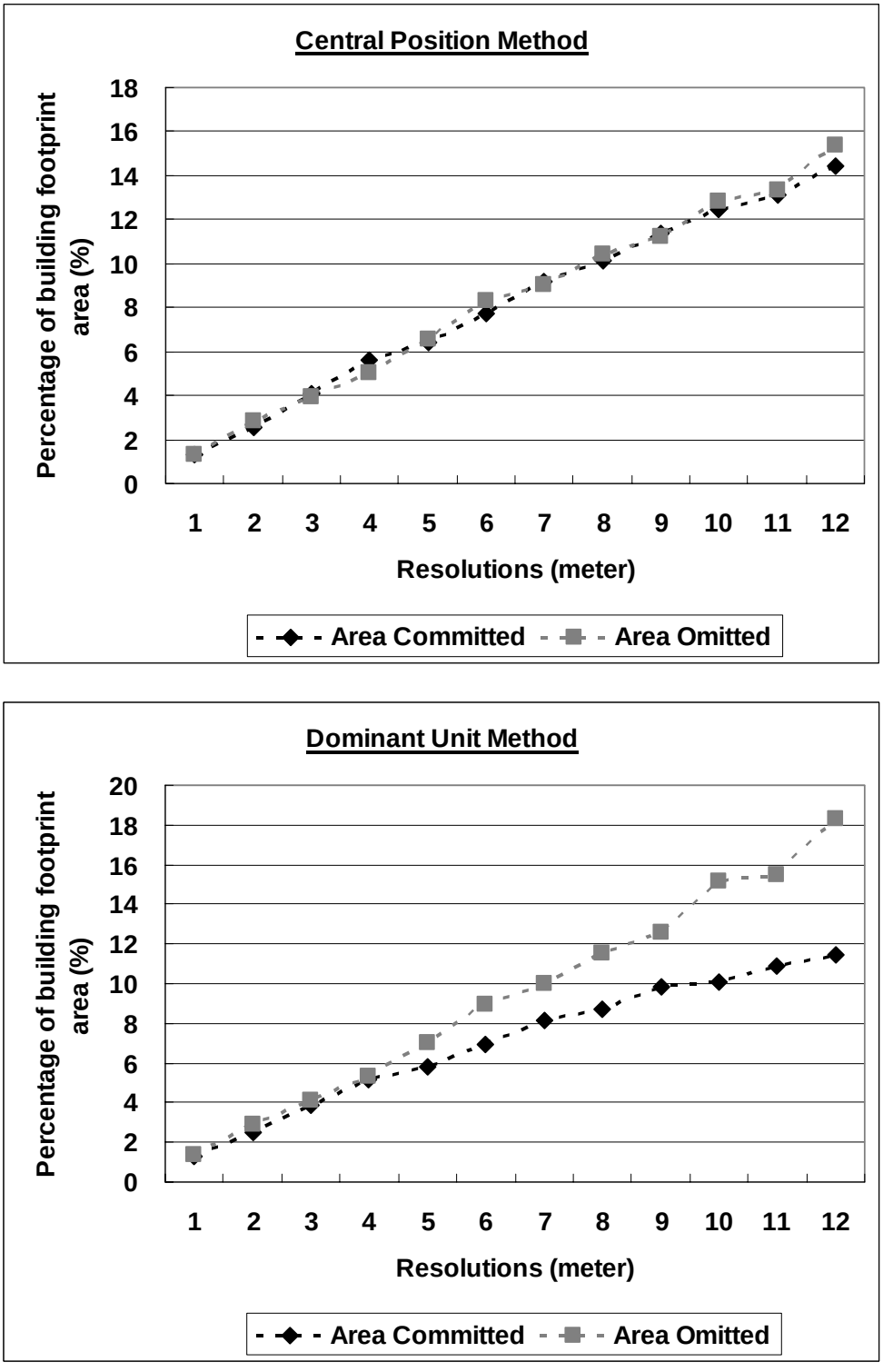


Figure 3.14. The percent of footprint area omitted (Grey dotted line) and committed (black dotted line) after conversion.

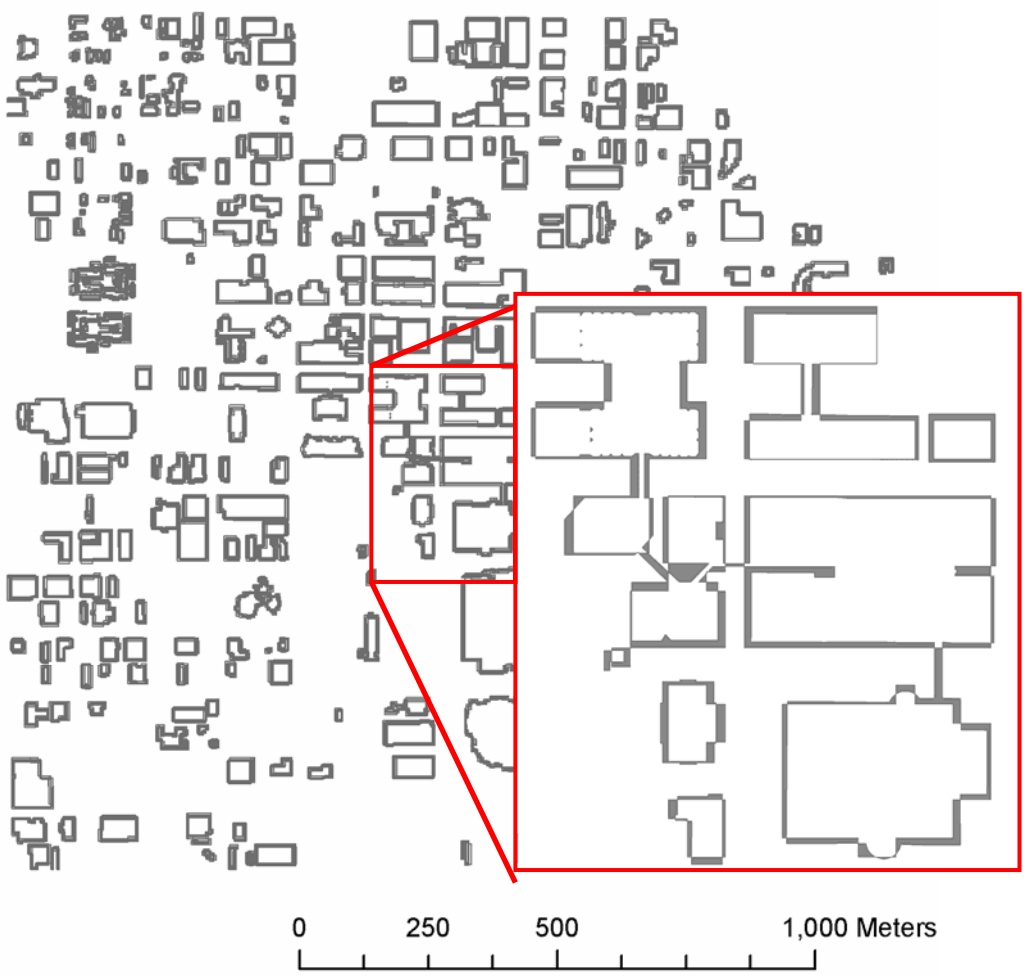


Figure 3.15. The spatial distribution of omitted and committed footprint area after conversion. Area in grey is a combination of omitted and committed footprint area at twelve meter resolution.

around the original building footprints. Conversely, linear features like skywalks or gaps between buildings might become disappear after conversion (Table 3.3).

When the central position method was chosen to rasterize polygons, there was no fixed pattern for the absence of skywalks after conversion. Unlike the results from dominant unit method, skywalks would not be presented again once it absented from previous resolution. Exceptional case was found at skywalk

Table 3.3. The resolutions that missed skywalks after conversion. Comparing two methods of rasterization.

Resolution (meter)	Skywalk ID (Central Position)							Skywalk ID (Dominant Unit)						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1														
2														
3							x							x
4				x		x	x				x		x	x
5		x		x		x			x		x		x	x
6		x		x		x			x		x		x	x
7				x		x	x		x		x		x	x
8	x	x	x			x		x	x	x	x		x	x
9	x	x	x		x	x	x	x	x	x	x	x	x	x
10	x			x			x	x	x	x	x		x	x
11	x		x	x			x	x	x	x	x	x	x	x
12			x	x	x	x	x	x	x	x	x	x	x	x

x – Absented of the skywalk

number five where the width of the skywalk was about five meters. Therefore, the presence of this skywalk at ten meter resolution would be depended on the spacing of the grid. Although Wedhe (1982) stated that the position of the cell was not important for overall map accuracy, it would become an important factor if the study area was composed of polygons of similar sizes and regular patterns, or polygons in linear shape. For example, buildings in a residential area usually were similar in sizes and were distributed in a regular pattern. If these buildings were converted to a raster using the same cell size as the building area, the position

of the cell could determine whether or not most of the buildings would be absented.

Moreover, the central position method for rasterization could be problematic in an urban atmospheric dispersion study. For example, the skywalks in downtown Oklahoma City could be absented at nine meter resolution but presented at ten meter resolution. However, for atmospheric dispersion modeling, beyond a certain resolution, the size of skywalks were considered insignificant to the dispersion process. Alternatively, the dominant unit method, where a polygon was converted to raster cell when the polygon occupied the dominant area of the cell, might be a better option for building rasterization when the building data was used for urban atmospheric dispersion modeling.

4.4 Change in building volume

The changes in total building volume were not explicit from one to six meter resolutions. From seven meter resolution onward, the total building volume fluctuated up to three percent for central position method and nine percent for dominant unit method (Figure 3.16). As the horizontal resolution became coarser, the percentage of buildings volume unchanged steadily decreased from 99 percent at one meter resolution to 77 percent and 74 percent at twelve meter

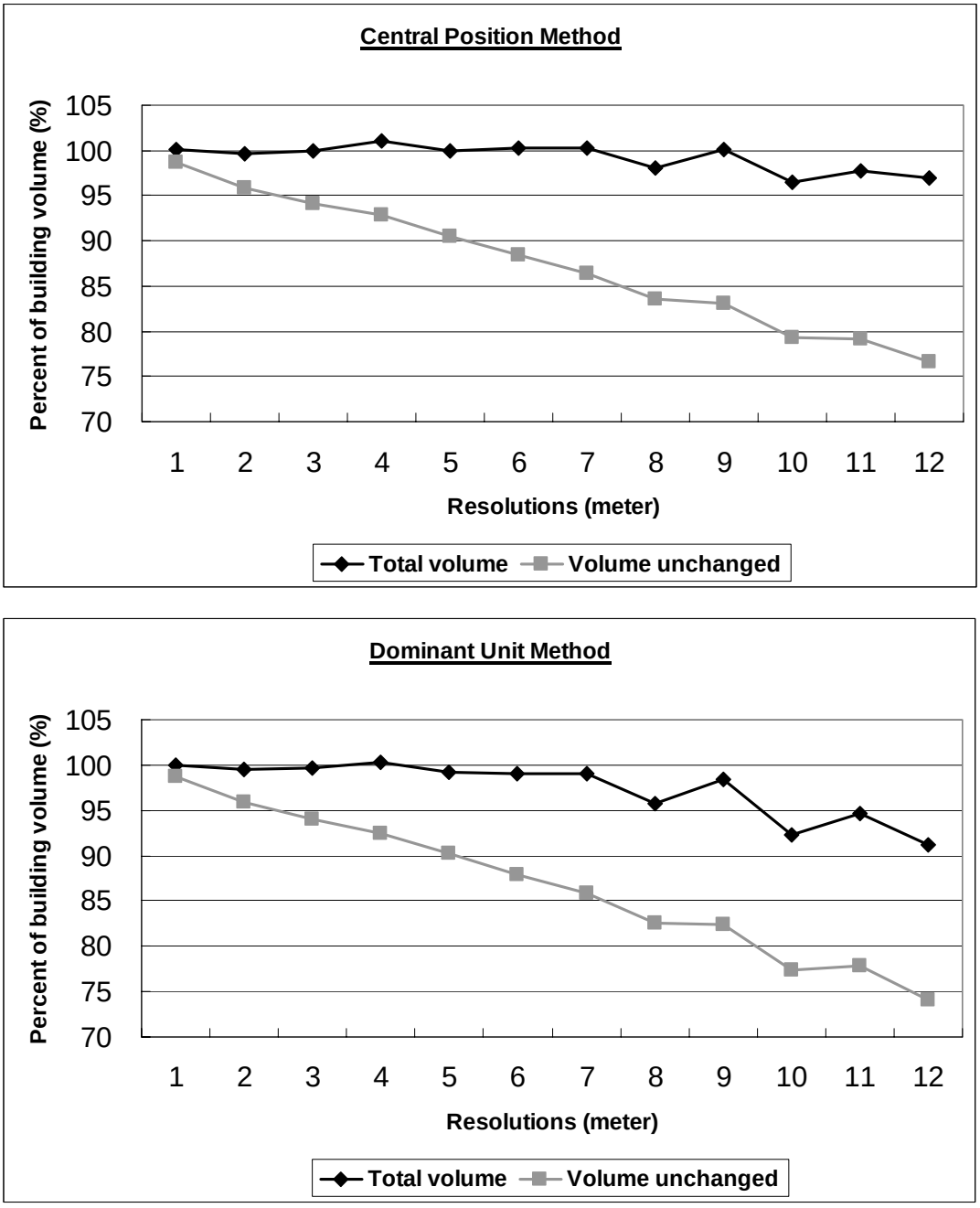


Figure 3.16. The change in total building volume after the conversion (solid lines) and the building volume unchanged (dotted lines) across twelve resolutions, comparing two methods of rasterization.

resolution for central position method and dominant unit method respectively. The changes in percentage of building volume omitted and committed were similar

for central position method (Figure 3.17). Both percentages increased from one percent to about 21 percent across twelve resolutions. For dominant unit method, the changes in percentages percentage between building volume omitted and volume committed increased from one percent to 26 percent and 19 percent respectively. Comparing two methods of rasterization, greater differences in building volume omitted and committed could only be found at resolution beyond ten meter.

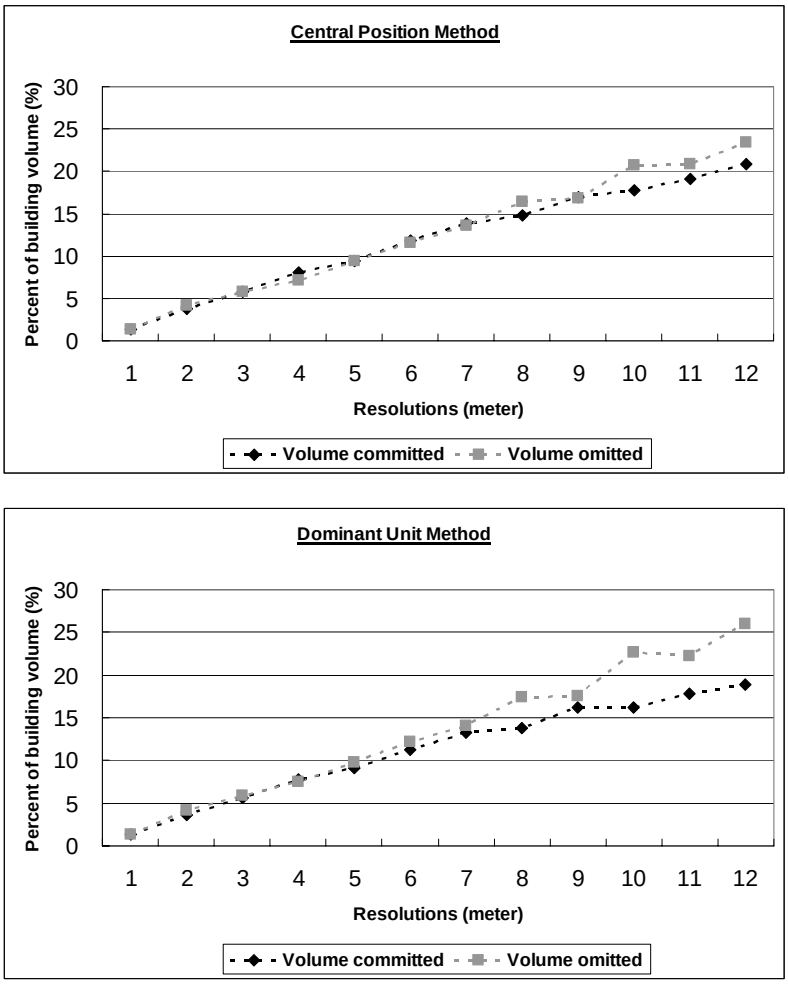


Figure 3.17. The percent building volume omitted (grey dotted line) and committed (black dotted line) after conversion.

At one meter resolution, most building volume omitted and committed was found at building edges. As the horizontal resolution became coarser, building volume omitted and committed was found not only at the edge of the buildings, but also across the study area because there was a remainder from dividing the elevation of the building by vertical resolution (Figure 3.18). Similar patterns were obtained from central position method and dominant unit method.

Figure 3.18. The spatial distribution of building volume omitted (right) and committed (left) at one meter (up) and six meter (bottom) resolutions.

In comparison to the change in building footprint, more buildings disappeared after the conversion incorporating a vertical resolution that affected building elevation. Those omitted buildings were mostly located at the outer range of the study area, where buildings were lower in elevation. However, unlike the changes in building footprint, buildings with elevation smaller than half of the resolution would not be occurred again once it disappeared at a vertical resolution. For modeling atmospheric dispersion in urban area, whether or not the source of emission was located near the absented buildings became more important because buildings near the emission source might change the dispersion pattern significantly. In order to reduce the influences of omitted or committed building volume towards the atmospheric dispersion modeling, the dispersion models should be able to assign a vertical resolution separately.

5. Conclusions

Based on the results of uncertainty assessment in building location, footprint area and volume, this study finds that coupling GIS with UADMs involves two main issues. First issue is concerned with practicality. Although a fine spatial resolution provides a more detailed representation of the building model, it is not always practical for UADMs. Also, data outputs from the

dispersion model output conform to the raster GIS data, but many GIS software do not readily display 3 dimensional data with an additional temporal dimension. Future studies may focus on enhancing displays of the 4D atmospheric dispersion results in GIS.

The second issue concerns data accuracy after conversion. As the spatial resolution becomes coarser, spatial uncertainty grows in the building model. A linear relationship is found between the spatial resolution and the mean shifting distances of building vertices. In the case of rasterization using the central position method, linear features with an area or height similar to the resolution may be absented without a predictable pattern. An alternative option is the central position method for rasterization when dealing with building models. The results of the spatial uncertainty assessment for coupling ArcMap and QUIC can be used to model spatial uncertainty in building models. Future work should assess the influence of building model uncertainty on the results of urban atmospheric dispersion modeling.

References

Bernard, L., and T. Kruger. 2000. Integration of GIS and spatio-temporal simulation models: Interoperable components for different simulation strategies. *Transactions in GIS* 4 (3):197-215.

Bregt, A. K., J. Denneboom, H. J. Gesink, and Y. v. Randen. 1991. Determination of rasterizing error: a case study with the soil map of the Netherlands. *International Journal of Geographic Information Systems* 5:361-368.

Congalton, R. G. 1997. Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion. *Photogrammetric Engineering & Remote Sensing* 63 (4):425-434.

Fedra, K. 1996. Distributed models and embedded GIS: Integration strategies and case studies. In *GIS and Environmental Modeling: Progress and Research Issues*, ed. M. F. Goodchild, 413-417. Colorado: GIS World Books.

Goodchild, M. F., ed. 1996. *GIS and Environmental Modeling: Progress and Research Issues*. Colorado: GIS World Books.

Goodchild, M. F., B. Buttenfield, and J. Wood. 1994. Introduction to visualizing data quality. In *Visualization in Geographical Information Systems*, eds. H. M. Hernshaw and D. J. Unwin, 141-149. New York: John Wiley and Sons.

Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind. 2001. Uncertainty, error, and sensitivity. In *Geographic Information Systems and Science*, 325-344. New York: John Wiley & Sons, LTD.

Martin, P. H., E. J. LeBoeuf, J. P. Dobbins, E. B. Daniel, and M. D. Abkowitz. 2005. Interfacing GIS with water resource models: A state-of-the-art review. *Journal of the American Water Resources Association* 41 (6):1471-1487.

Nyerges, T. L. 1993. Understanding the scope of GIS: Its relationship to environmental modeling. In *Environmental Modeling with GIS*, eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert. NewnYork: Oxford.

Pardyjak, E., and M. J. Brown. 2002. Fast-response modeling of a two building urban street canyon. Paper read at 4th AMS Symposium on the Urban Environment, May 20-24, at Norfolk, Virginia, USA.

Piwowar, J. M., E. F. Ledrew, and D. J. Dudycha. 1990. Integration of spatial data in vector and raster formats in geographical information system. *International Journal of Geographic Information Systems* 4:429-444.

Vardoulakis, S., B. E. A. Fisher, K. Pericleous, and N. Gonzalez-Flesca. 2003. Modelling air quality in street canyons: a review. *Atmospheric Environment* 37:155-182.

Wedhe, M. 1982. Grid cell size in relation to errors in maps and inventories produced by computerized map processing. *Photogrammetric Engineering & Remote Sensing* 48 (8):1289-1298.

Wentz, E. A. 2000. A shape definition for geographic applications based on edge, elongation and perforation. *Geographical Analysis* 32 (2):95-112.

Williams, M., M. J. Brown, and E. R. Pardyjak. 2002. Development of a dispersion model for flow around buildings. Paper read at 4th AMS Symposium on the Urban Environment, at Norfolk, Virginia, USA.

Zhu, A.-X. 2005. Research Issue on uncertainty in geographic data and GIS-based analysis. In *A Research Agenda for Geographic Information Science*, eds. R. B. McMaster and E. L. Usery, 167-223. Boca Raton, Fla: CRC Press.

Chapter 4 : The Influences of Spatial Uncertainty toward Urban Atmospheric Dispersion Model

Abstract

Urban atmospheric dispersion models (UADMs) estimates the atmospheric circulation and predict how pollutants are dispersed within an urban environment. Many important decisions are made in reference with the output from UAMDs. However, UADMs output may be unreliable due to uncertainties in input data, parameter settings and model assumptions. Many research efforts focus on examining the sensitivities of parameter settings toward UADMs. Complementarily, this paper compares the influences of uncertainties from spatial input data and meteorological parameters toward UADMs, using Quick Urban and Industrial Complex dispersion model (QUIC) as an example. Buildings are one of the fundamental spatial data sets for UADMs. Often, spatial uncertainty is introduced to the UADMs through changes in data resolution and formats, such that locations, shapes and heights of buildings are modified during the transformation. With a field experiment in Oklahoma City and Monte-Carlo simulation, this study examines the influences of uncertainties from both meteorological parameters and building data. The building data are perturbed base on previous assessment of spatial uncertainty. Wind speed and wind direction are perturbed according to the uncertainty obtained from field

observation. Then, under the same scenario, two groups of tracer chemical (i.e. Sulfur Hexafluoride - SF6) dispersion simulations are generated: one group with uncertainty from meteorological parameters only and the other group with uncertainties from both building data and meteorological parameters. Each group contains thirty simulations. Two groups of simulations are then compared with ground observation data. Results show that locations near the release point and building edges show significantly higher simulated concentration than observations by ground sensors for both simulation groups. The overestimation of concentration is particularly apparent within the first two minutes of the release and near the release points. Over-estimation is also prominent from ground level up to twelve meters around the buildings. Furthermore, at five out of eight sampling sites, simulations with uncertainties from both meteorological parameters and building data generate a closer concentration than the other simulation group that represents uncertainty only from meteorological parameters. The finding seems suggest that uncertainty from meteorological data and building data may have canceling effects in dispersion modeling. Future research should examine additional scenarios and with other UADMs and identify the mechanisms of uncertainty effects from multiple sources.

1. Introduction

Urban atmospheric dispersion models (UADMs) are commonly used to monitor air quality, predict air pollutants dispersion, establish environmental regulations, and response to accidental hazardous gas releases in urban area. One of the key challenges in UADMs is to account for the complex building structures and arrangement in an urban environment.

As all process models, UADMs are subject to uncertainties from various sources. For instance, the results of UADMs can be influenced by meteorological uncertainties such as changes in wind speed, wind direction, and wind profile; spatial uncertainties such as inaccuracy in building dimension, location and terrains; and scenario uncertainty such as emission type, duration and amount (USEPA, 1992). Therefore, any UADM outcome cannot be confidently interpreted and used for decision making without understanding the model's behavior towards the influences of uncertainties from potential sources. Uncertainty in this paper refers to the discrepancies between the data and the reality the data represent, without ascertain true differences. Therefore, it is usually represented by statistical probability at a confidence level. For example, at a 90 percent confidence level, the accuracy of the digital elevation model (DEM) falls within +/- 30 meters. From the perspective of environmental modeling, uncertainty is

inevitable and presented in every aspect such as model assumptions and data representation (Goodchild, 1993).

A good understanding of the effects of data uncertainties on UADM outcomes provides the basis of model reliability. Consequent decision making processes such as planning of evacuation route in case of emergency and planning of buildings in urban downtown area will need to consider the inherit uncertainty in UAMS estimates. Fundamentally, understanding uncertainty provides references to model development for improved estimation of atmospheric dispersion in urban environments.

While there are many sources of uncertainty, this paper focuses on examining the influences of spatial uncertainty toward the results of UADMs, using a newly developed UADM-QUIC (Quick Urban and Industrial Complex atmospheric dispersion system), as an example. This study examines the influences of spatial uncertainty using a combination of Monte Carlo approach and sensitivity analysis. First, two sets of uncertainties are generated: one is meteorological uncertainty and the other is spatial uncertainty. Second, under a controlled dispersion scenario, QUIC model is run with two sets of parameters. First set of parameters is perturbed with meteorological uncertainty and second set of parameters is perturbed with uncertainty from meteorological and building data.

Finally, the results of the QUIC model are examined and validated with ground observations.

2. Background

UADM estimates atmospheric circulation and pollutants dispersion within an urban environment. Urban dispersion modeling requires advanced computation power, detailed information of urban environment and knowledge of fine-scale atmospheric circulation. With the improvement in computer and remote sensing technologies, UADMs have become a common tool in studying air quality and atmospheric dispersion. However, uncertainty remains a challenged issue for scientists, modelers and decision-makers.

According to Fox (1984), uncertainty in atmospheric dispersion models originates from two main sources: one is the stochastic nature of atmospheric motion, and the other is the errors in the input data. Stochastic nature of atmospheric motion such as turbulence is by large random and cannot be predicted precisely. Therefore, every model's prediction inherits some degree of uncertainty. Meanwhile, urban dispersion modeling also accounts for errors in input data, such as model parameters, emission characteristics, meteorological and terrain conditions. Errors in these data and inadequate settings of model physics

lead to uncertainty in model's results. Theoretically, uncertainty can be reduced through improvement in input data accuracy and computation algorithms.

In practice, however, improvement in data accuracy does not necessarily reduce uncertainty in model results. A typical case is the uncertainty from spatial data, or we called spatial uncertainty. Spatial uncertainty is introduced mainly during the data collection and conversion processes. For examples, an urban environment with complex building structures and constructions can challenge the development of an accurate, representative three dimensional building model. Even with current remote sensing technology such as LiDAR (i.e. Light Detection and Ranging), which can achieve one-meter horizontal resolution and sub-meter accuracy on building heights, the building data accuracy can also be varied greatly due to different environment settings and spatial arrangements of buildings and urban landscape (Burian et al., 2004). Moreover, most UADMs requires model specific data formats (USEPA, 1992). Data transformation to the specified data format inevitably leads to data modification. In addition, the computation-intensive nature of UADMs with spatial data at a fine resolution can exhaust computation resources. Therefore, some UADMs are designed for coarse building data to reduce the processing time. Consequently, these UADMs encounter an additional degree of uncertainty. Issues with the data accuracy, data

requirements and model operation further raise the importance of understanding the influence of spatial uncertainty on urban atmospheric dispersion modeling. Research is aiming to address the influence.

While research on the influence of spatial uncertainty on UADMs is lacking, several studies examined uncertainty effects from other input data such as meteorological condition, emission characteristics and model parameters. For examples, Yegnan et.al. (2002) analyzed the influences of wind speed and ambient temperature toward the results of ISCST dispersion model using Taylor series approach and Monte Carlo approach. They found that the dispersion results were more sensitive to wind speed than ambient temperature. Sullivan et.al. (2004) combined two dispersion modeling system (ISCST3 and TOXST) to predict the emission of fumigants. By conducting a sensitivity analysis on key input parameters, they found that emission rate was a critical input parameter to modeling outcomes. Manomaiphiboon and Russell (2004) evaluated the uncertainties of five model parameters towards a Lagrangian particle model. Among the friction velocity, mean surface turbulent heat flux, surface roughness height, mean surface temperature and a universal constant in model equations, they found that the friction velocity was the most influential factor that affects the mean ground-level concentration. Sax and Isakov (2003) estimated the uncertainty of

hexavalent chromium concentrations from welding operating using ISCST3 and AERMOD. By comparing emissions, spatial and temporal allocation of emissions, model parameters and meteorology, they found that Gaussian models are sensitive to all components but most sensitive to emissions amount. Together, these studies highlight the influential factors and model assumptions to model sensitivity and outcome uncertainty.

With uncertainties coming from different input data sources, uncertainty analysis is the key to informed decision making based on the results of UADMs. Analysis of uncertainty has been an important research topic in different fields of atmospheric study such as air quality models (Britter and Hanna, 2003), exposure assessment (USEPA, 1992; Cullen, 1999), and weather forecasting (Brooks et al., 1995; Krzysztofowicz, 1998). In general, Taylor series analysis and Monte Carlo simulations are two common approaches for uncertainty analysis. Taylor series approach analyzes the uncertainty by evaluating the derivatives of the model output function while Monte Carlo approach simulates the uncertainties and evaluates the perturbed outputs of the model. In comparison of the two approaches, Monte Carlo approach is more suitable for stochastic models such as UADMs, while Talyor series approach is more appropriated for deterministic models (Heuvelink, 1998).

However, Monte Carlo approach is often limited by computation cost since UADMs usually consume huge data storage and require long computation time. Besides the intensity of computation needs, Monte Carlo approach describes only the variation of the model results caused by the uncertainties of input data. Addition analyses are needed to distinguish the influence of uncertainty from different input data. Expanding upon the Monte Carlo approach, this research examines the influence of spatial uncertainty towards a newly developed UADM – QUIC, using Monte Carlo approach with sensitivity analysis. The Central Business District in downtown Oklahoma City (OKC) is used as an example and the study simulates part of the field dispersion experiment taking place in OKC at 2003. The detailed research design is described in the following section, followed by results, discussions and a conclusion.

3. Research Design

Figure 4.1 shows the conceptual design of the research. Using QUIC dispersion model as an example, two groups of simulations are generated under the same scenario. The first group (Group A) of simulations associates with

meteorological uncertainty only (i.e. changes in wind direction and wind speed).

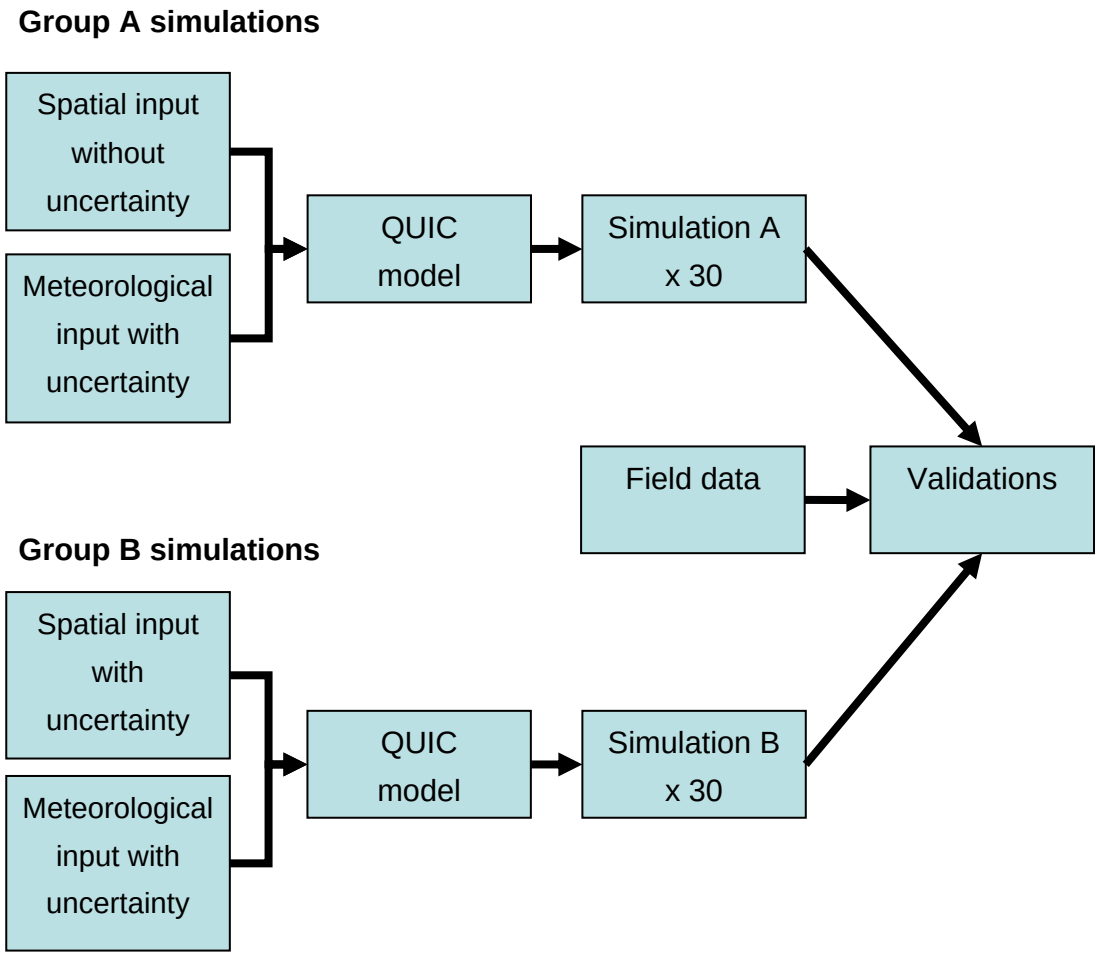


Figure 4.1. Conceptual flowchart of the research design.

This group of simulations assumes with no spatial uncertainty (i.e changes in building locations and dimensions) which are presumed by the most UADMs. The second group (Group B) of simulations associates with both meteorological and spatial uncertainty. This group of simulations is assumed to be more representative of the reality as uncertainty can be involved in both meteorological and building data. Then, the model outcomes from two groups of simulations are

examined and the model outcomes are compared with field observation. Thirty simulations are chosen based on the consideration of computation power and the Central Limit Theorem. Without prior knowledge on the distribution of tracer gas concentration, the statistical rule of thumb suggested thirty simulations as a starting point.

The simulations are run on a Dell desktop with 512 RAM and Pentium 4 2GHz CPU. Each group of simulations includes thirty runs of the model under the considerations of the limitation in computation power and the requirement for statistical significance. Only wind speed and wind direction are perturbed for simulating meteorological uncertainty because these variables are the basic meteorological inputs for UADMs. For spatial uncertainty, it is assumed that the major source of spatial uncertainty is attributed to data conversion to meet the data requirements for model input. This research assumes that there was no scenario uncertainty because scenario uncertainty is infeasible to quantify, given that each scenario varies greatly in terms of emission type, location, amount and duration. The detailed description of QUIC dispersion model, settings of scenario, generation of spatial and meteorological uncertainty and the examinations of the models results are outlined as follows.

3.1 QUIC dispersion model

QUIC is a new dispersion model developed by Los Alamos National Laboratory. The model is designed to provide fast predictions on the dispersion of air contaminants in an urban environment with limited computation power (i.e. laptop). QUIC is composed of two parts: QUIC-URB and QUIC-PLUME. QUIC-URB calculates the three dimension wind field among buildings using empirical algorithms for certain flow regions near buildings (such as wakes, cavities, street canyons etc.). In these empirical parameterizations, the dimensions of the different flow regions are expressed as functions of the building dimensions. It can thus be expected that QUIC-URB is particularly sensitive to uncertainties in the GIS building database (Pardyjak and Brown, 2002). QUIC-PLUME simulates the dispersion of air contaminates using wind field from QUIC-URB and Langevin random walk equations (Williams et al., 2004). Dispersion towards building surfaces is handled as the particles are elastic. Reflection angle is determined by calculating the direction with the largest penetration into the wall. The reflected distance is equal to the penetrated distance. Moreover, the model is tested thrice with identical input parameters to ensure that the model outputs are reproducible. The model also comes with a graphical user interface (GUI) for data input, parameter settings and results

visualization that are built inside MATLAB software. Figure 4.2 shows the QUIC user interface. Detailed descriptions of the QUIC could be found on the QUIC's user manual.

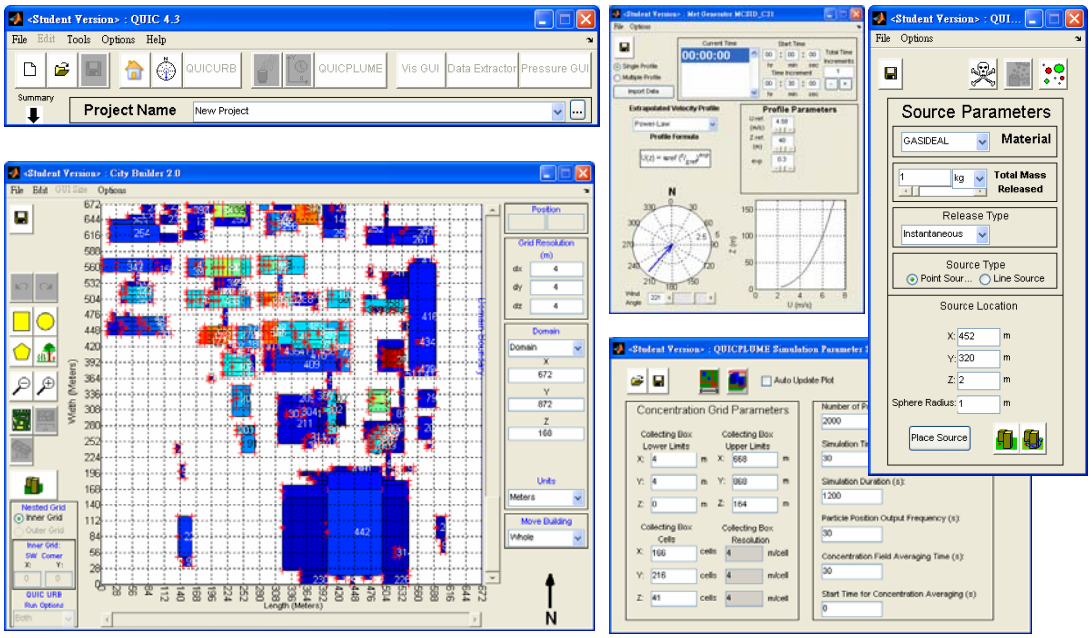


Figure 4.2. QUIC user interface modules.

3.2 Scenario settings

Table 4.1 shows the detail settings of the scenario. The dispersion field experiment took place in downtown OKC, July 2003, also known as Joint Urban 2003. The study was one of the largest dispersion studies in United States, sponsored by U.S. Department of Energy, U.S. Department of Homeland Security and U.S. Department of Defense. The study area was shown in Figure 4.3. It was located at downtown Oklahoma City with an extent about 612 meters by 830

meters. Flat terrain and well-defined central area reduced the complexity of UADM. Given the scale of tracer gas release, this study assumed that the concentration of tracer gas was not significant outside the extent of the study area. With over 200 tracer gas samplers and wind samplers covered in downtown OKC, tracer gas (i.e. sulfur hexafluoride, SF6) was released at three different locations over ten intensive operation periods (IOP). The scenario settings mimicked part of the second IOP, conducted at July 2nd 2003, where one kilogram of SF6 was released instantaneously near Westin Hotel, on the west side of Broadway Avenue. Under the scenario, the QUIC ran the simulated dispersion for twenty minutes with thirty seconds interval. The simulations were then compared with the results from fieldworks.

Table 4.1. Settings of dispersion scenario.

Simulation parameters	
Number of particles:	2000
Simulation time step:	30 seconds
Simulation duration:	1200 seconds
Concentration field average time:	30 seconds
Source parameter	
Total mass release:	1 kg
Source type:	Point
Release type:	Instantaneous

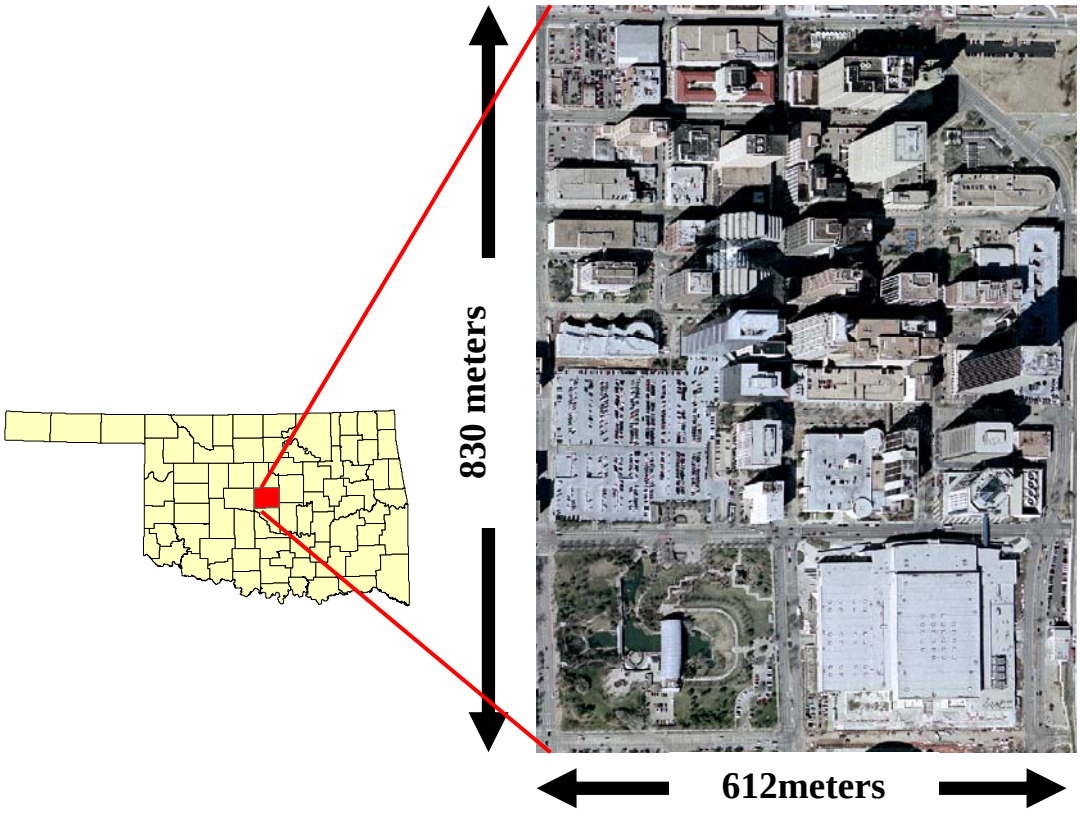


Figure 4.3. The study area – Downtown Oklahoma City. Date souce: U.S.G.S., date: March-2002.

3.3 Spatial and meteorological uncertainties

An assessment of the influence of spatial uncertainties towards results of QUIC requires probability distributions of the uncertainties. This study assumes that the data conversion is the key source of spatial uncertainty as a comprehensive study of the 3D building model has identified that locations and sources of uncertainty in building data (Cheuk and Yuan, 2008). QUIC employs a unique data format for building input, and the building data in QUIC are stored as either

rectangles or circles. Each record contains height, width (length in y co-ordinate), length (length in x co-ordinate), xfo (minimum x co-ordinate), yfo (middle y co-ordinate), and zfo (base height). Nevertheless, buildings, in reality, are often in complex shapes. As a result, when data are converted into QUIC format, the location of the buildings may be shifted. Based on the spatial uncertainty involved during the data conversion from ArcGIS to QUIC, a frequency distribution of spatial uncertainty at four meter resolution in terms of direction and magnitude is computed. Then, the distribution is used to perturb the building data (Figure 4.4, also see Appendix I). The four meter resolution is chosen because of the computational limitations of the computer system used in the study. With resolutions finer than four meter, dispersion simulation using QUIC is not feasible because the computer will freeze.

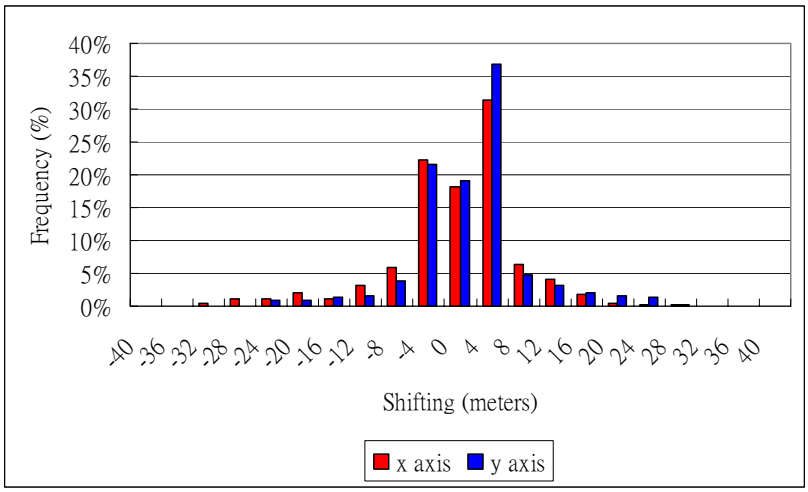


Figure 4.4. Frequency distribution of spatial uncertainty at four meter resolution.

For meteorological uncertainty, this study assumes that the main source of uncertainty comes from variations in wind speed and wind direction. Therefore, the meteorological data obtained during the Joint Urban 2003 project is used to derive meteorological uncertainty. Figure 4.5 shows the locations of two wind samplers used to generate the meteorological uncertainty. These samplers are chosen because they are located at relatively open space compared to other samplers. Site fourteen is at the roof of the buildings, and site fifteen is at a sampling tower forty meters above ground. Figure 4.6 shows the wind speed and wind direction variations during the second IOP. Table 4.2 shows the meteorological parameters used in the thirty simulations.

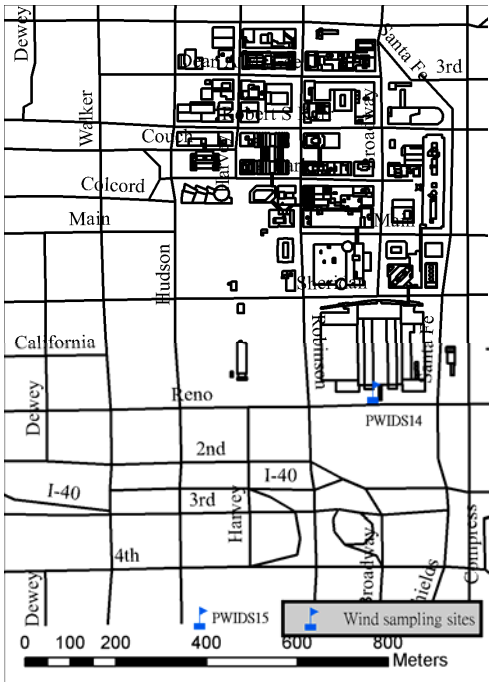


Figure 4.5. Locations of two wind samplers which are used for generating meteorological uncertainty.

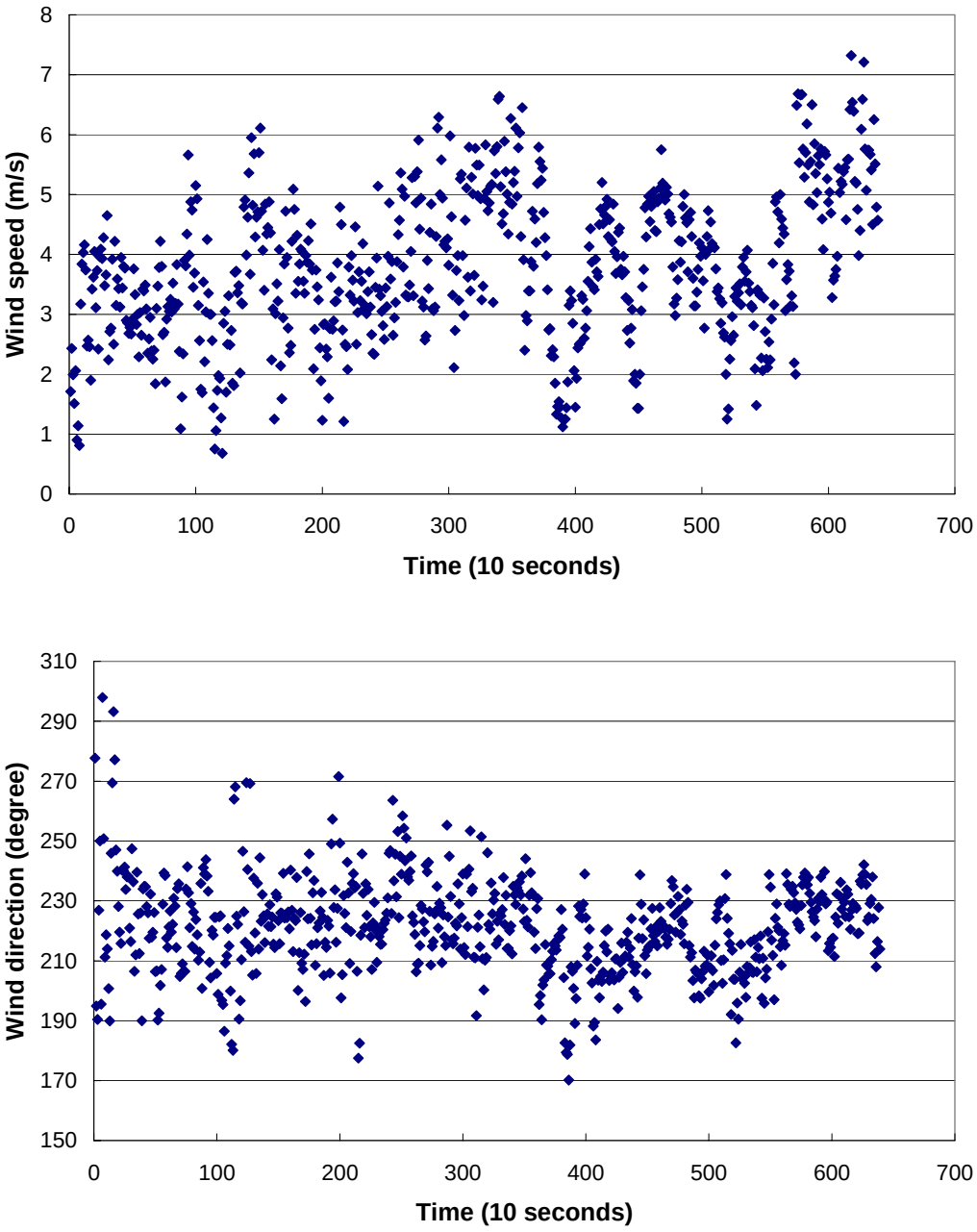


Figure 4.6. Variations of wind speed and wind direction (meteorological uncertainty) during second IOP.

Table 4.2. Wind speeds and wind directions used in the simulations.

Simulation ID	Wind speed (m/s)	Wind direction	Height (m)
A1	4.65	232	26
A2	3.75	235	40
A3	2.7	225	26
A4	4.88	222	40
A5	2.65	234	26
A6	3.39	258	26
A7	2.7	223	26
A8	4.84	214	26
A9	4.71	257	26
A10	4.08	239	40
A11	4.88	207	26
A12	3.23	228	40
A13	2.89	217	26
A14	5.26	227	40
A15	4.3	216	40
A17	4.09	184	40
A18	1.85	209	40
A20	5.67	227	40
A21	4.58	221	40
A22	1.23	232	26
A23	2.52	206	40
A24	1.43	229	40
A25	2.81	232	26
A26	5.8	222	26
A27	4.84	243	26
A28	3.23	203	40
A29	2.69	216	40
A30	4.86	240	26
A31	4.05	217	40
A32	4.83	217	40

3.4. Examinations of model outputs

With Monte Carlo method, this study perturbs the spatial and meteorological inputs and generated two groups of the model outputs. Each group contains thirty simulations and each simulation is composed of a 3D field (612m x 830m x 150m) of concentration over twenty minutes. A cube in the 3D field of concentration is four by four by four meters. Concentration at a particular height and time is expressed as a raster layer in GIS, and therefore one simulation contained 1640 raster layers (40 time steps x 41 layers in height). Additional three layers in height are required by the QUIC model in order to compute the wind circulation. With massive amount of output data, the study focuses on two aspects: 1) the spatial distribution of integrated concentration over time and, 2) the spatial distribution of concentration over time.

First, this study looks at the spatial distribution of integrated concentration, the sum of concentration across simulation time. Since each group of output data sets contains thirty simulations, the mean and variance of integrated concentrations are calculated. The absolute differences in mean integrated concentration are also computed and student T test is conducted in order to identify locations with significant differences at 95 percent confidence interval based on mean integrated concentration. The null hypothesis is that there are no significant differences

between the mean integration concentrations between two groups of outputs. After prescreen of the simulation, examination of mean and variance integrated concentration are stopped beyond twenty meters in height as very small concentrations are found above this level.

Second, this study examines the spatial distribution of concentration over time. For each group of simulations, the mean and variance concentration for every time step are computed. Student T tests are also performed to determine whether there are any significant differences in the concentrations between two groups of output data sets at every time steps up to 95 percent confidence interval. The null hypothesis is that there are no significant differences between the mean concentrations between two output data sets at every time step. However, prescreen of the simulations determines to discard concentrations after the tenth time step as the concentrations are relatively small and dispersed.

3.5. Evaluations of model simulations

This study compares the concentration between two sets of simulations and evaluates the simulation results with ground observations in Joint Urban 2003 in order to examine the influence of spatial uncertainty towards model accuracy. Figure 4.7 shows the sensor locations. Each observation point is equipped with a

real-time sampler operated by Lawrence Livermore National Laboratory (LNLL). These samplers are chosen for validations because they are close to the release point and are determined by meteorology collaborators with good data integrity.

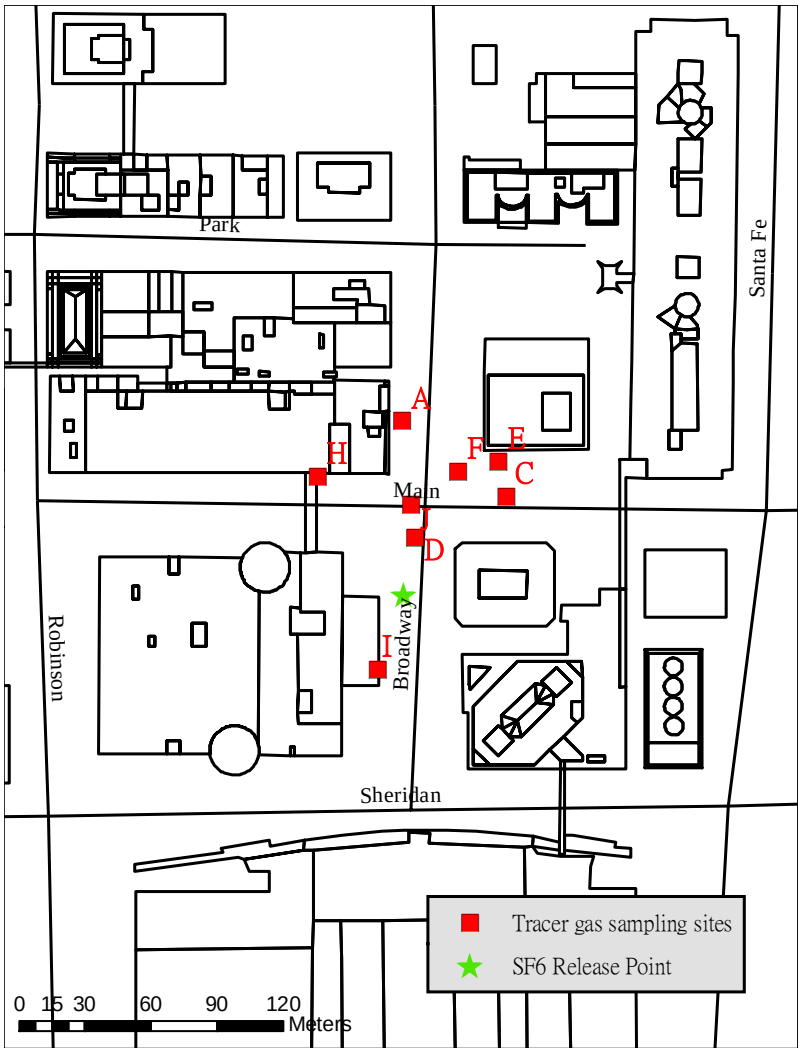


Figure 4.7. Locations of the ground observation sites.

However, there are only eight real-time sampling sites, and the observations are only available at ground level.

There are many different methods to compare the ground observation with model simulations. Chang and Hanna (2004) reviewed and presented various measures that used to evaluation atmospheric dispersion models. For examples, fractional bias, geometric mean bias, the normalized mean square error, correlation coefficient and the faction of predictions within a factor of two of observation. They suggested multiple measures for model evaluation since each measure carried its own limitations and advantages.

This paper first compares the simulated integrated concentration with ground observations for each sampling site. Then, the predicted concentration is compared with observatory data by pairing them with both time and space. Four common evaluation measurements are used in the evaluation and their equations are shown as follows:

1) Fractional bias (FB) = $\frac{\left(\overline{Co} - \overline{Cp}\right)}{0.5\left(\overline{Co} + \overline{Cp}\right)}$ ·····equation 1

2) Geometric mean bias (MG) = $\exp\left(\overline{\ln Co} - \overline{\ln Cp}\right)$ ·····equation 2

3) Normalized mean square error (NMSE) = $\frac{\overline{(Co - Cp)^2}}{\overline{Co} \cdot \overline{Cp}}$ ·····equation 3

4) Geometric variance (VG) = $\exp\left[\overline{(\ln Co - \ln Cp)^2}\right]$ ·····equation 4

, where Co is the observed values; Cp is the predicted values; overbar represents the average over the dataset.

4. Results and Discussions

4.1 Spatial distribution of integrated concentration

Figure 4.8 shows the mean integrated concentration for both Group A and B simulations at the ground level. The mean integrated concentrations of both groups were classified into quintiles in order to compare their spatial distributions. Locations with the highest mean integrated concentration were shown in red while the lowest mean integrated concentration were shown in dark green. In general, both groups simulated similar dispersion patterns. Dispersion progressed along the north-east side of the study area. The mean integrated concentration gradually declined from release points toward northeast as the simulated wind direction went from south-west (i.e. between 184 and 258 degree). The location for the highest mean integrated concentrations were also very similar, both appeared at the northwest of release point, mainly caused by eddies generated by the building west of release point.

Nevertheless, this study identified three main differences between the two simulation groups by subtracting the mean integrated concentration of Group B from Group A at the ground level (Figure 4.9). The blue color represented area where Group B generated a higher mean integrated concentration than Group A,

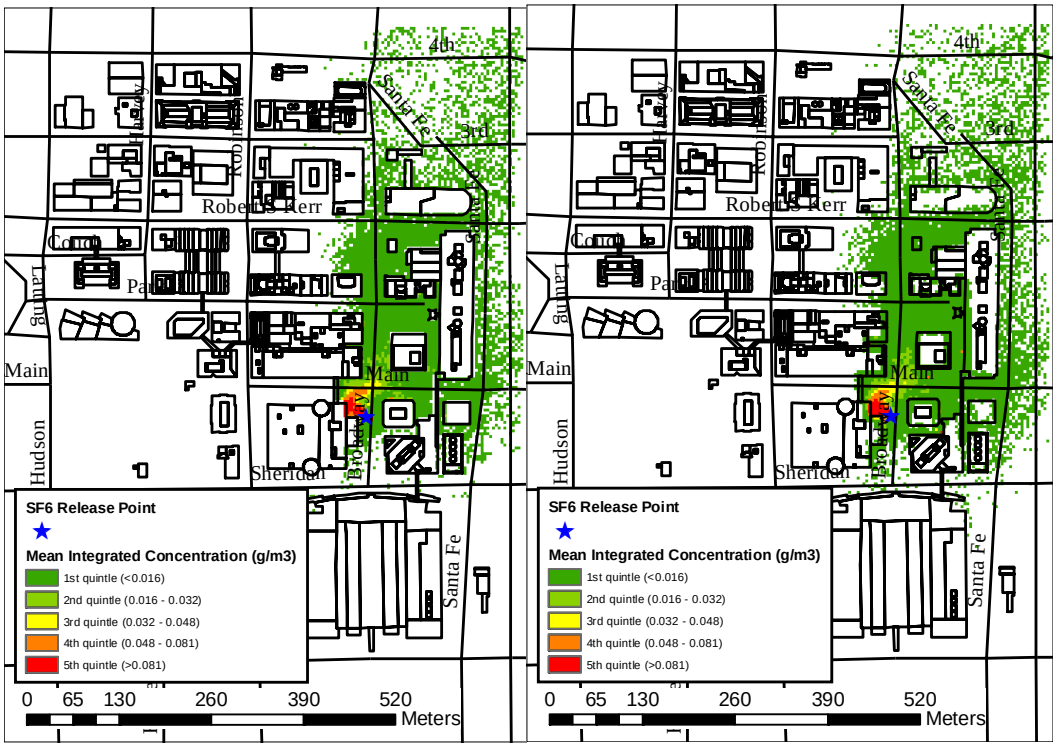


Figure 4.8. Mean integrated concentration at the ground level for simulation Group A (left) and Group B (right).

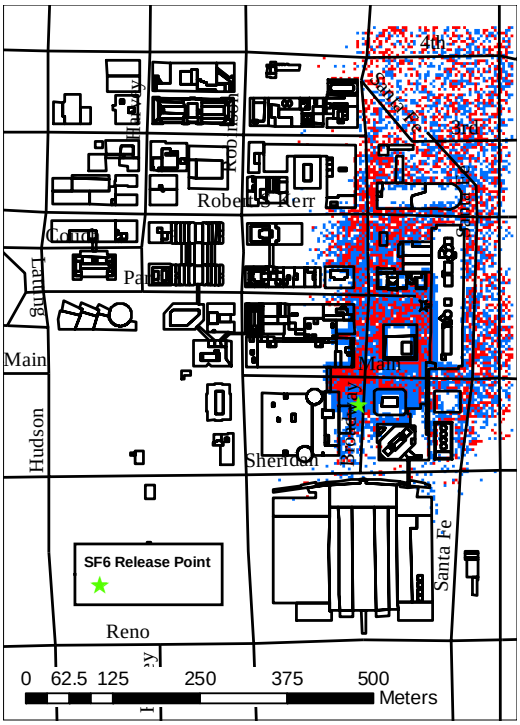


Figure 4.9. Differences in mean integrated concentration at the ground level. Area with blue color represents Group B generated a higher concentration; while area with red color represents Group A generated a higher concentration.

while the red color represented area represented the opposite. Higher mean integrated concentration could be found inside the building footprints for Group B as the location of buildings in Group B were perturbed and shifted. Moreover, Group B generated a higher concentration around the edge of building as well as around the release point (i.e. about 26m).

Similar comparisons between the two groups of simulations were identified at different height levels. Figure 4.10 displayed results from both groups at five different height levels. Only five height levels were displayed as the mean integrated concentration above level five (i.e. twenty meters) was negligible when compared to the concentration at ground level. Since Group B simulated higher mean integrated concentration adjacent to the release points at various height levels, this study suggested that there might be more vertical mix of tracer gas when spatial uncertainty was accounted.

This study also performed student T tests to find out whether there were significant differences between the mean integrated concentrations from both simulations at various height levels. Figure 4.11 shows the location where the mean integrated concentrations were significantly different between two simulation groups at the 95 percent confidence interval. The blue color

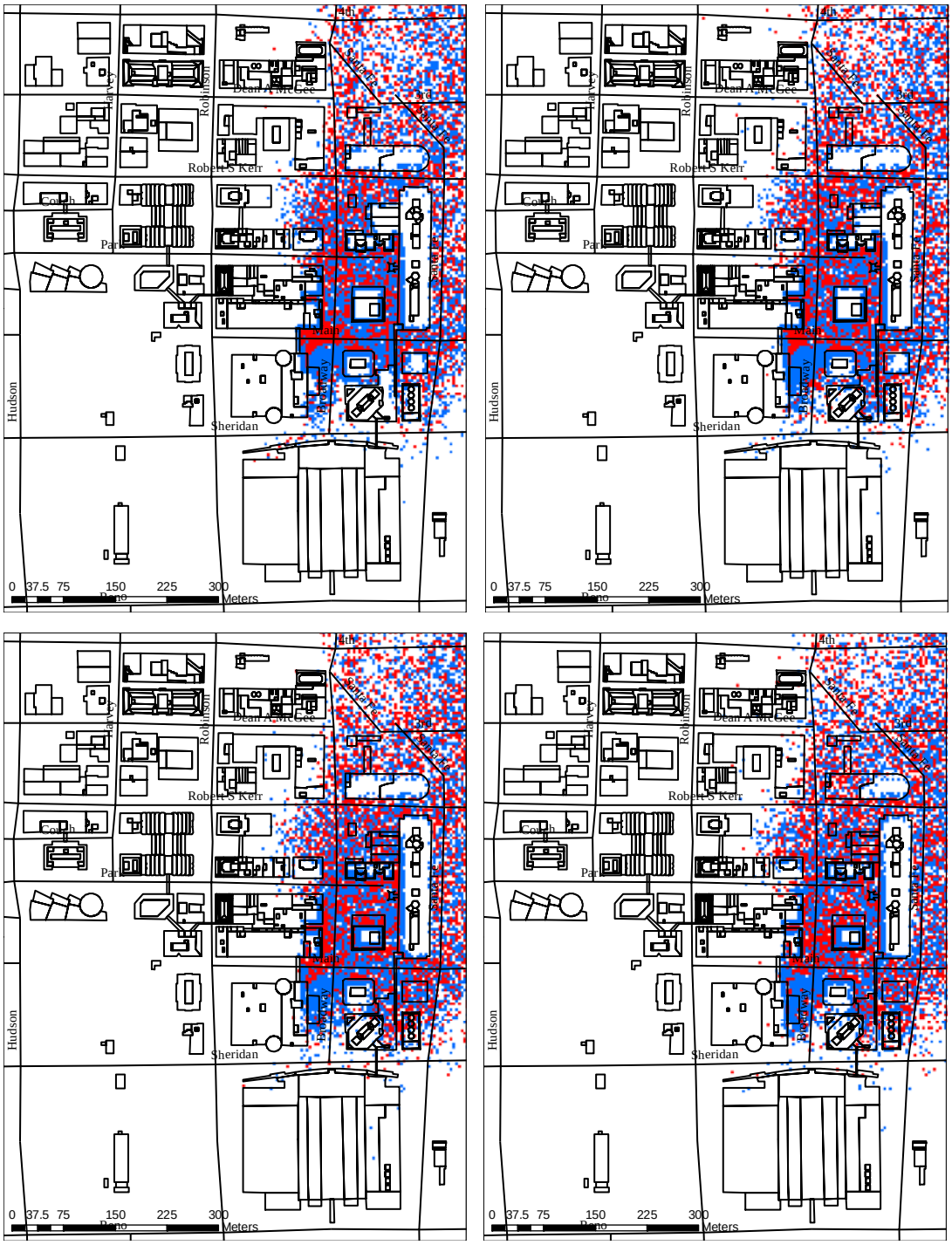


Figure 4.10. Differences in mean integrated concentration at various height level. From left to right, top to bottom, represents height level 2, 3, 4 and 5 respectively.

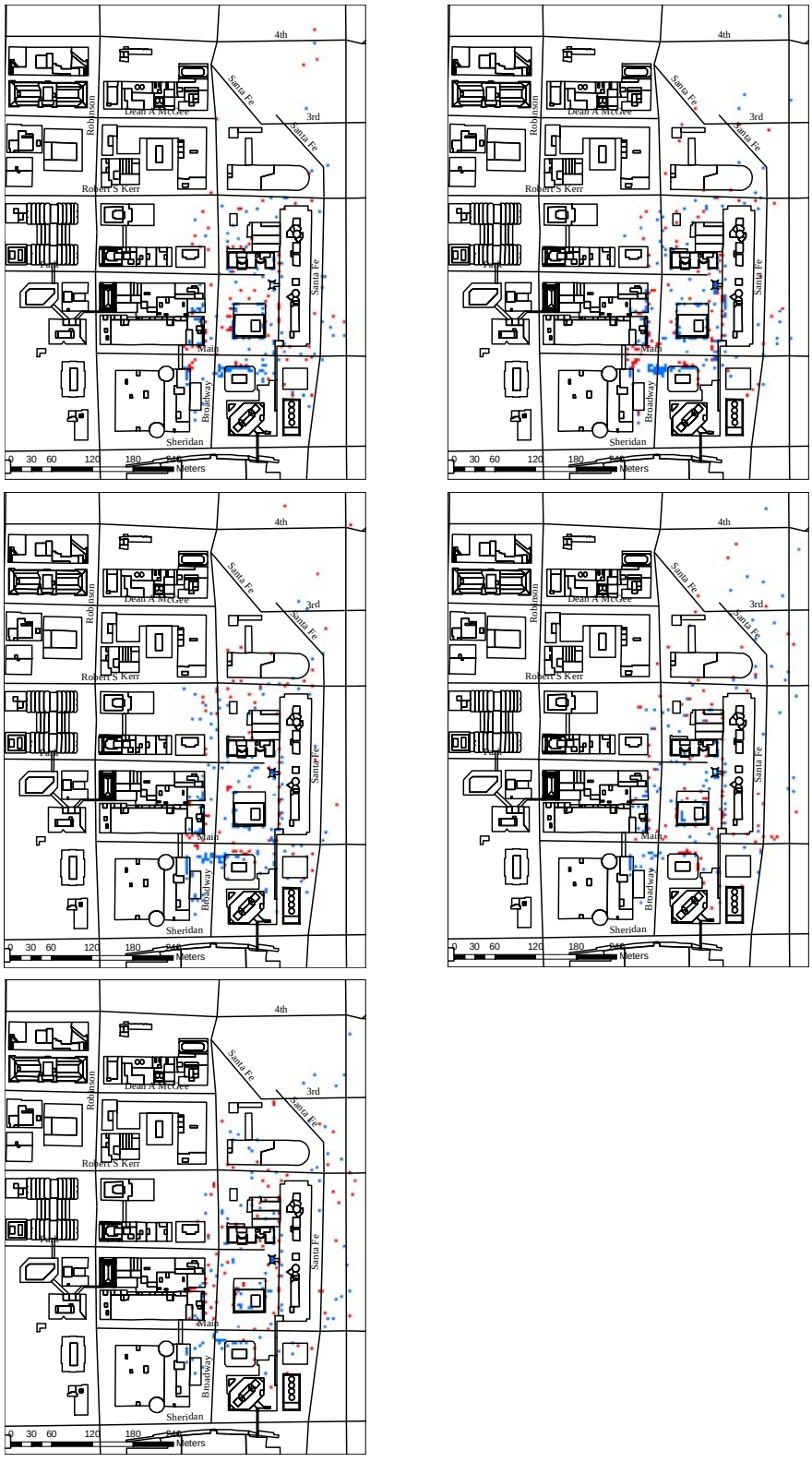


Figure 4.11. Locations with significant differences in mean integrated concentration at various height level. From left to right, top to bottom, represents height level 1, 2, 3, 4 and 5 respectively.

represented location with significantly higher mean integrated concentration from Group B, while the red color represented locations with significantly higher mean integrated concentrations from Group A. Across most of the study area, there were no significant differences in mean integrated concentrations between Group A and B except at the edges of buildings near the release point where small clusters of blue and red points appeared. The differences were more extensive at the first, second and the third height levels. Away from the release point, individual red and blue points were found but those could be treated as differences caused by the randomness of dispersions between two simulation groups.

4.2 Spatial distribution of concentration over time

Figure 4.12 shows the differences of mean concentrations between two groups of output data sets at the ground level, from time-steps two to nine. For consistency, the same color scheme was applied in Figure 4.9 to Figure 4.12. Step one was the initial condition, the same for every simulation, so it was ignored. From step two to step four, differences between output Group A and Group B were apparent. As the tracer gas dispersed from the release point to northeast area, higher concentration accumulated near the release point in Group B simulations. From step five to nine, the tracer gas was dispersed to the northeast side of the

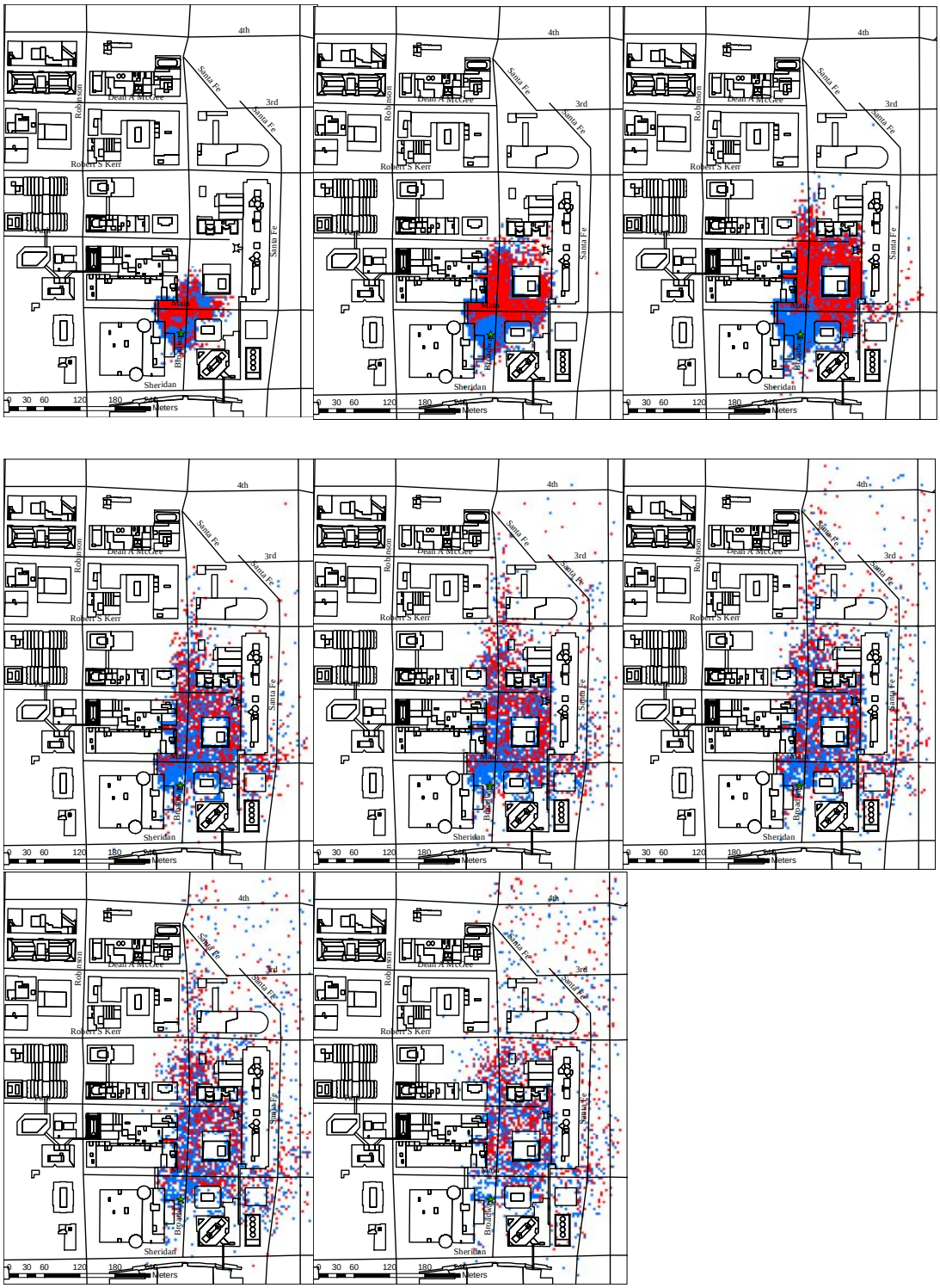


Figure 4.12. Differences in mean concentrations between two groups of simulations at ground level (0 to 4 meters), from time step two to nine. Sequence from left to right, top to bottom.

release point in no particular pattern, except for higher concentrations from Group B simulations. Concentration differences at other heights level were examined. Similar patterns at the lower altitudes appeared at height level-two to level-four. The concentrations simulated from Group B adjacent to the release point were higher than concentration simulated from Group A at time step two, three and four. Beyond step four, the concentrations were dispersed and mixed inside the street canyon northeast to the release point. Around twenty meters above the ground (i.e. height level five), this pattern dissipated but was still discernible. (Figure 4.13 to 4.16). These results suggested that the spatial uncertainties accounted in the QUIC dispersion model (i.e. Group B) might have hampered the dispersion of tracer gas compared to those simulations without spatial uncertainties (i.e. Group A).

Although higher concentration could be found around the release point with spatial uncertainties, not all of the results were significantly different. At the 95 percent confidence interval, significant differences in concentration could only be found at locations near the release point (within thrity meters), and within the first three time-steps (within 1.5 minutes). Figure 4.17 shows the maps at different height levels and time steps which significant differences could be observed. The

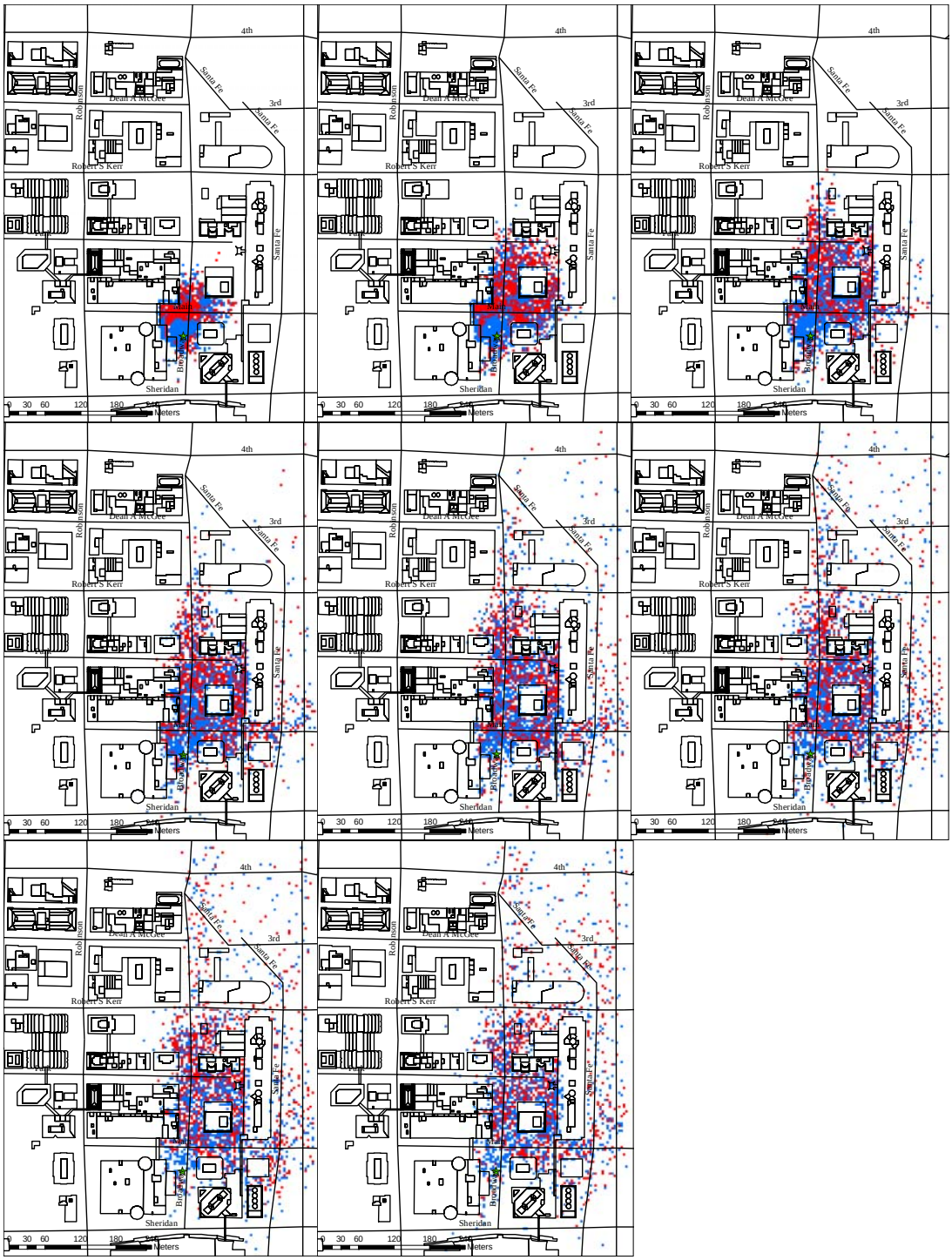


Figure 4.13. Differences in mean concentrations between two groups of simulations at level two (four to eight meters), from time step two to nine. Sequence from left to right, top to bottom.

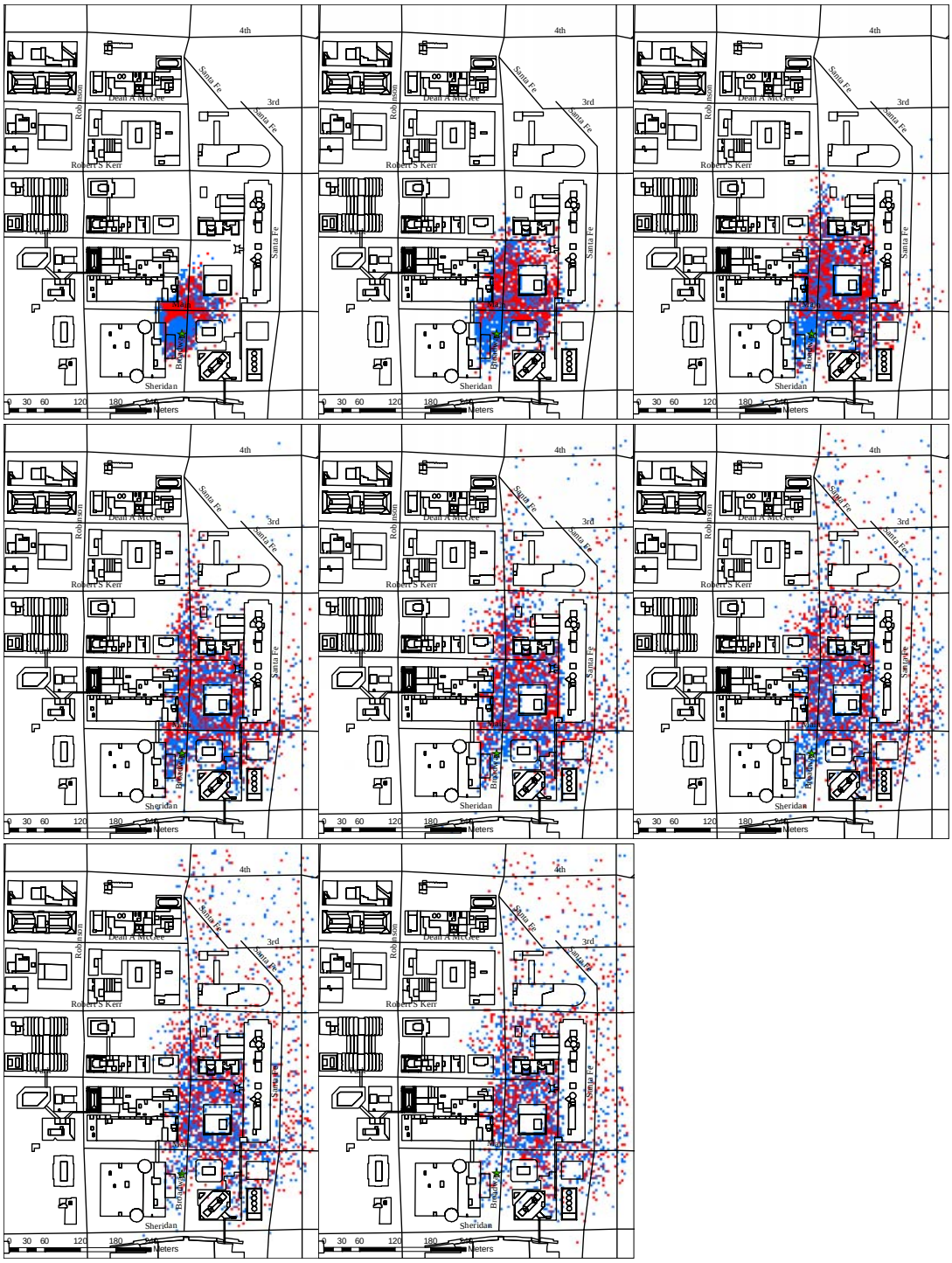


Figure 4.14. Differences in mean concentrations between two groups of simulations at level three (eight to twelve meters), from time step two to nine. Sequence from left to right, top to bottom.

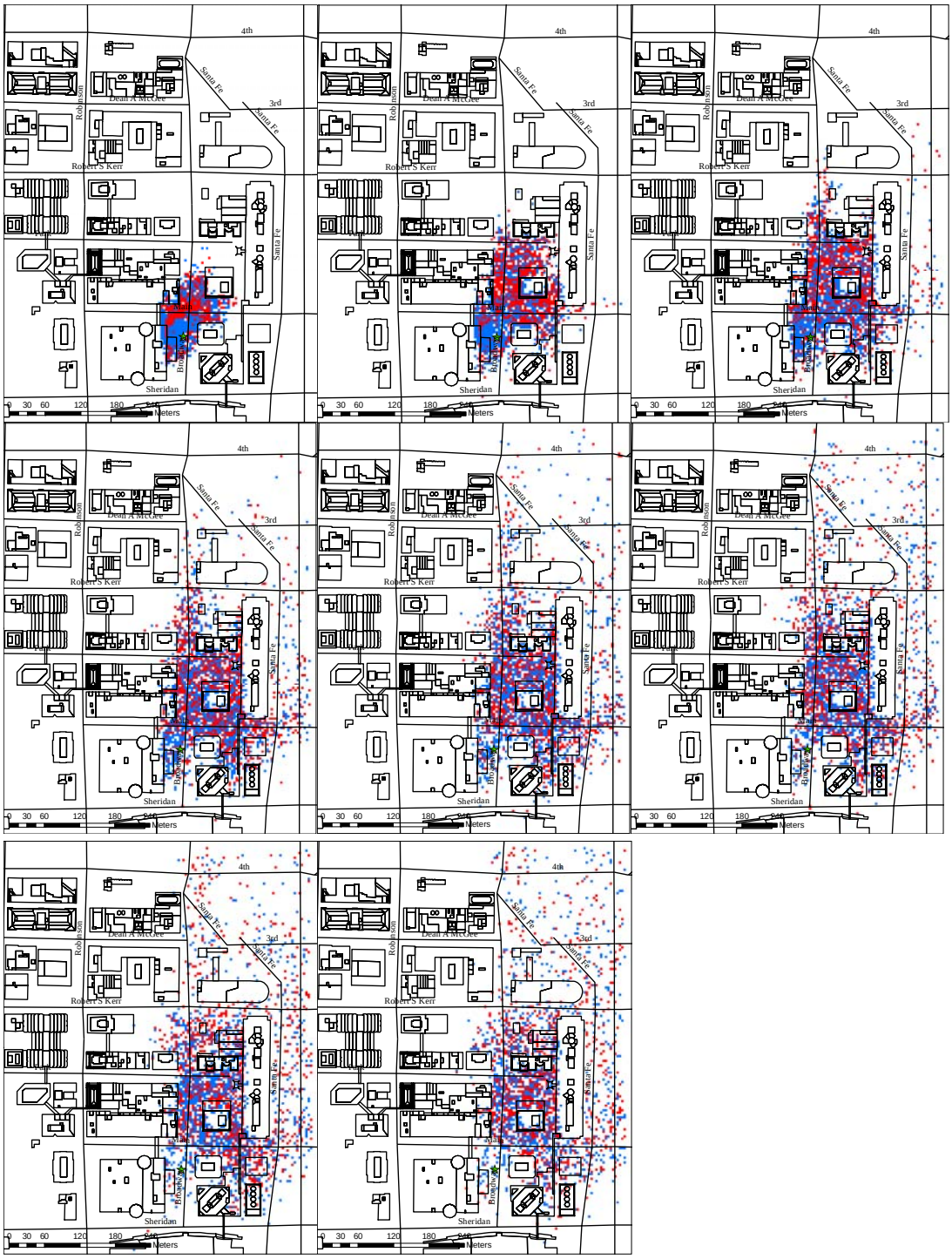


Figure 4.15. Differences in mean concentrations between two groups of simulations at level four (twelve to sixteen meters), from time step two to nine. Sequence from left to right, top to bottom.

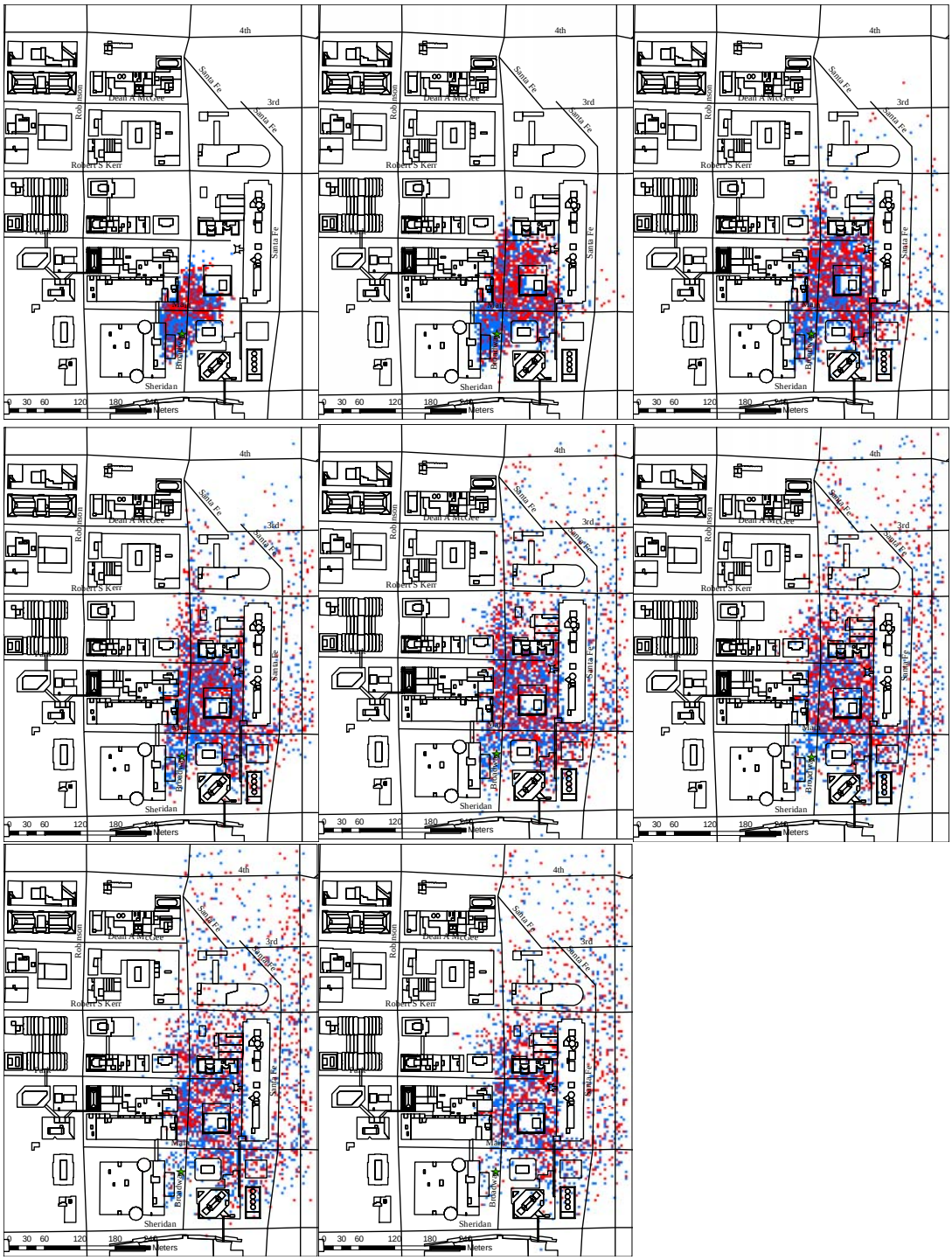


Figure 4.16. Differences in mean concentrations between two groups of simulations at level five (sixteen to twenty meters), from time step two to nine. Sequence from left to right, top to bottom.

Figure 4.17. Locations and time steps with significant differences in mean concentration between two groups of simulation outputs. Top row represents time step two, three and four at ground level. Bottom row represents time step two at height level two and three respectively.

differences were only apparent at the first three time-steps at ground level, and second time-step at level-two (four to eight meters) and second time-step at level-three (eight to twelve meters). Above height level-three, no significant differences appeared.

4.3 Validations of model simulations

Figure 4.18 shows the mean integrated concentrations of model simulations and ground observations across eight sampling sites. Both Group A and Group B over-estimated the concentrations at all the sampling sites, except for site A, where observed concentration was under-estimated by a factor of five. Among the over-estimated sites, only estimates at site D and H fell within a factor of three. Other sampling sites (C, E, F, I and J) were over-estimated exceeding a factor of ten.

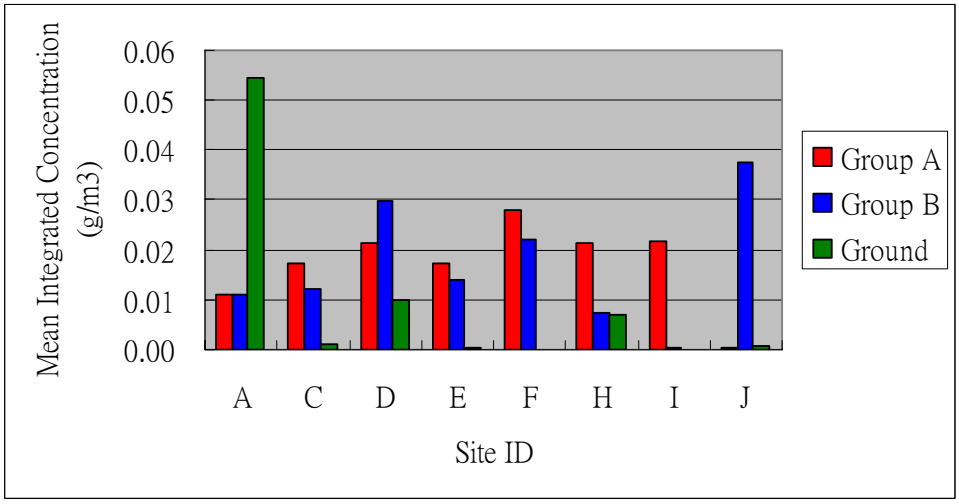


Figure 4.18. Mean integrated concentration at eight sampling sites.

Four common measurements were calculated for both groups of simulations (Table 4.3). A perfect model would have zero for FB and NMSE and one for MG and VG. Both groups of simulation received negative FB which meant over-estimation. A MG of 0.03 from both groups of simulation also

referred an over-prediction of around 30 times between mean estimated concentration and mean observed concentration. Relative large NMSE and VB from both groups of simulation represented the large differences between estimated and observed values, which were mainly contributed by the wrong estimation of a log-normal and relatively scattered distribution of concentration.

Table 4.3. Four common evaluation measurements for two groups of simulations.

	FB	**MG**	**NMSE**	**VG**
Group A	-0.54	0.03	14.17	0.170E+11
Group B	-0.25	0.03	13.07	0.550E+10

To further understand the influence of spatial uncertainty on the results of QUIC, the mean concentrations of model simulations were plotted with ground observation over time (Figure 4.19, also see Appendix II). Green solid lines with diamonds represented mean concentration from the ground observations. Red and blue solid lines with squares and triangles represented mean concentrations from Group A and Group B simulations, respectively. Three patterns appeared in comparison of the concentrations between ground observations and simulations. First, both groups of simulations over-estimated the concentrations at five out of eight sampling sites (i.e. Site C, E, F, I and J). Among those sampling points,

Group B usually generated lower concentrations than Group A and closer to the ground observations. Second, at sampling sites D and H, both groups of simulation over-estimated the concentration at the beginning (i.e. time steps one to four) but then under-estimated the concentrations later (i.e. time step five to eight). There seemed to be a two minute time lag of tracer gas arrival at sites D and H between the QUIC's simulations and ground observations. Again, concentration from Group B was smaller than Group A and closer to the ground observations. Third, both groups of simulations under-estimated the concentrations at site A. The concentration observed at ground level was much higher (almost ten times higher) than the both sets of simulation at the fourth time step.

4.4 The influences of spatial uncertainty

Under the same dispersion scenario settings with an instantaneous release of one kg tracer gas, QUIC simulated a different dispersion pattern with the influence of spatial uncertainty. First, higher mean integrated concentration appeared near the building edges. Second, higher mean concentration appeared near the release point at the first two minutes of dispersion. Similar results persisted up to twenty meters in height. Third, at seven out of eight sampling sites, the mean concentration over time was lower than the mean concentration

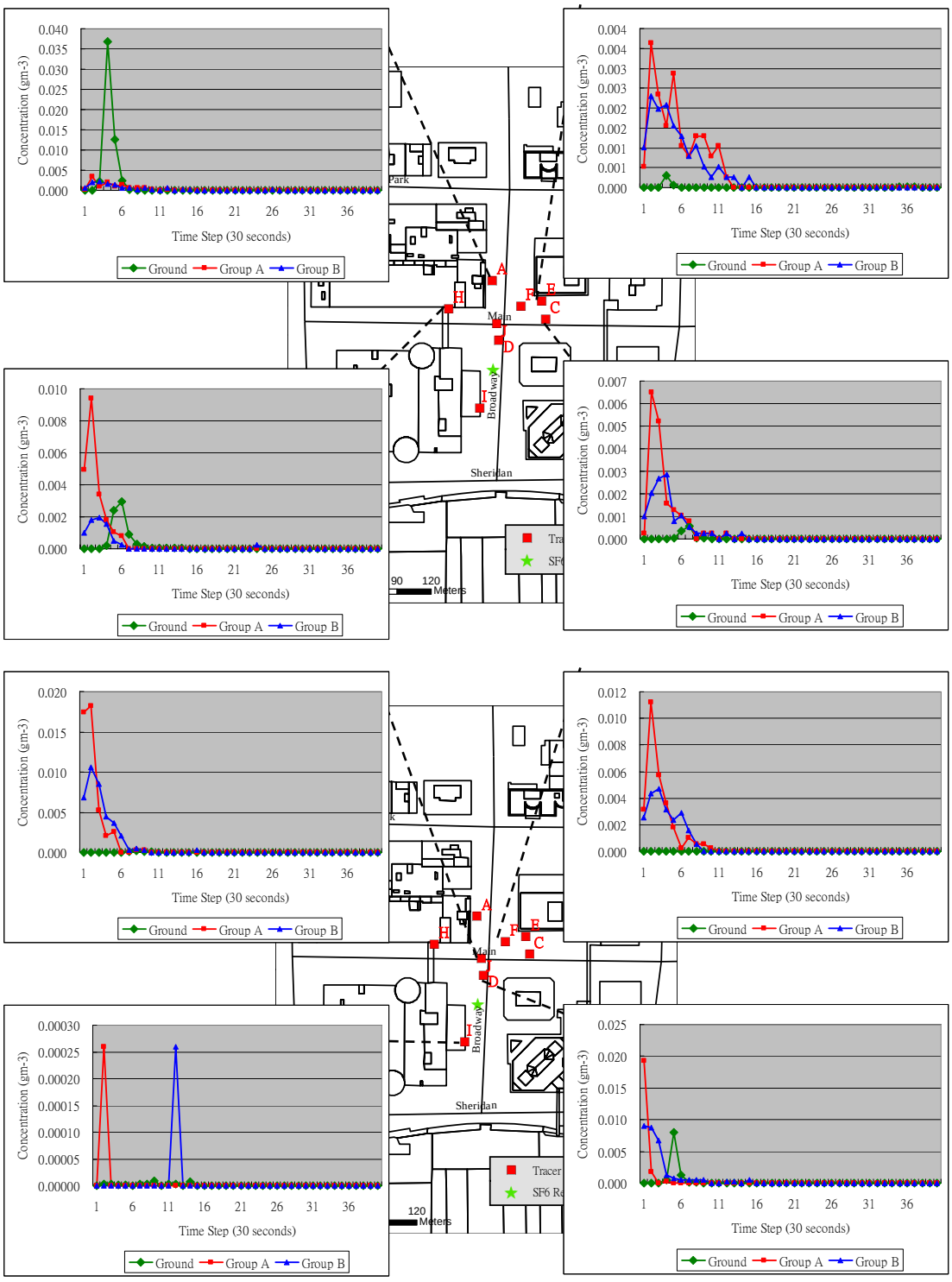


Figure 4.19. Comparison between model simulations and ground observations. Top showing results from site A, C, E & H, bottom showing results from site D, F, I, & J.

simulated without accounting for spatial uncertainty. However, after account for spatial uncertainty, the mean concentration overtime was closer to the ground observation. All three differences could be explained by the shifting of building locations which generated different sizes of urban canyons. Smaller urban canyons near the release point might trap the gas inside and lead to a higher concentration at the beginning of dispersion. After a period of dispersion, the concentration was diluted to the atmosphere and the differences became less significant. Since higher concentrations were found near the release point, lower concentrations were found at the other part of the study area. Therefore, lower concentrations were observed from simulation with spatial uncertainty as the ground observation points were located farther away from the release point.

However, there were limitations in this research and further studies were required on the influence of spatial uncertainty towards UADMs. First, only one dispersion scenario were tested on one atmospheric dispersion model. Additional tests on other scenarios and urban atmospheric dispersion models would be desirable. Continuous release of gas was not tested which might be also important to decision making for emergency managers in real chemical dispersion scenarios. Second, the simulations with spatial uncertainty generated a closer concentration than simulations without spatial uncertainty. The potential

canceling effect of uncertainty sources was intriguing and needed future validation with other scenarios, models, and sources of uncertainty, such as model physics and accuracy of other model parameters. Third, the complexity of urban environments required a sufficient number of ground observations and data modifications before we could run the UADMs with spatial uncertainty. Other approaches that minimized preprocessing cost were desirable to model the influence of spatial uncertainty.

5. Conclusions

Many studies on UADMs have emphasized meteorological and scenario uncertainties. Yet spatial uncertainty should not be ignored. This paper demonstrated the influences of spatial uncertainty towards QUIC dispersion model by conducting two groups of Monte Carlo simulations: Group A with meteorological uncertainties in wind speed and direction and, Group B with meteorological uncertainties and spatial uncertainty in building location. With simulations of an instantaneous release of one kg tracer gas (SF6) in downtown OKC area, this study found that Group B simulations generated a significantly higher mean integrated concentration near the release point when compared to Group A simulations. The significant differences (up to 95 percent confidence

interval) in mean integrated concentration could be found up to three height levels (i.e. up to twelve meters in height). Moreover, by examining the mean concentration over time, Group B simulations generated higher mean concentration near the release point but lower mean concentration away from the release point. However, significant differences (up to 95 percent confidence interval) could only be found during the first two minutes of simulation at the ground level. With field observations from eight real-time tracer gas samplers, both groups of simulations over-estimated the mean concentration except one location. Overall, Group B simulations were closer to the field observations. Although more uncertainties in model input generated simulations closer to field observations, the potential canceling effect of uncertainty sources required further validations on various scenarios and different UADMs. Furthermore, due to the complexity of urban environment, processing cost was high to examine the influence of spatial uncertainty toward UADMs. Development in new approaches that determined whether or not the introduction of spatial uncertainty was necessary to UADMs at certain scales was also desirable.

References

Britter, R. E., and S. R. Hanna. 2003. Flow and dispersion in urban areas. *Annual Review of Fluid Mechanics* 35:469-496.

Brooks, H. E., M. S. Tracton, D. J. Stensrud, G. DiMego, and Z. Toth. 1995. Short-range ensemble forecasting: report for a workshop, 25-27 July 1994. *Bulletin of American Meteorological Society* 76:1617-1624.

Burian, S. J., M. J. Brown, J. Ching, M. L. Cheuk, M. Yuan, W. Han, and A. T. McKinnon. 2004. Urban morphological analysis for mesoscale meteorological and dispersion modeling applications: current issues. In *Fifth Symposium on the Urban Environment, American Meteorological Society*. Vancouver, BC, Canada.

Cullen, A. C. 1999. *Probabilistic techniques in exposure assessment: a handbook for addressing variability and uncertainty in models and inputs*. New York: Plenum Press.

Fox, D. G. 1984. Uncertainty in air quality modeling. *Bulletin of American Meteorological Society* 65:27-36.

Goodchild, M. F. 1993. Data models and data quality: Problems and prospects. In *Environmental Modeling with GIS*, eds. M. F. Goodchild, B. O. Parks and L. T. Steyaert, 94-103. New York: Oxford.

Heuvelink, G. B. M. 1998. *Error Propagation in Environmental Modelling with GIS*. Bristol: Taylor & Francis.

Krzysztofowicz, R. 1998. Probabilistic hydrometeorological forecasts - toward a new era in operational forecasting. *Bulletin of American Meteorological Society* 73 (2):243-251.

Manomaiphiboon, K., and A. G. Russell. 2004. Effects of uncertainties in parameters of a Lagrangian particle model on mean ground-level concentrations under stable conditions. *Atmospheric Environment* 38:5529-5543.

Pardyjak, E., and M. J. Brown. 2002. Fast-response modeling of a two building urban street canyon. Paper read at 4th AMS Symposium on the Urban Environment, May 20-24, at Norfolk, Virginia, USA.

Sax, T., and V. Isakov. 2003. A case study for assessing uncertainty in local-scale regulatory air quality modeling applications. *Atmospheric Environment* 37:3481-3489.

Sullivan, D. A., M. T. Holdsworth, and D. J. Hlinka. 2004. Monte Carlo-based dispersion modeling of off-gassing releases from the fumigant metam-sodium for determining distances to exposure endpoints. *Atmospheric Environment* 38:2471-2481.

USEPA. 1992. Guidelines for Exposure Assessment.

Williams, M. D., M. J. Brown, B. Singh, and D. Boswell. 2004. *QUIC-PLUME Theory Guide.* LA-UR-04-0561. Los Alamos National Laboratory.

Yee, Y., M. Bustillos, S. Chang, R. Cionco, and E. Creegan. 2004. Wind and turbulence observations in joint urban 2003. In *Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone*. Seattle, WA.

Yegnan, A., D. G. Williamson, and A. J. Graettinger. 2002. Uncertainty analysis in air dispersion modeling. *Environmental Modelling & Software* 17:639-649.

Chapter 5 : Conclusion

1. Introduction

Knowing the output uncertainty in environmental models is crucial for decision-makers because important policies are made with the aid of environmental models. Lack of complete knowledge in natural phenomena, errors in input data and inadequate setting of model parameters are examples that contribute to the uncertainty of the model results. This dissertation traces back the sources of uncertainty in the geo-spatial data and summarizes how the uncertainty of geo-spatial data influences the results of an urban atmospheric dispersion model (UADM), using LiDAR-derived building data and Quick Urban and Industrial Complex (QUIC) dispersion model as an example.

Advances in remote sensing technologies have acquired geo-spatial data at high resolution and multi-dimensions. For example, light detection and ranging (LiDAR) technique allows users to capture detailed terrain elevation information up to sub-meter resolution. Combined with feature extraction algorithms, surface features such as buildings elevation, tree branches, sand dunes and even light pole can be captured by LiDAR. Previous studies show that LiDAR data also contributes uncertainty to environmental models because of the limitations in LiDAR technique and feature extraction algorithms (Burian et al., 2004). A

comprehensive understanding of LiDAR data accuracy helps identify potential source of uncertainty before inputting the data into the environmental models.

Another source of uncertainty arises during the data transference between geographic information systems (GIS) and environmental models. Unique data models and formats in environmental models require data conversion. Data conversion processes such as changing in scale and re-sampling of data alter the data location and further introduce uncertainty to environmental models. Previous studies have been focused on the shifting of objects location in a two-dimensional surface (Huising and Pereira, 1998; Ahokas, Kaartinen, and Hyyppa, 2003; Alharthy, Bethel, and Mikhail, 2004; Hopkinson et al., 2001). However, as three-dimensional data are now common, understanding the uncertainty in the third-dimension is also necessary.

A final source of uncertainty relates to model parameters. Settings of model parameters such as meteorological factors and equations variables can be uncertain due to missing information or natural variability. Using UADM as an example, model parameters such as wind profile parameters and pollutant release information can greatly influence the results of the models (Williams, Brown, and Pardyjak, 2002). However, most studies have focused on examining uncertainties from model parameters without acknowledging the uncertainty from geo-spatial

data (Sax and Isakov, 2003; Sullivan, Holdsworth, and Hlinka, 2004; Yegnan, Williamson, and Graettinger, 2002; Manomaiphiboon and Russell, 2004). To fully understand the influence of spatial uncertainty toward the results of UADM, research on comparing influence of uncertainty between model parameters and geo-spatial data is necessary.

This chapter first summarizes the findings from: 1) spatial uncertainty from LiDAR-derived building data, 2) spatial uncertainty from the linkages between GIS and QUIC and, 3) influence of spatial uncertainty toward the QUIC dispersion model, and then the chapter concludes with discussions on research contributions and suggestions for future research.

2. Summary of Findings

2.1 Spatial uncertainty from LiDAR-derived building data

Three main sources of uncertainty were accounted for using LiDAR data to derive a three-dimensional building model. Environmental settings in urban area, feature extraction algorithm and manual digitizing contribute to the sources of uncertainty.

By comparing the LiDAR data with field observation, this study found that environmental settings such as complex building structure, sky-rise buildings,

different surface materials and frequent changing of the environments reduced the accuracy in capturing three-dimensional data. Building gaps smaller than six meters prevented laser pulse from reaching to the ground and hence produce a higher elevation than actual elevation (Figure 2.9, p.51). Similarly, high-rise buildings could block laser pulses and cause missing data in surrounding area, resulting in significant data errors (Figure 2.10, p.52). Buildings with glassy surface also generated signal noises to surrounding area and caused a rougher surface (Figure 2.11, p.53). Vegetation covers near the buildings could hinder determination of building boundary. Sometimes, trees could grow over or on top of the building and mask the actual building heights (Figure 2.12, p.55). Construction sites were in transition and, therefore, could result in significant differences in spatial configuration and height in an urban environment. By the time that a three-dimensional model was established, the construction might be completed and hence, the 3D model became outdated.

Raw LiDAR data consisted of massive amount of points with x, y and z co-ordinates, which required feature extraction algorithm to identify useful information. Assumptions built in any feature extraction algorithm introduced additional spatial uncertainty. For example, building extraction algorithms might assume that building is rectangular objects, and consequently many smaller

rectangular objects in urban area could be mis-classified as buildings. Cargo containers, trees in the park, cover parking canopy and bridges were commonly mis-classified objects (Figure 2.16, p.59). Moreover, the feature extraction algorithm only extracted building's footprints, detailed building outlines required laborious inspections and manual adjustments.

While manual adjustments were necessary to finalize the construction of 3D building models, the procedure was also a source of uncertainty. Our human subject test showed that discrepancies among digitizing operators might lead to significant uncertainty in building locations in two ways (Figure 2.18, p.61): First, when buildings were too close to each other with separation less than six meters, and second where there were vegetation outgrowth near buildings.

Auxiliary remote sensing data, such as aerial photos, could be useful to reduce spatial uncertainty in LiDAR-derived 3D building models to supplement information about geography. Features in a complex urban environment could be difficult to discern with LiDAR data alone. The dissertation research offered guidelines to recognize spatial uncertainty and area where significant uncertainty was most likely and needed detailed analysis in development 3D building model with LiDAR data under the consideration of time and cost limitations.

2.2 Spatial uncertainty from the linkages between GIS and QUIC

More often than not, data transformation was necessary to ingest geospatial data to environmental models. In this research, the data transformation between ArcGIS and QUIC resulted in discrepancies of both location and shape. The LiDAR data were converted using Visual Basic for Applications (VBA) inside ArcGIS. In ArcGIS, LiDAR-derived building data were represented as polygons with irregular shapes, composed of lines that made up of multiple points with x and y co-ordinates. In QUIC, building data were stored as either rectangle or circle, which were composed of building ID, width, length, height, base-height, and location in terms of middle x and y co-ordinates. Therefore, the main source of spatial uncertainty came from the conversion from irregular polygons to rectangular shapes. Other environmental models might require different forms of transformation, but shape and location were the two common elements to which data transformation introduce uncertainty.

According to the QUIC model needs, our algorithm converted the building data from irregular polygons to rectangular shape, utilizing two types of vector-to-raster conversion in ArcGIS, dominant unit method and central position method. The central position method assigned values to the grids by taking the polygons that fell at the center of the grids. The dominant unit method assigned

values to the grids by considering the polygons that shared the dominant unit of the grid. However, two issues arose during the conversion. First, the number of records increased dramatically (i.e. 6.3 times) after conversion with the finest resolution. Second, location, area and volume of the building data were changed after conversion. First issue could be solved by reducing the study area or increasing the resolution. As resolution became coarser, the number of records decreased from thirteen percent to fifty percent. In the study area, the number of buildings became smaller than that of original records when the resolution is seven meters or coarser.

Second issue related to uncertainty in spatial location. The location of building shifted about half a unit of resolution after data conversion (Figure 3.6, p.93). As a result, there were also changes in building footprint area and building volume. From one meter to twelve meter resolutions, the total building footprint area changed within one percent for both methods of conversion. Footprint area omitted and committed increased from one percent to fourteen percent and fifteen percent for central position method, while the area omitted and committed increased from one percent to eighteen percent and eleven percent with dominant unit method. Total building volume fluctuated around three percent and nine percent for central position method and dominant unit method respectively (Figure

3.16, p.109). The percentage of volume omitted and committed were similar for central position method, both increased from one percent to twenty-one percent across twelve resolutions. However, the percentage of volume omitted and committed for dominant unit method changed from one percent to twenty-six percent and nineteen percent, respectively. Most of the changes occurred at the edge of buildings except two conditions. First, linear features such as skywalk and buildings smaller than half a unit of the resolution might disappear after data conversion with dominant unit method but might not disappear using central position method. Second, building heights were altered after data conversion and the effects happened across the entire study area for any building which height could not be evenly divided by the chosen vertical resolution (Figure 3.18, p.111).

2.3 Influences of spatial uncertainty toward the QUIC dispersion model

Based on the spatial uncertainty from LiDAR-derived building data and data conversion, Monte-Carlo simulations were used to examine their influences toward QUIC dispersion model. Two groups of simulations were generated. Group A simulations were perturbed with meteorological uncertainty that were composed of variations in wind speed and wind direction. Group B simulations were perturbed with both meteorological and spatial uncertainty. The results of

the simulations were summarized in terms of integrated concentration, concentration over time and concentration versus ground observations.

The main difference between Group A and Group B simulations could be found near to the building edges (Figure 4.9, p.140); whereas, the integrated concentrations from Group B simulations showed further dispersion across the downwind side of the study area and they were also higher around the release point than Group A. Similar results could be observed at different height levels. However, significant differences between the two groups of simulation only appeared near the building edges.

By comparing concentration estimates over time, apparent differences were found during the first two minutes of simulation (i.e. four simulation time steps) (Figure 4.12, p.145). The release gas generally dispersed from southwest to northeast. With the influence of spatial uncertainty, higher concentrations could be observed near the release point, while lower concentrations could be found at the street block northeast to the release point. Similar pattern could also be observed at height level two to level four (i.e. eight to sixteen meters).

Eight sampling sites near the release points were chosen for model evaluations (Figure 4.19, p.155). Overall, both groups of simulations over-estimated the integrated concentration by a factor of three or above at seven

out of eight locations. Sampling sites farther away from the release point received larger differences between model prediction and ground observation. By comparing the mean concentration over time, both groups of simulations overestimated the concentration mainly at the first-two minutes. At two locations, both groups of simulations over-estimated the concentration at first and then under-estimated it later. Simulations with spatial uncertainty usually generated lower concentrations at the sampling points when compared to simulations without spatial uncertainty. However, the predictions were closer to the ground observation.

3. Concluding Remarks

Uncertainty is inevitable and ubiquitous in GIS applications, propagating through how we conceptualize the earth's surface and objects using computer models, associates with the data through data collections, manipulations, analyses modeling, and interpretations. A good understanding of the degree to which spatial uncertainty influence the decision-making processes is critical to proper use of geospatial data and modeling outcomes. This dissertation research has shown a proof in the research hypothesis that the uncertainty embedded in the

LiDAR-derived building model and the data conversion process, has influences on the results of an UADM.

Although improving remote sensing technology delivers richer sources and more accurate spatial data, spatial uncertainty is still involved in every stages of the data life cycle. Understanding the influence of spatial uncertainty in each stages of the data life cycle provides a clearer scene for the data applications and hence reduces the chance of mis-leading results from environmental models. Future research could further examine the influence of spatial uncertainty, such as under what scenario or what scale of study, that may change the decisions-making processes.

References

Ahokas, E., H. Kaartinen, and J. Hyyppa. 2003. A quality assessment of airborne laser scanner data. Paper read at ISPRS Working Group 3 Workshop, 3-D reconstruction from airborne laserscanner and InSAR data, 8-10 October 2003, at Dersden, Germany.

Alharthy, A., J. Bethel, and E. M. Mikhail. 2004. Analysis and accuracy assessment of airborne laserscanning system. Paper read at XXth ISPRS Congress, Commission II, 12-23 July 2004, at Istanbul, Turkey.

Burian, S. J., M. J. Brown, J. Ching, M. L. Cheuk, M. Yuan, W. Han, and A. T. McKinnon. 2004. Urban norphological analysis for mesoscale meteorological and dispersion modeling applications: current issues. In *Fifth Symposium on the Urban Environment, American Meteorological Society*. Vancouver, BC, Canada.

Hopkinson, C., L. E. Chasmer, G. Zsigovics, I. F. Creed, M. Sitar, P. Treitz, and V. M. Rober. 2001. Errors in LiDAR ground elevation and wetland vegetation height estimates. Paper read at International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, at Bangkok, Thailand.

Huising, E. J., and L. M. G. Pereira. 1998. Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry & Remote Sensing* 53:245-261.

Manomaiphiboon, K., and A. G. Russell. 2004. Effects of uncertainties in parameters of a Lagrangian particle model on mean ground-level concentrations under stable conditions. *Atmospheric Environment* 38:5529-5543.

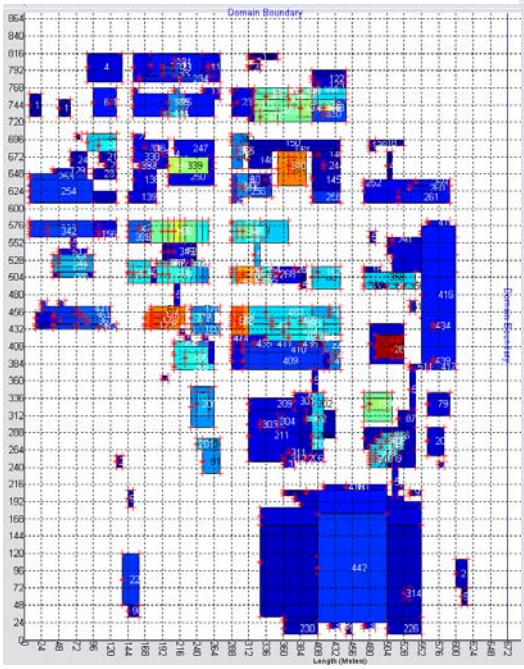
Sax, T., and V. Isakov. 2003. A case study for assessing uncertainty in local-scale regulatory air quality modeling applications. *Atmospheric Environment* 37:3481-3489.

Sullivan, D. A., M. T. Holdsworth, and D. J. Hlinka. 2004. Monte Carlo-based dispersion modeling of off-gassing releases from the fumigant metam-sodium for determining distances to exposure endpoints. *Atmospheric Environment* 38:2471-2481.

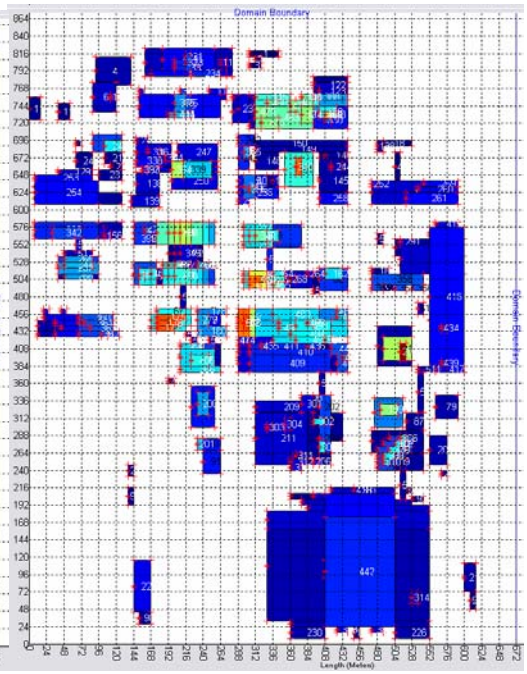
Williams, M., M. J. Brown, and E. R. Pardyjak. 2002. Development of a dispersion model for flow around buildings. Paper read at 4th AMS Symposium on the Urban Environment, at Norfolk, Virginia, USA.

Yegnan, A., D. G. Williamson, and A. J. Graettinger. 2002. Uncertainty analysis in air dispersion modeling. *Environmental Modelling & Software* 17:639-649.

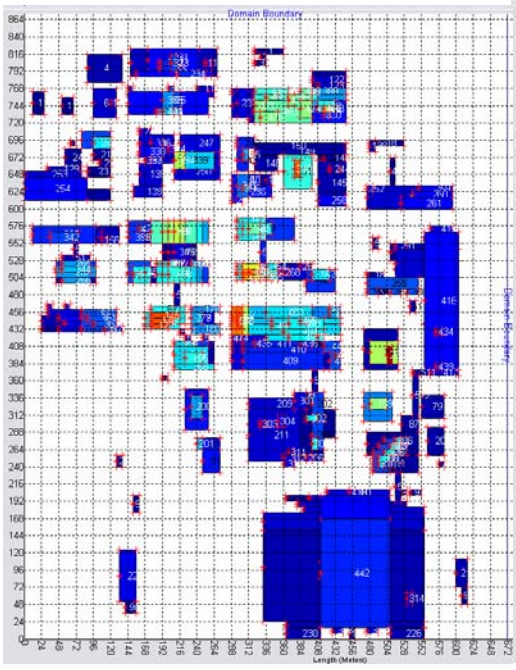
Appendix I – Building perturbations in Group B simulation



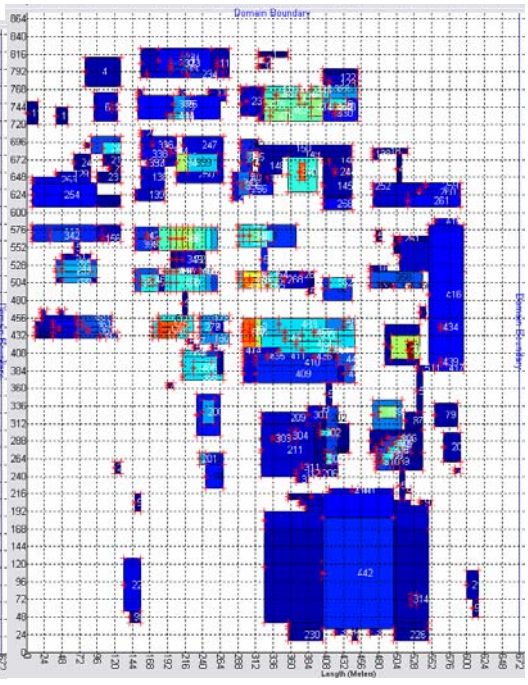
Original



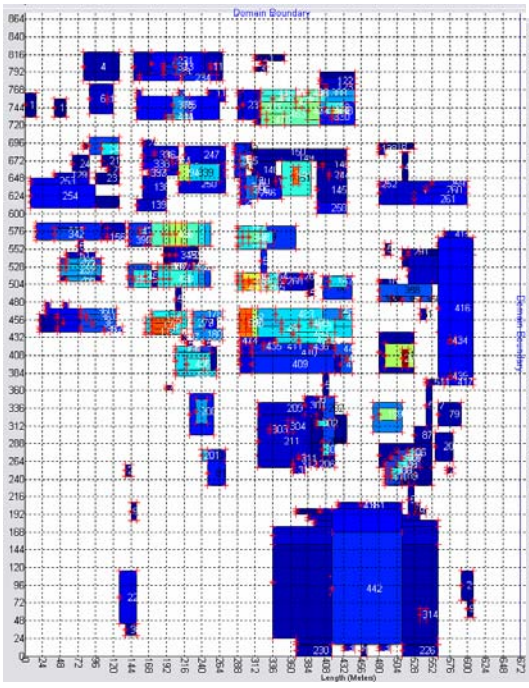
B01



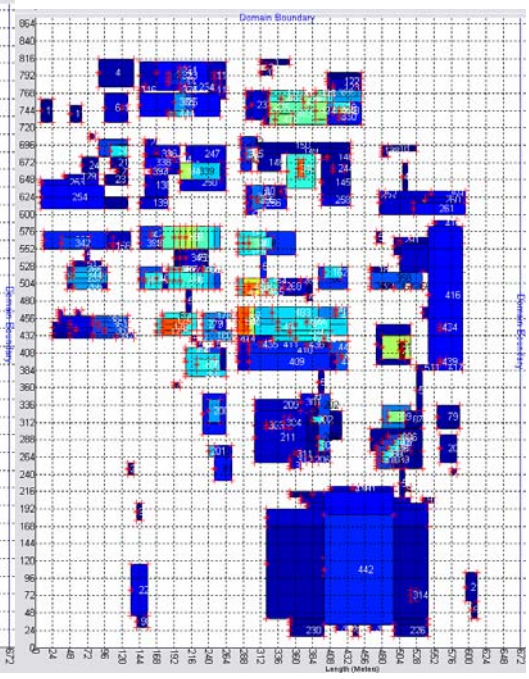
B02



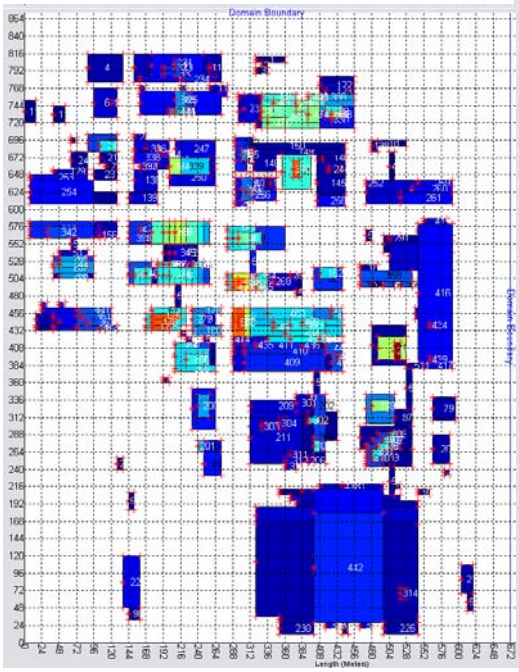
B03



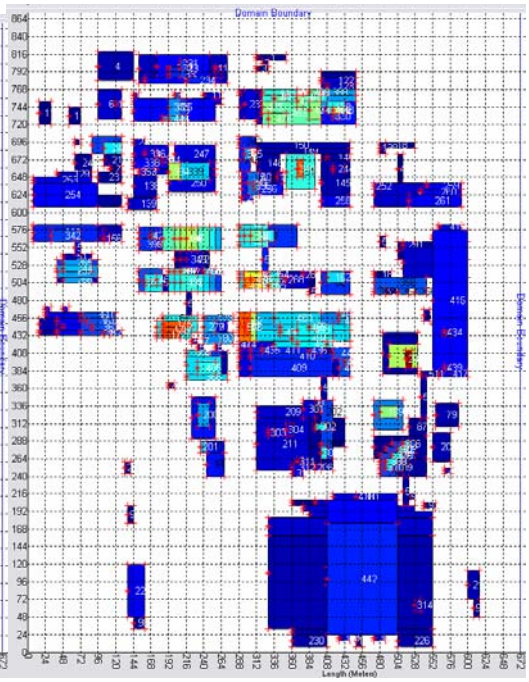
B04



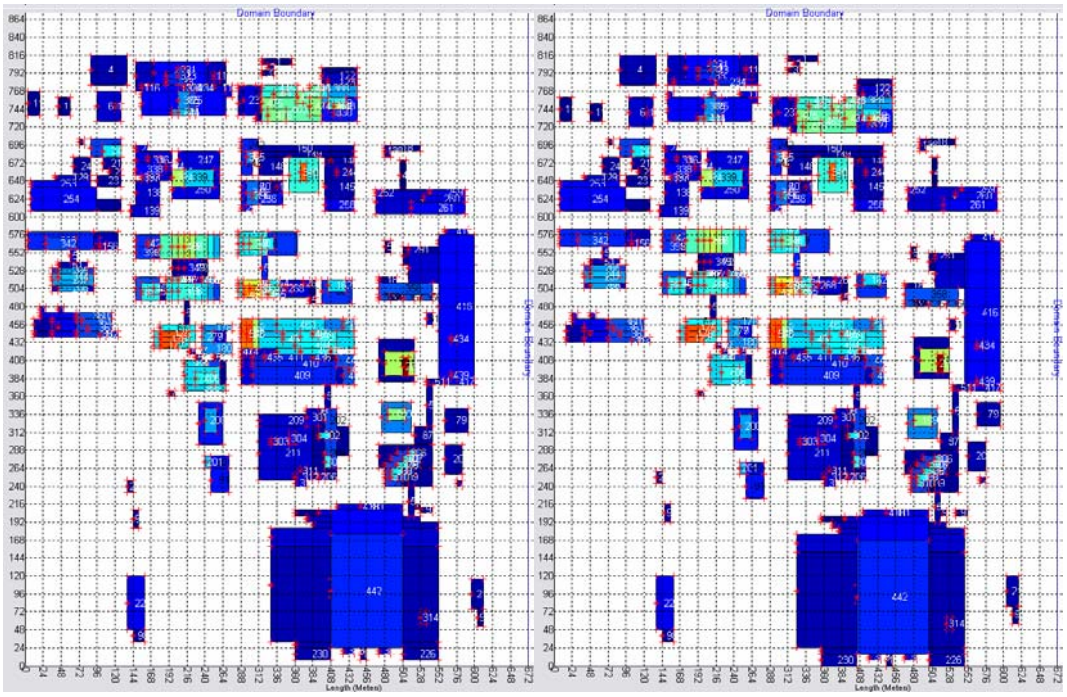
B05



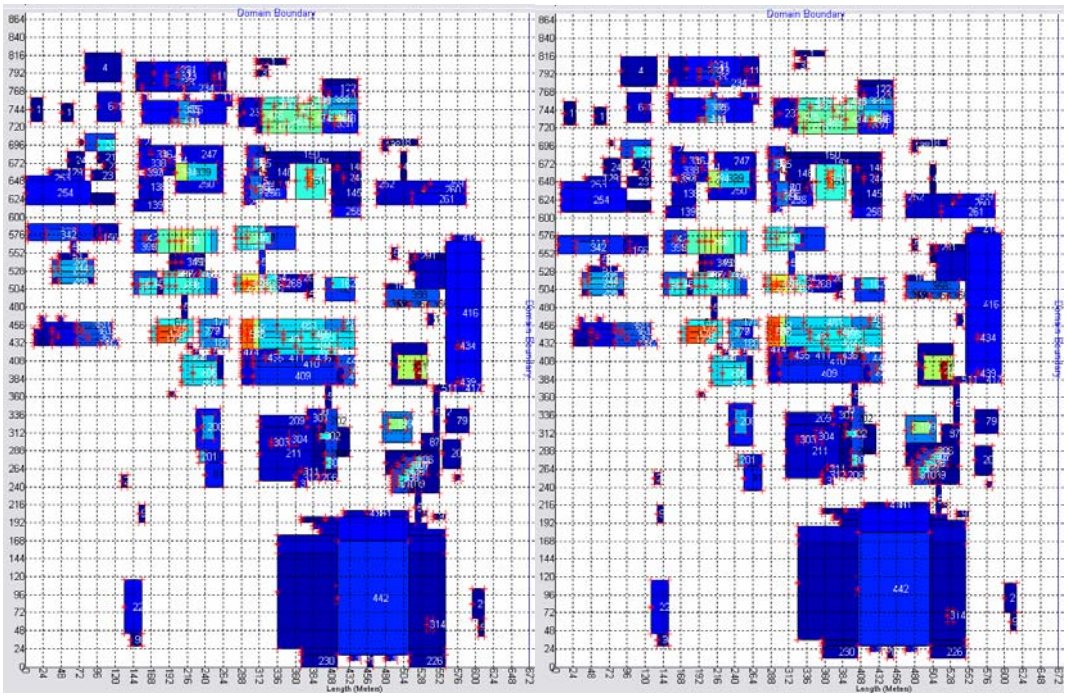
B06



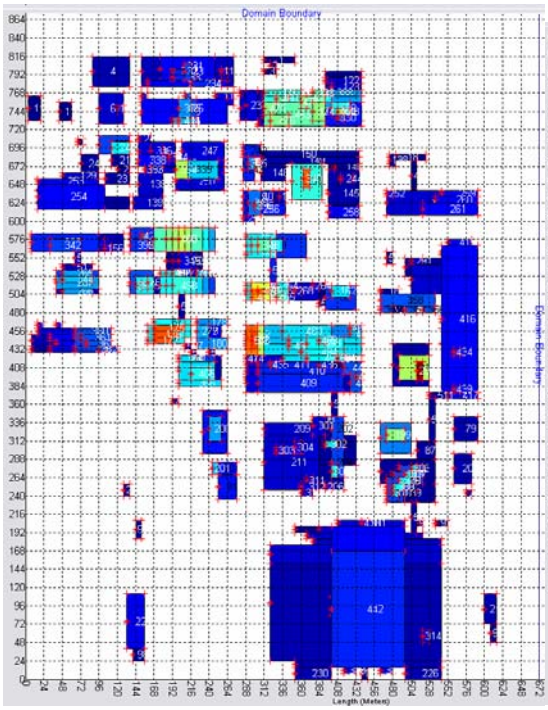
B07



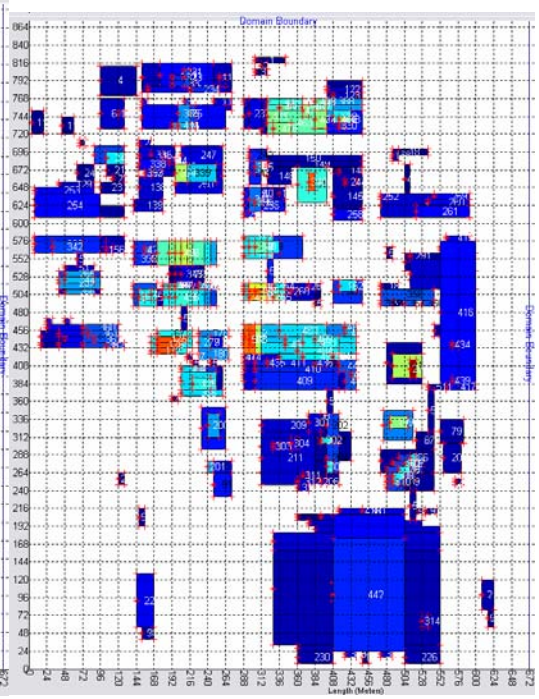
B08 B09



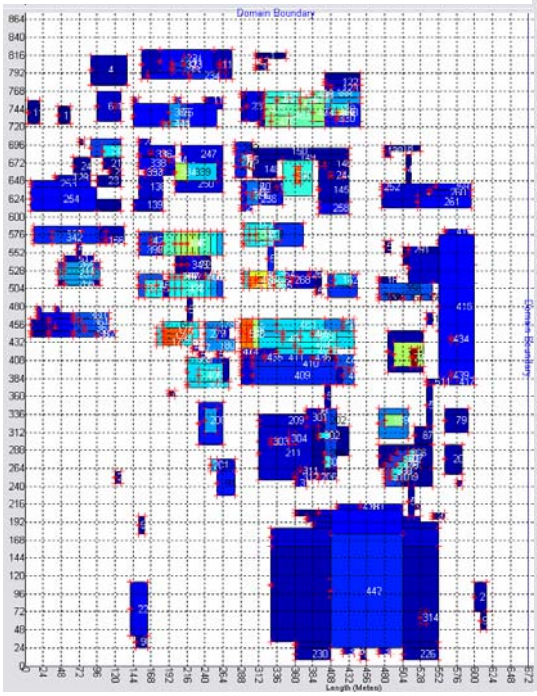
B10 B11



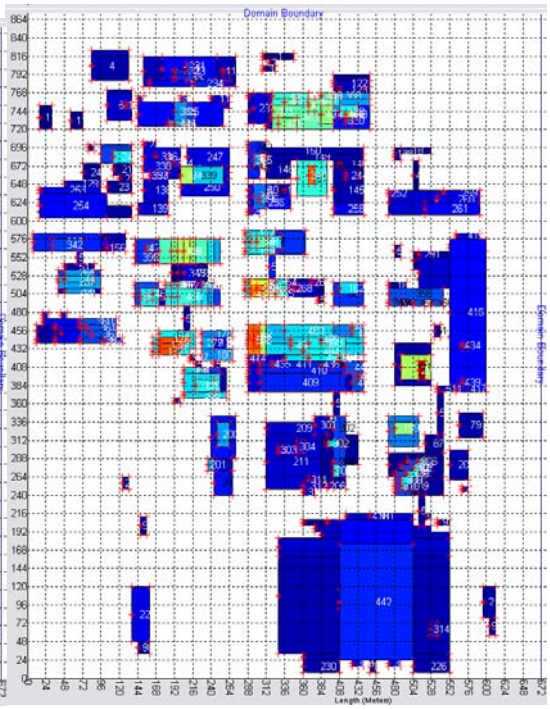
B12



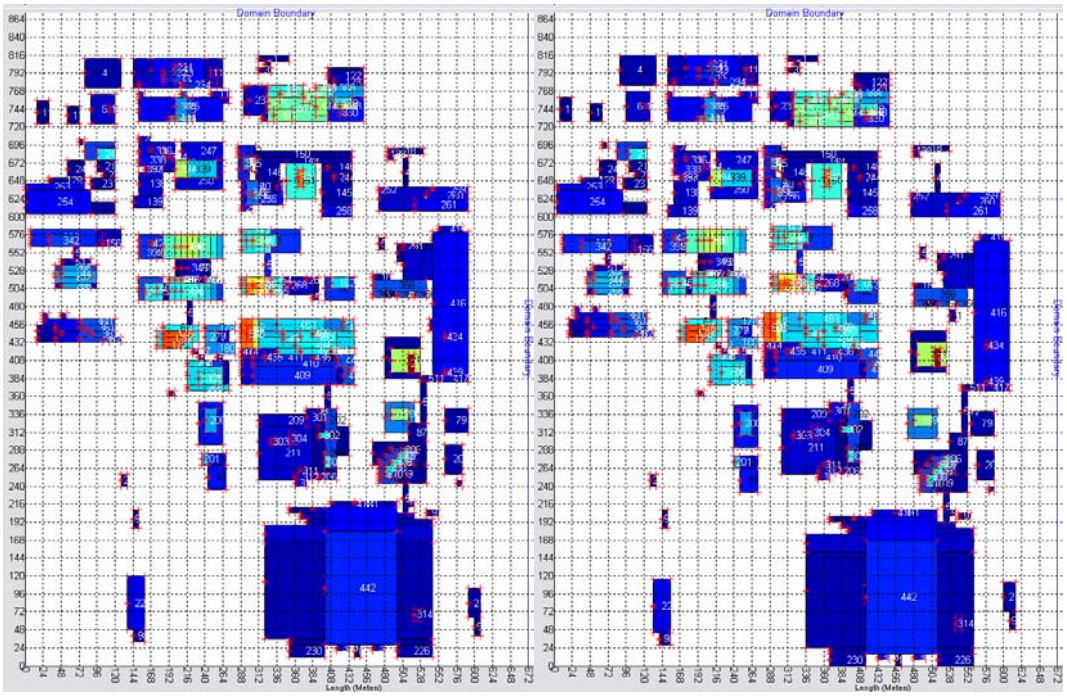
B13



B14

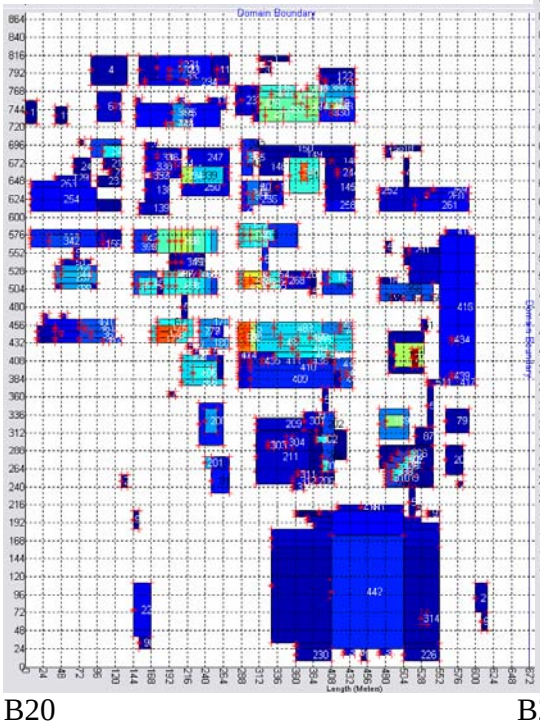


B15

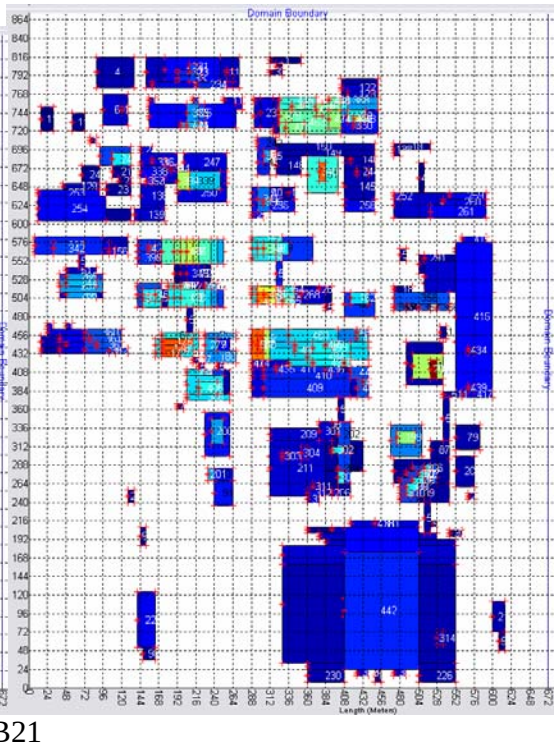


B17

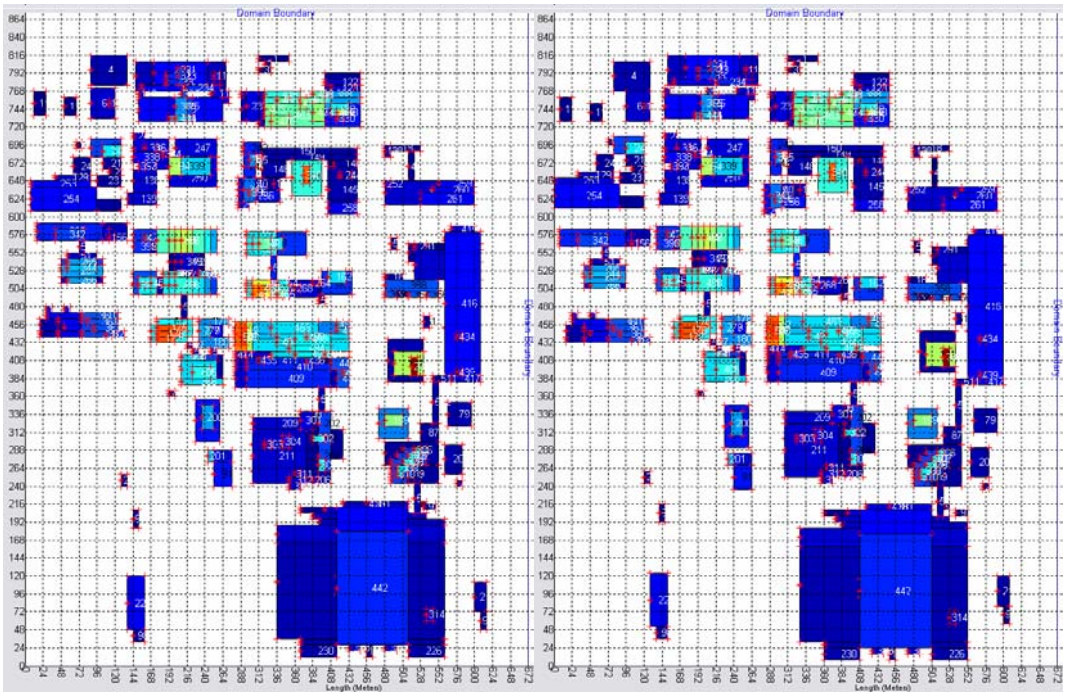
B18



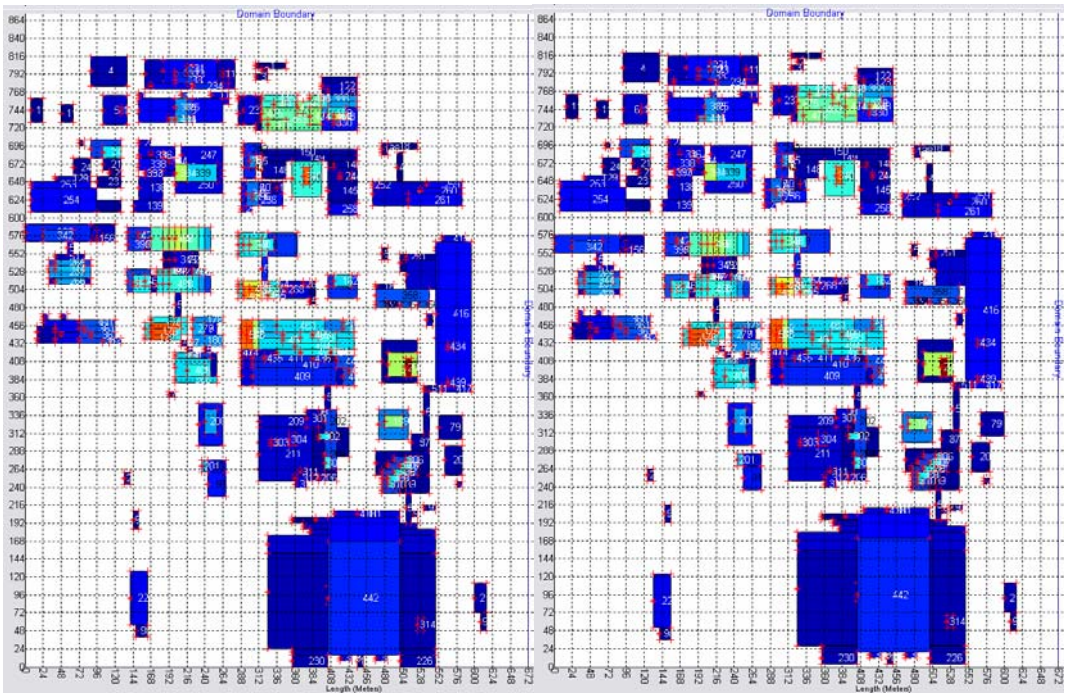
B20



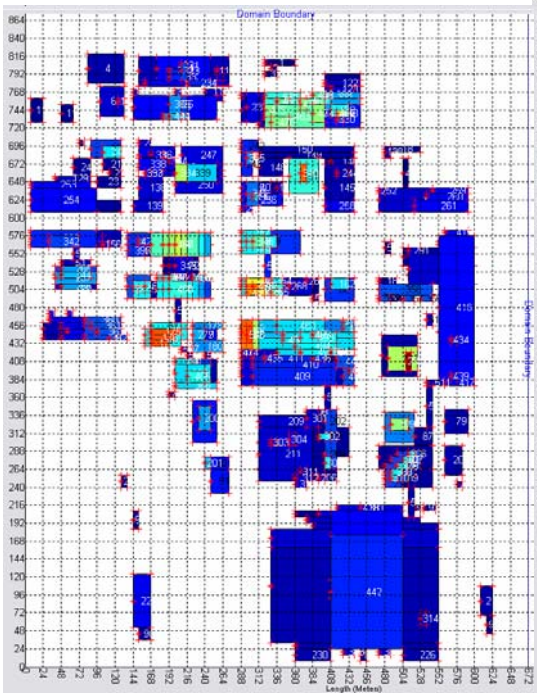
B21



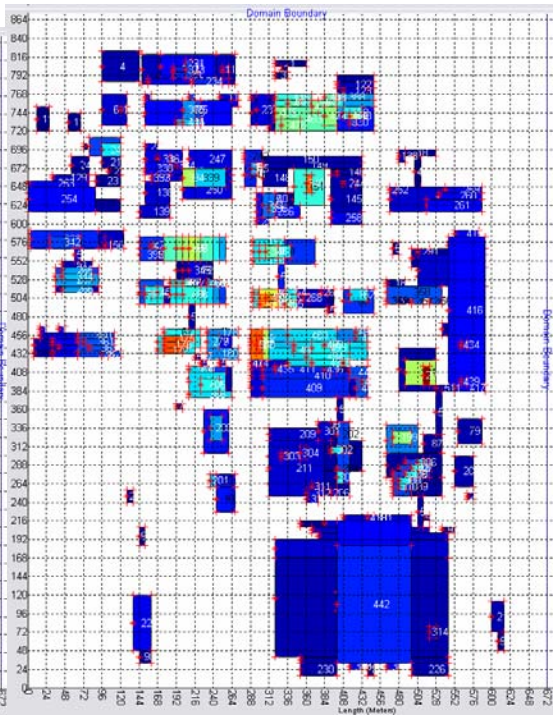
B22 B23



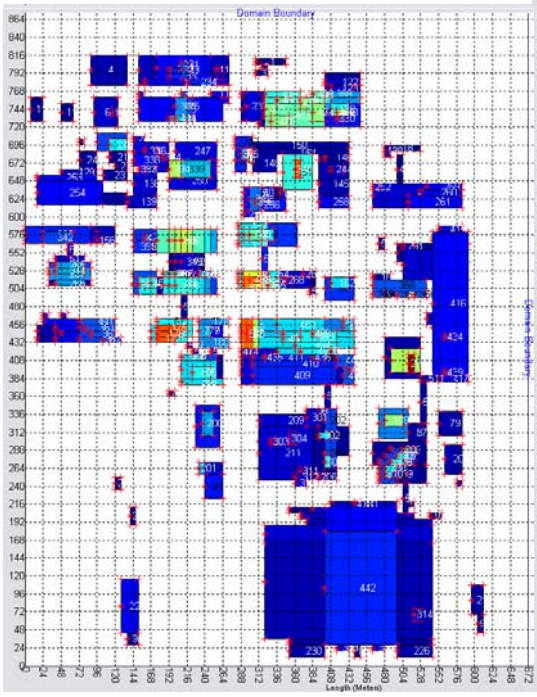
B24 B25



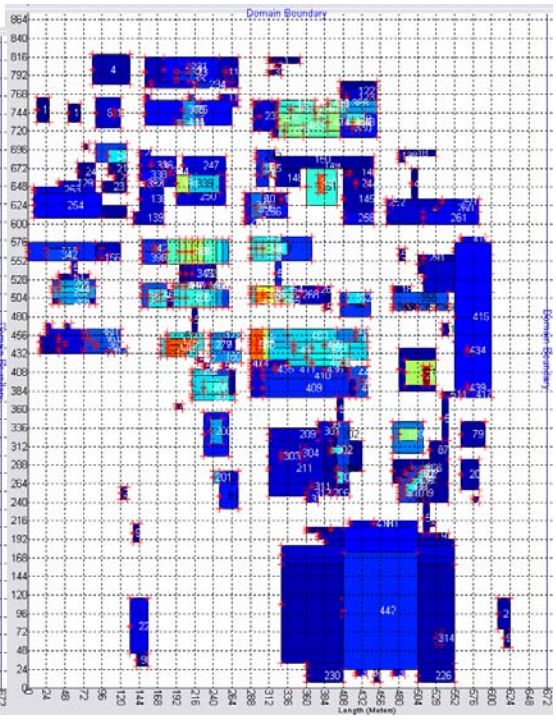
B26



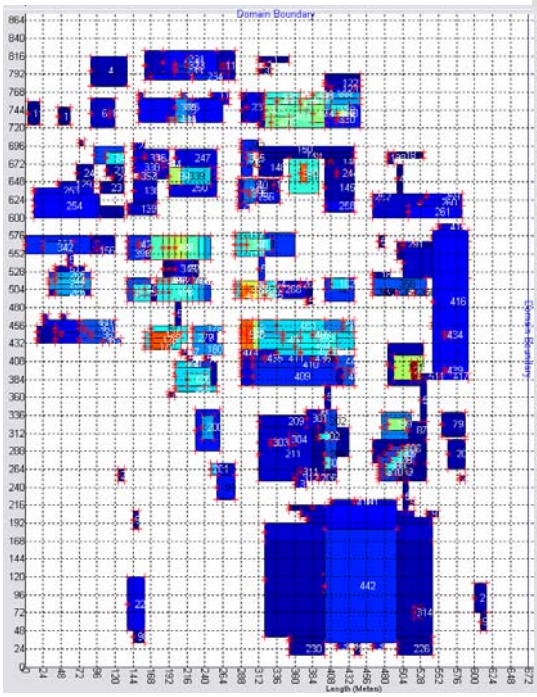
B27



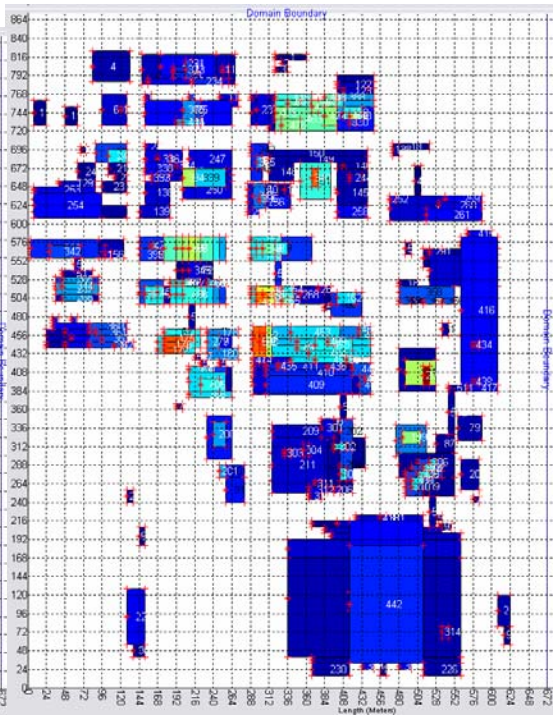
B28



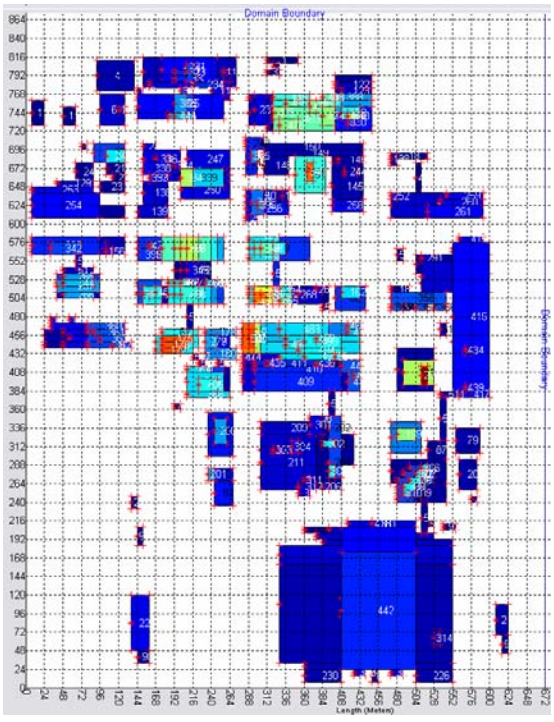
B29


Domain Boundary
Length (Meters)

B30

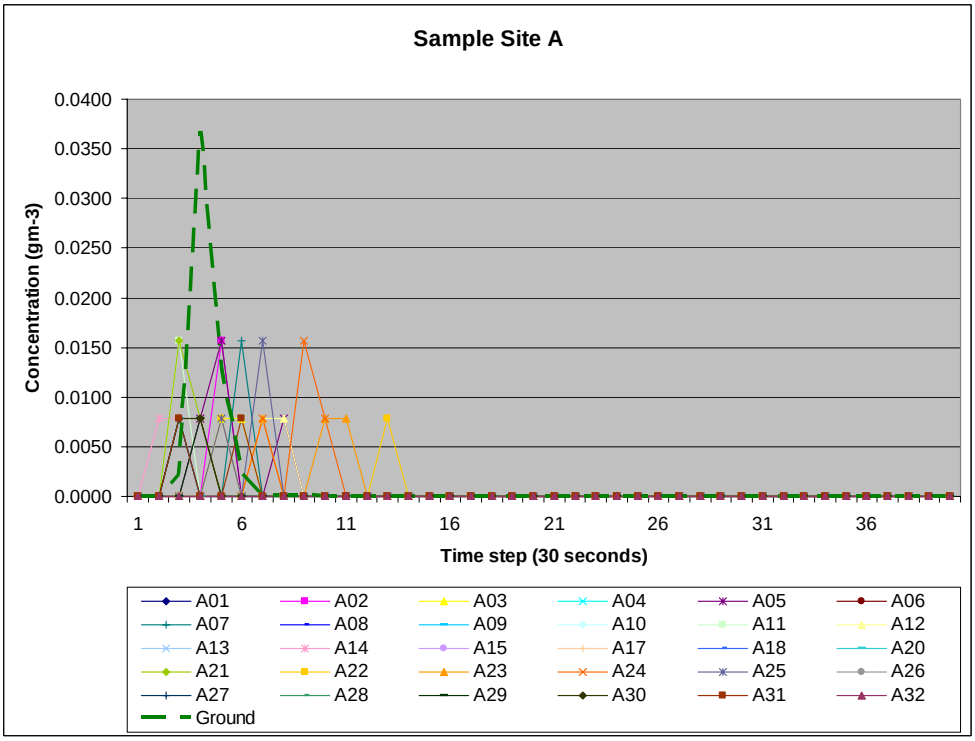

Domain Boundary
Length (Meters)

B31

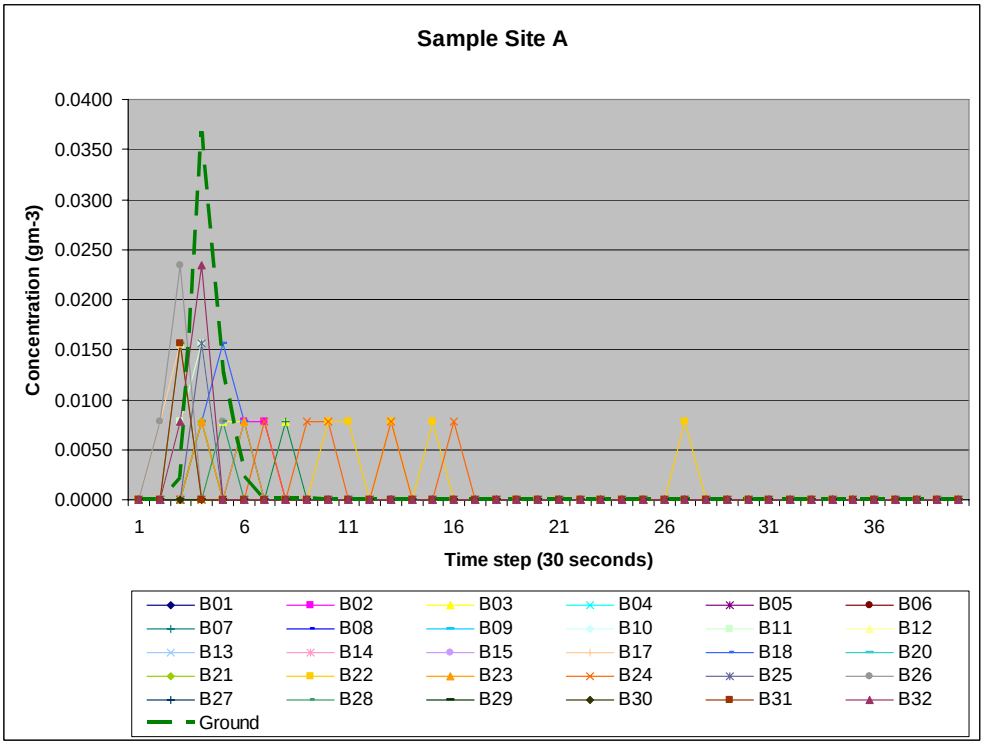

Domain Boundary
Length (Meters)

B32

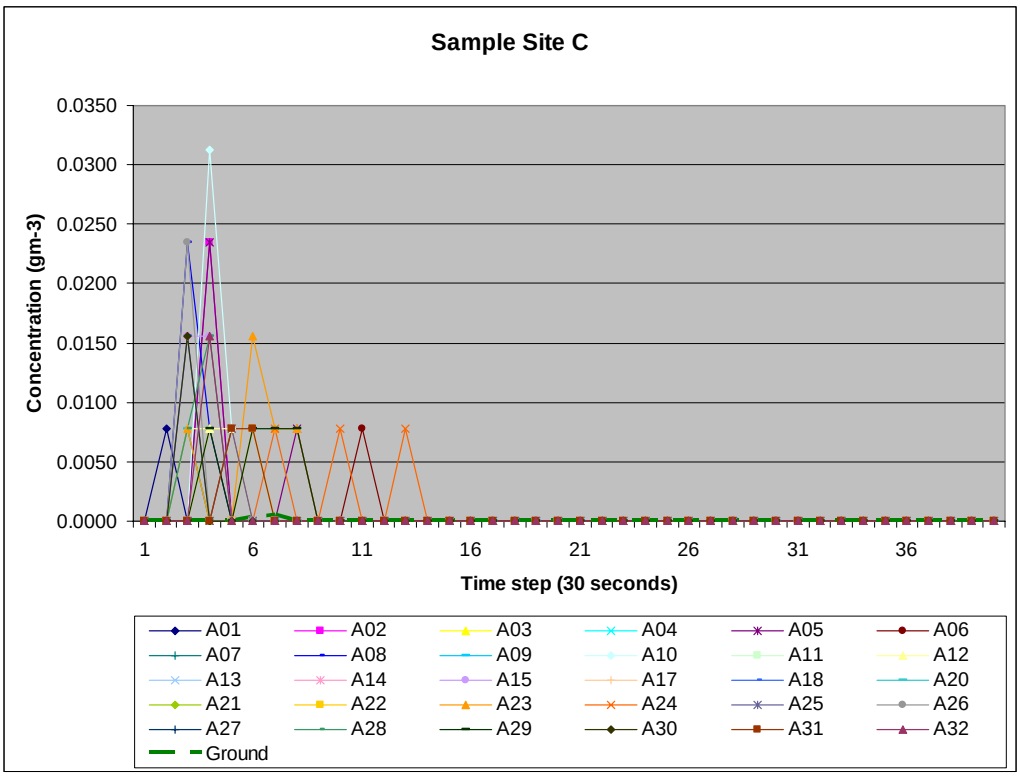
Appendix II – All simulation results at eight sampling sites.



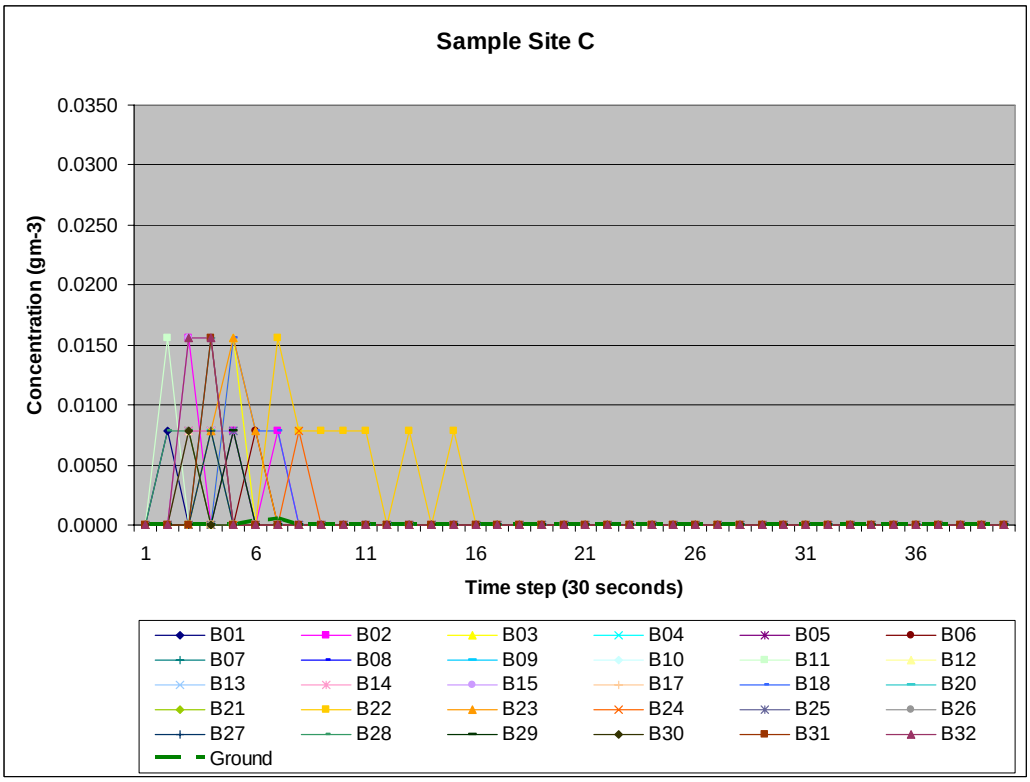
Group A simulations



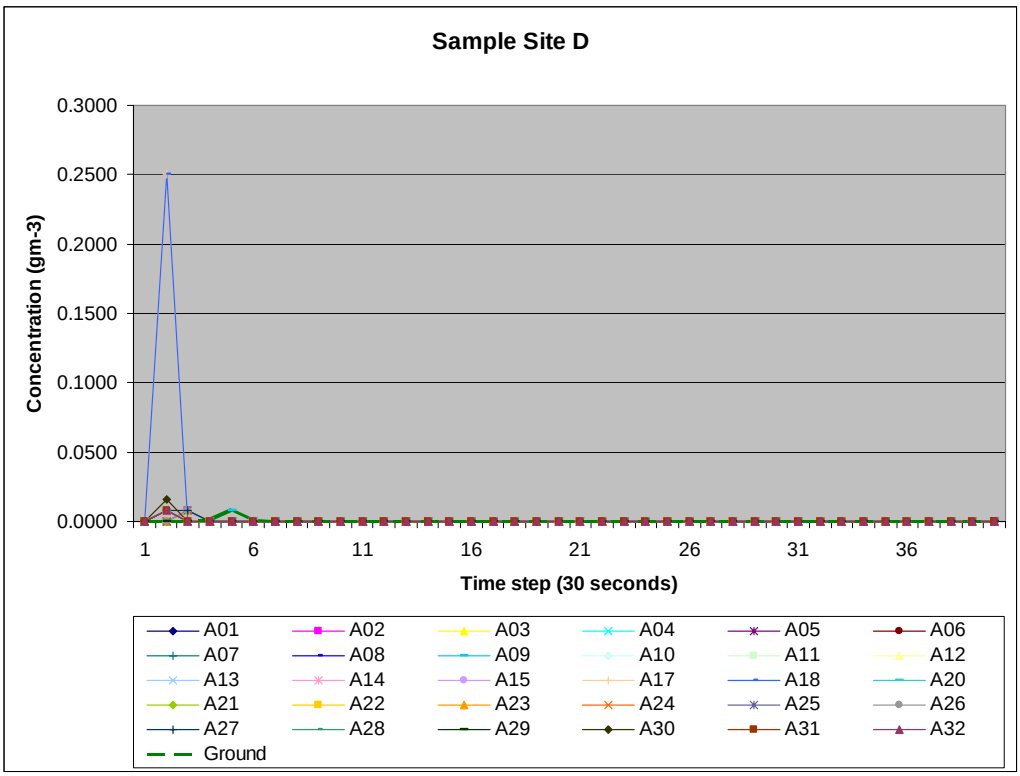
Group B simulations



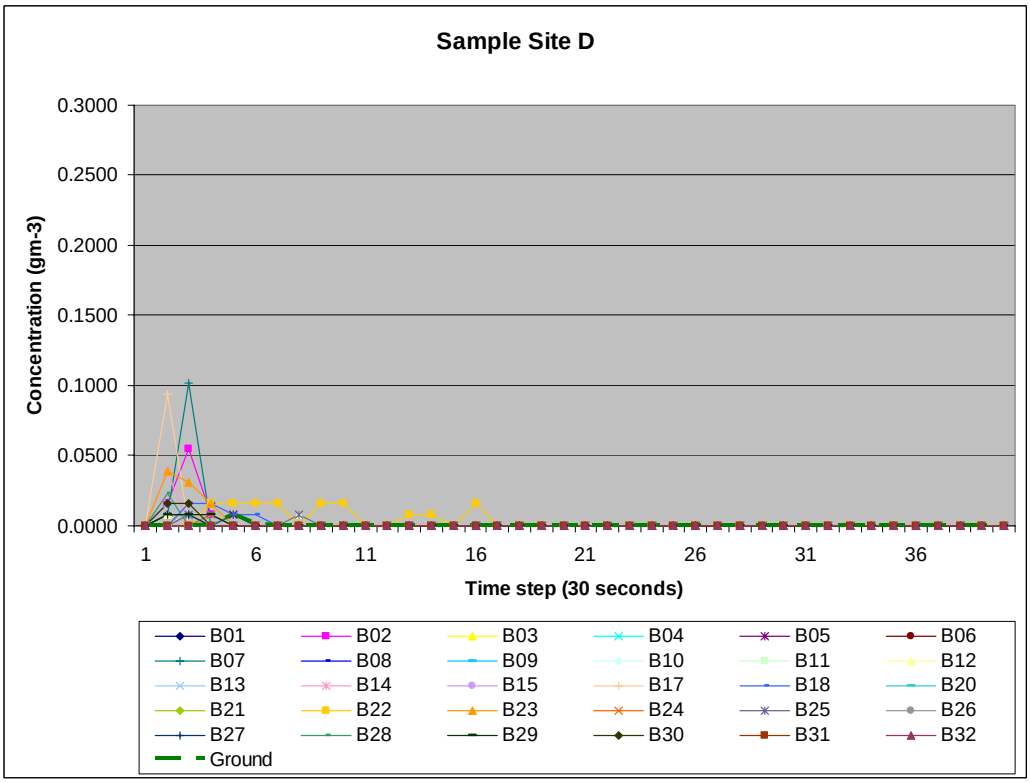
Group A simulations



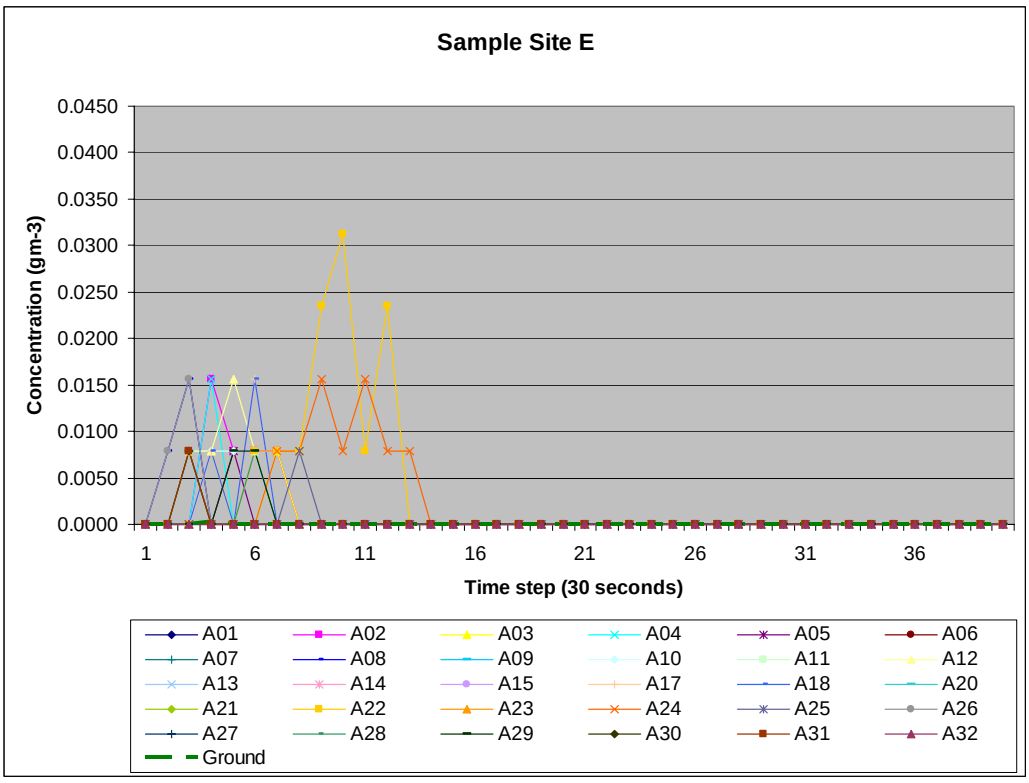
Group B simulations



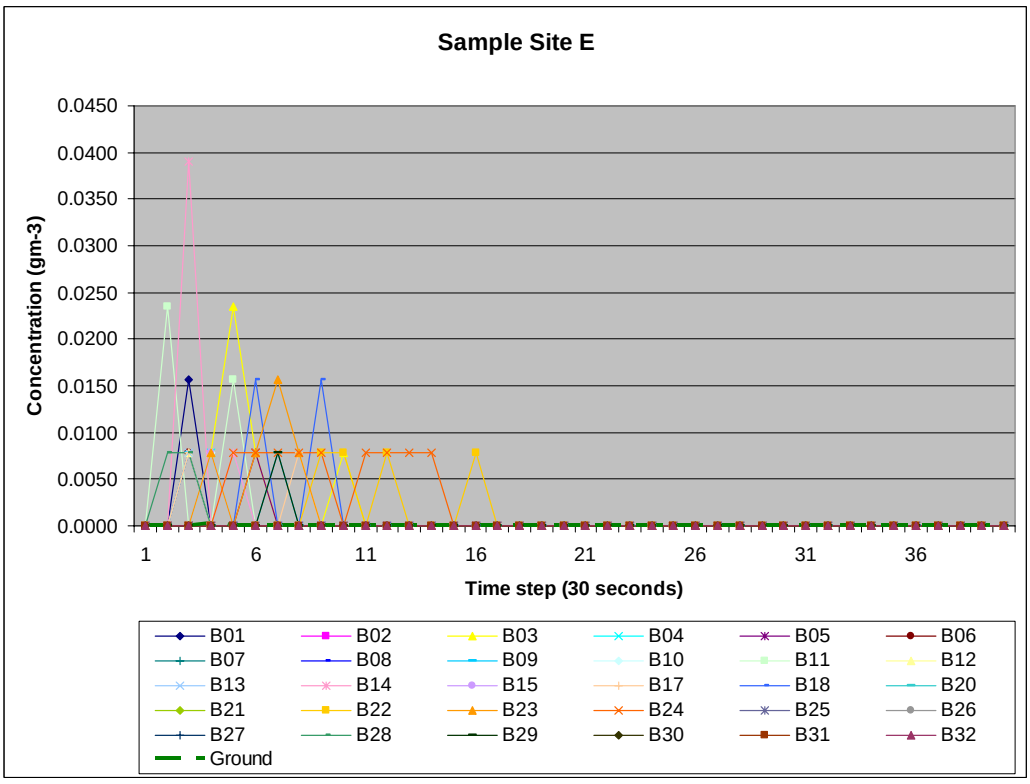
Group A simulations



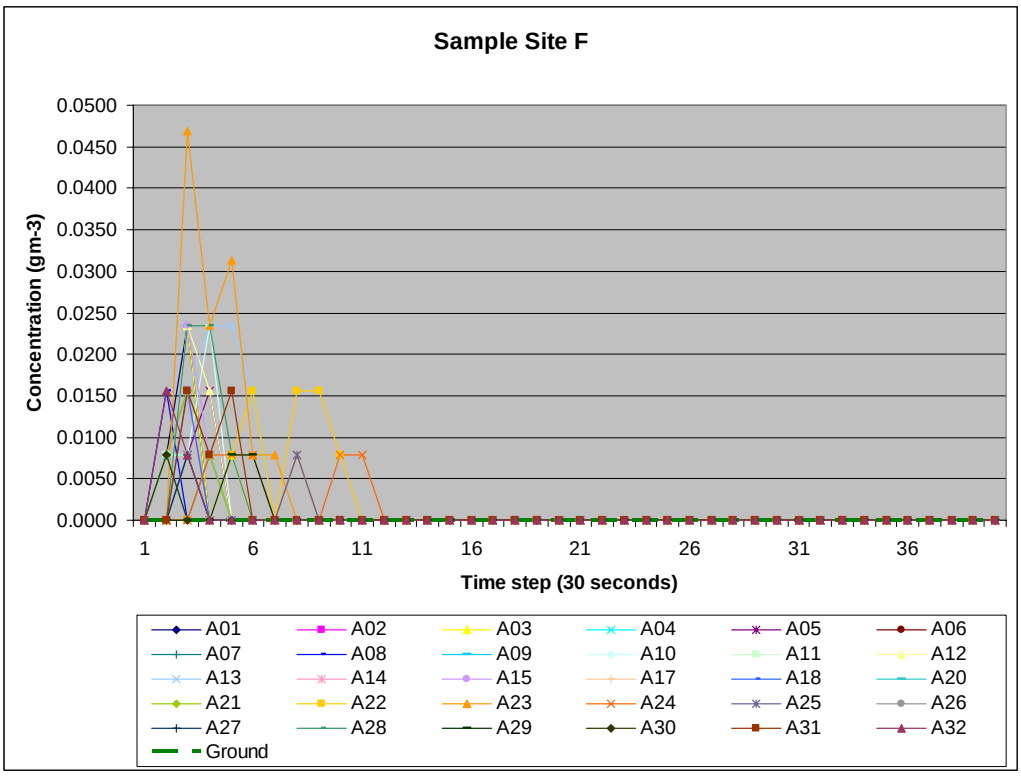
Group B simulations



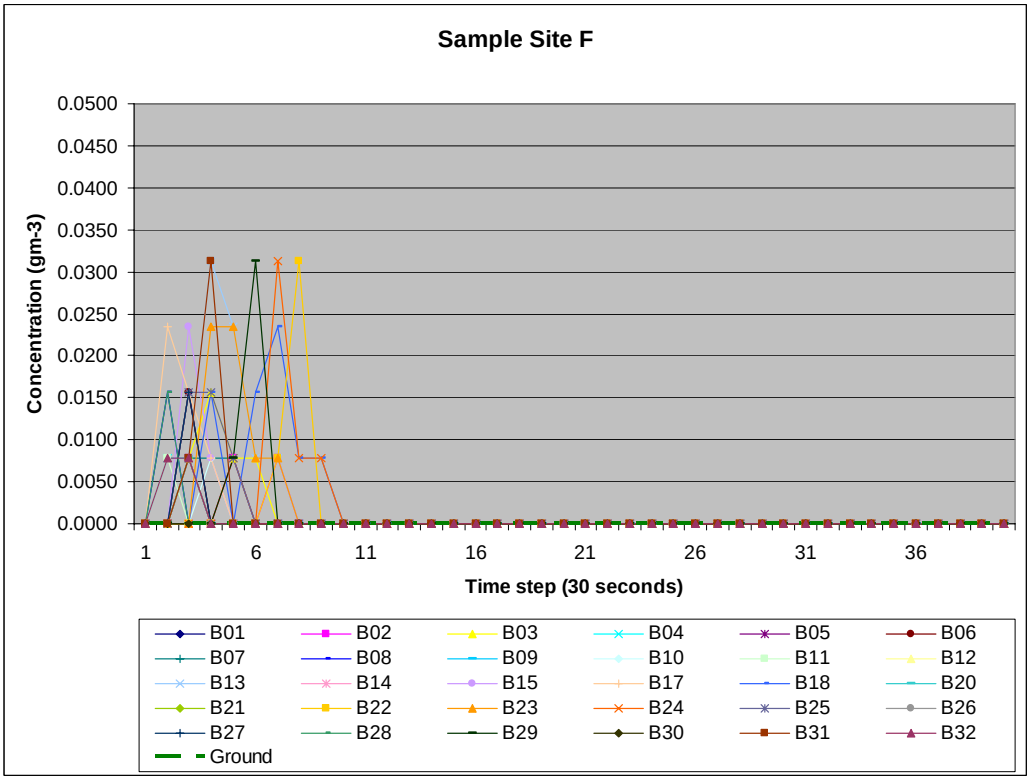
Group A simulations



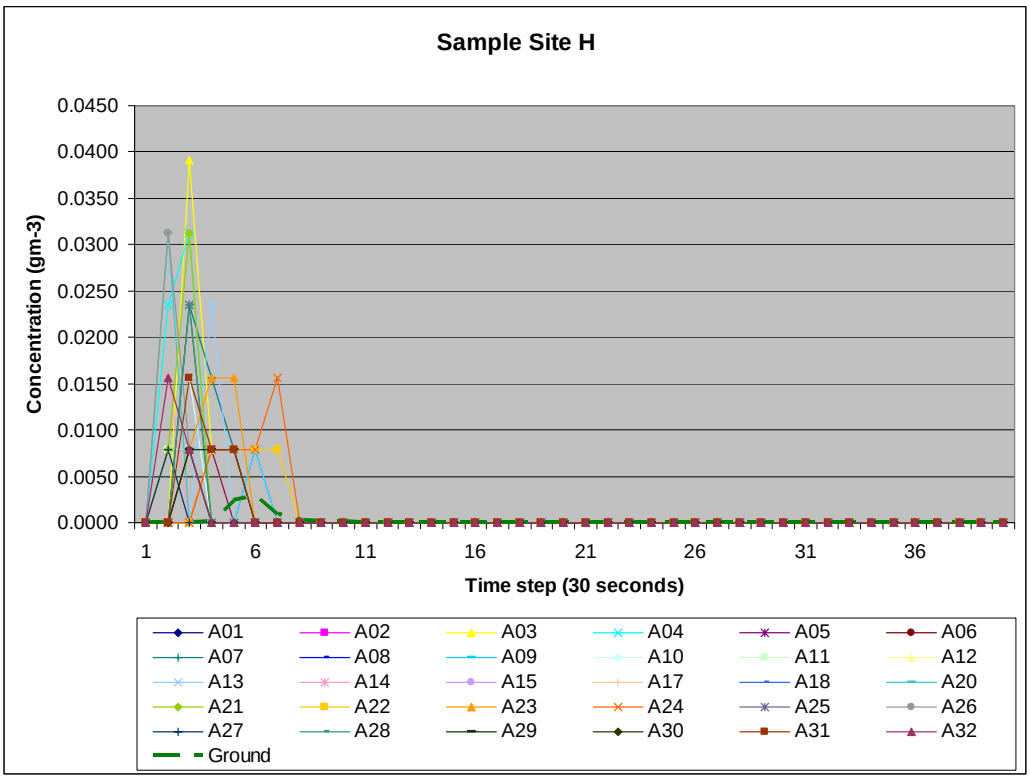
Group B simulations



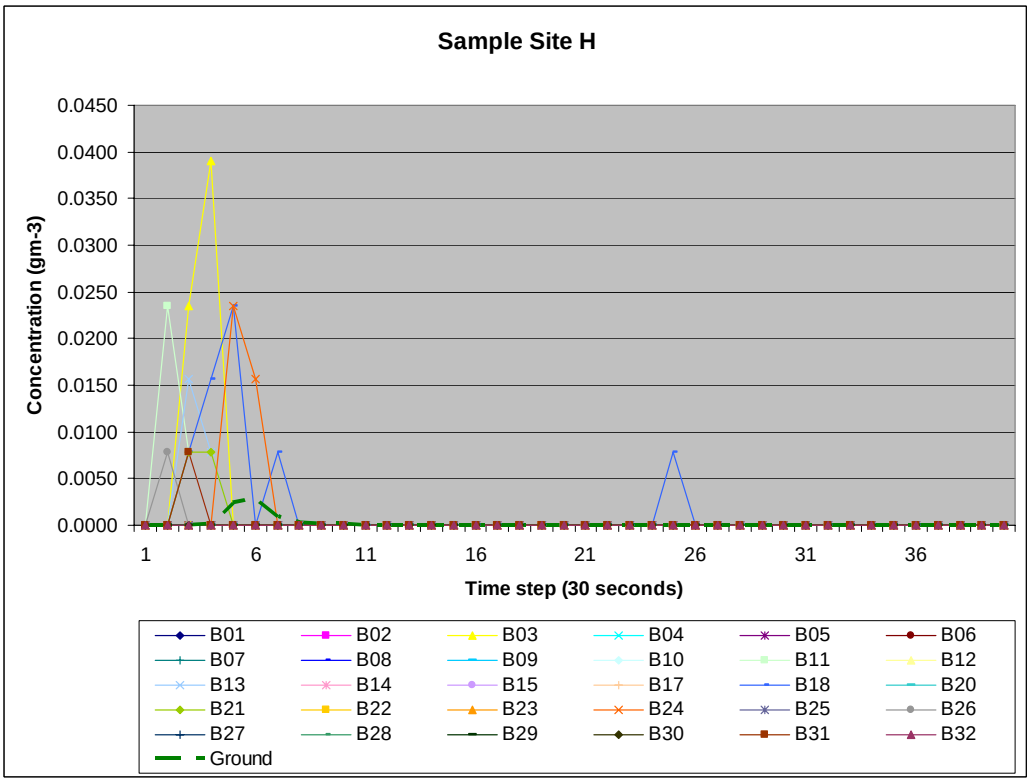
Group A simulations



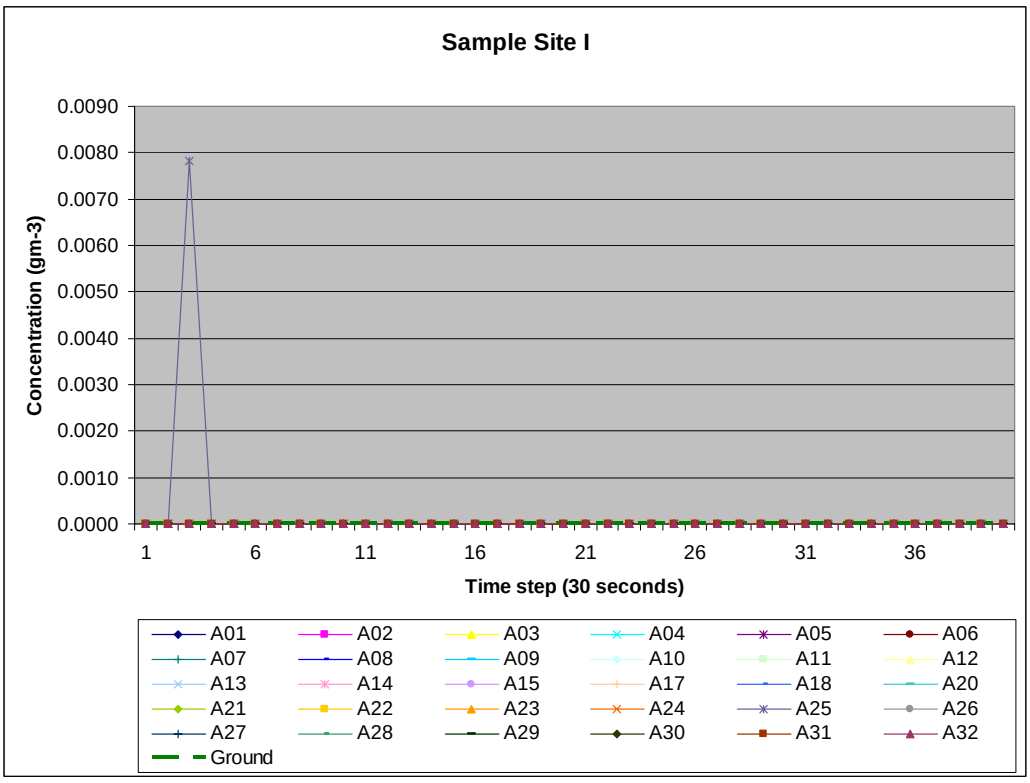
Group B simulations



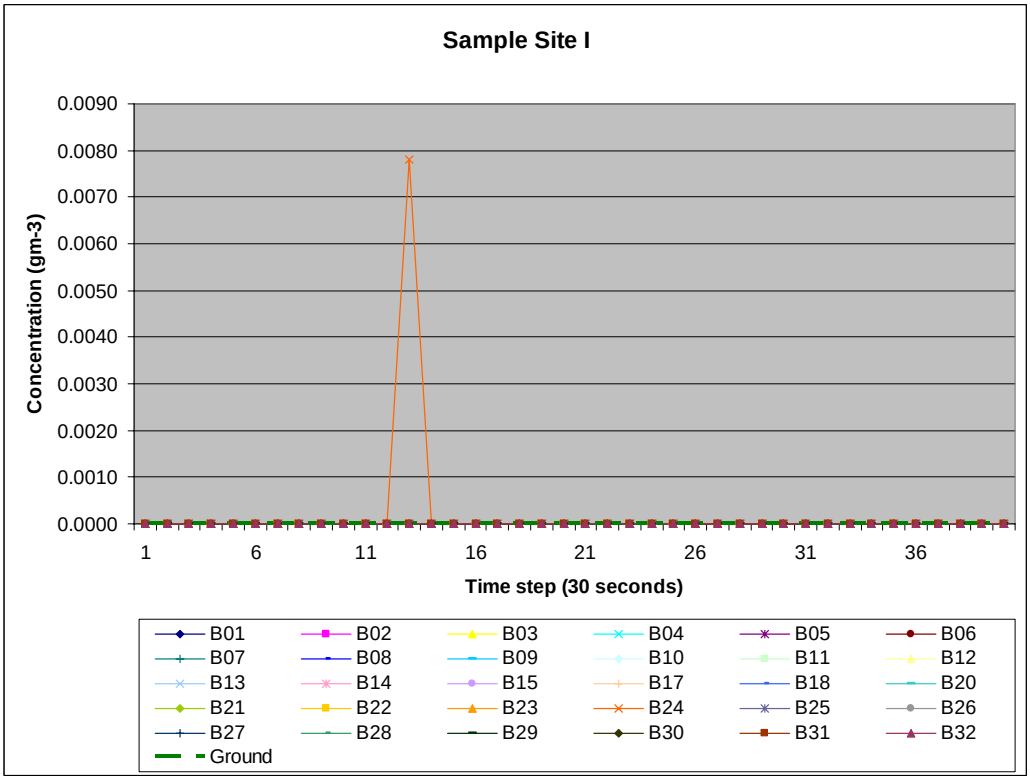
Group A simulations



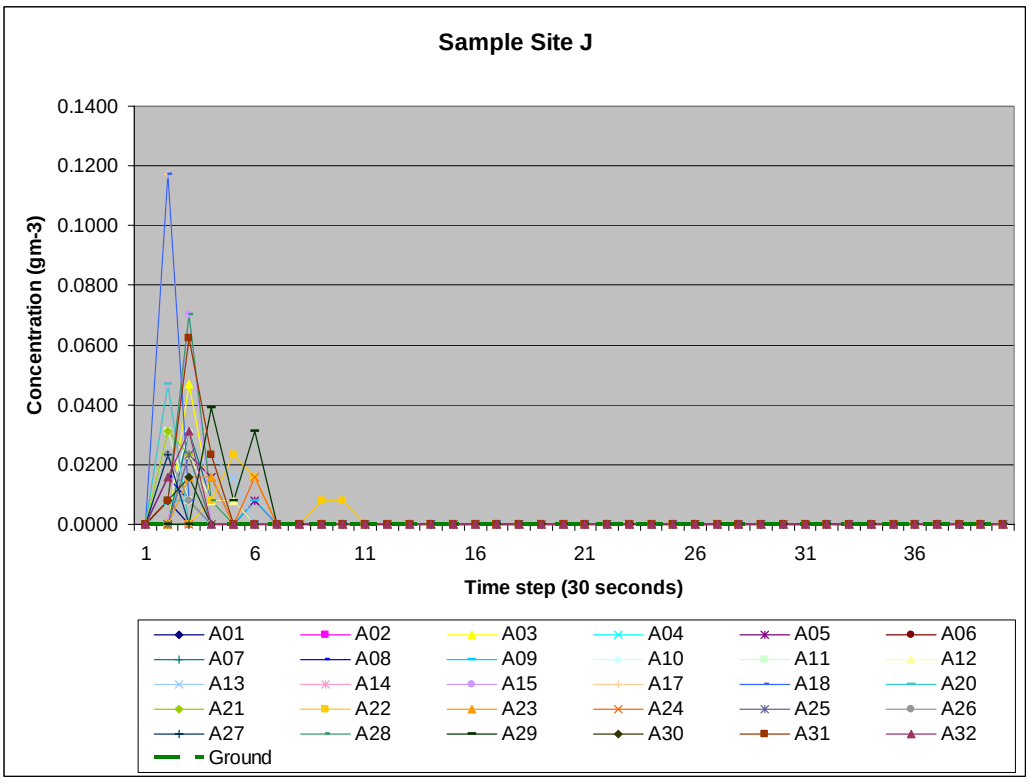
Group B simulations



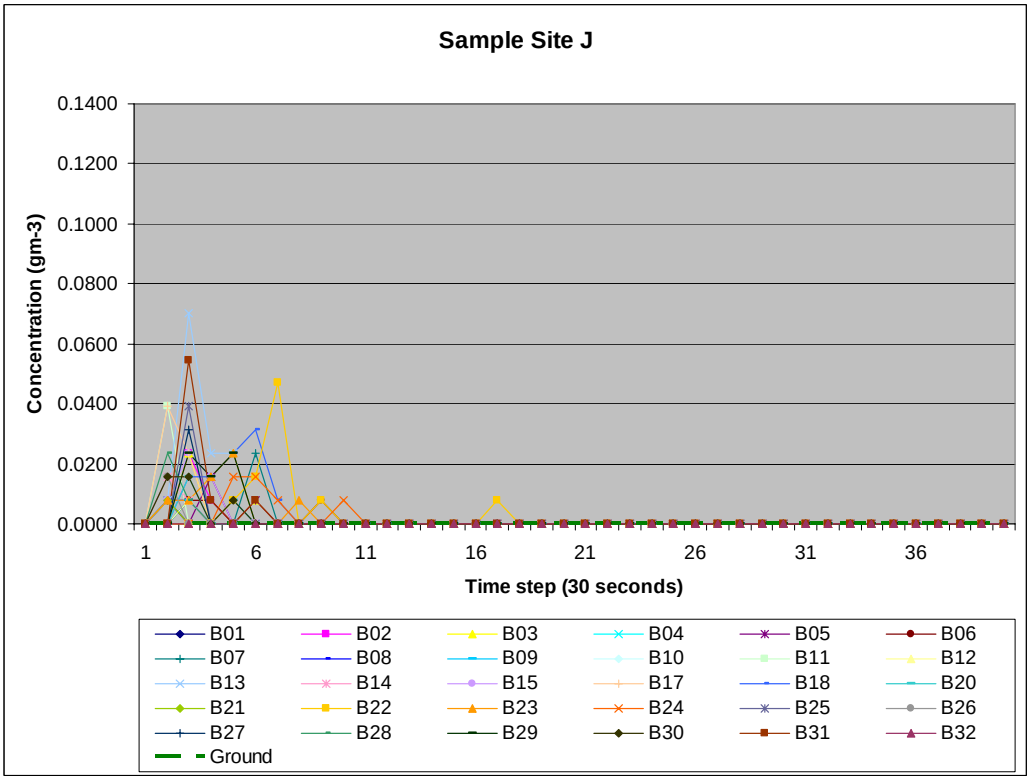
Group A simulations



Group B simulations



Group A simulations



Group B simulations